Predicting Obstetric Outcome Through a Web Application Using a Multinomial Logistic Regression Model

Márcia Oliveira[1(✉)], Ana Cristina Braga[2], and Rosete Nogueira[3,4]

[1] School of Engineering, University of Minho, 4710-057 Braga, Portugal
marcia27999@gmail.com
[2] ALGORITMI Research Centre, LASI, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal
[3] Life and Health Sciences Research Institute (ICVS), ICVS/3B's - PT Government Associate Laboratory, Campus de Gualtar, 4710-057 Braga/Guimarães, Portugal
[4] CGC Genetics, Unilabs, Embryo-fetal Pathology Laboratory, R. Sá da Bandeira, 706, 5o, 4000-432 Porto, Portugal

Abstract. Accurately predicting obstetric outcome is a critical challenge in maternal-fetal medicine, with significant implications for improving healthcare practices and minimizing adverse events. A dataset comprising fetoplacental biometric features was used, with the response variable stratified into three distinct categories: newborn, intrauterine fetal death, and neonatal death. In this context, the present paper aimed to develop an application capable of predicting one of the three possible obstetric outcomes. Multinomial logistic regression models are commonly used to identify the most influential predictors, as they enable the modeling of relationships between independent variables and a categorical dependent variable with more than two categories. In that regard, within the scope of model determination, various variable selection techniques were implemented, with the *Elastic net* method, using an alpha of 0.1, showing the best performance. For the intrauterine fetal death category, the model identified four significant variables, achieving an AUC-ROC of 0.841, with female fetal gender emerging as the most contributive factor. Regarding the neonatal death category, seven relevant variables were selected, with the model yielding an AUC-ROC of 0.620. The most impactful variables in this category were Diameter2, male gender, and maternal age. Following this stage, the chosen model, integrating the multinomial logistic regression algorithm with the *Elastic net* technique, was employed in the development of a web application, which is available at https://obstetricoutcome.shinyapps.io/mo_uminho/. This platform can be a significant contribution to obstetric practice, providing an intuitive tool for healthcare professionals.

Keywords: Obstetric outcome · Multinomial logistic regression · Variable selection · Application

This work has been supported by FCT Fundação para a Ciência e Tecnologia within the R&D Units Project Scope: UID/00319/2023.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2026
O. Gervasi et al. (Eds.): ICCSA 2025 Workshops, LNCS 15887, pp. 269–286, 2026.
https://doi.org/10.1007/978-3-031-97589-9_19

1 Introduction

Obstetrics plays a crucial role in maternal and fetal health, significantly impacting pregnancy outcomes [17]. The ability to accurately predict obstetric outcomes is vital for providing personalized medical care during gestation, minimizing adverse effects, and optimizing healthcare delivery [30].

Despite decades of intensive research in this area, predicting obstetric outcomes and associated complications, such as intrauterine growth restriction, preeclampsia, and stillbirth, remains a significant medical challenge [11]. Furthermore, placenta-mediated pregnancy complications are among the leading causes of maternal, fetal, and neonatal morbidity and mortality in developed countries, with increased risks of recurrence in subsequent pregnancies [20].

Placenta serves as the gateway and it is responsible for providing vital nutrients and oxygen supply to the fetus, while also produces hormones for maintaining pregnancy. It is essential to ensure that the fetus receives the appropriate supplement, being crucial for promoting optimal fetal growth. A successful pregnancy depends on a well-functioning placenta, and any compromise in its function has been suggested as a potential factor contributing to fetal demise. It is believed that a key determinant of placental function is its surface area, which facilitates the exchange of respiratory gases and nutrients necessary for optimal growth and well-being of the fetus during the gestational period [6,18,31]. Impaired placental function can lead to inadequate fetal growth, making it a focus of research aimed at understanding and mitigating these risks.

Over the past years, fetal and neonatal measurements have been widely used as indicators of fetal development. More recently, however, there has been growing interest in exploring placental biometric parameters and their relationship with obstetric outcomes [22,27]. Among these, placental weight (PW) has emerged as a potential indicator of placental function, reflecting both maternal and fetal environments, and varying in response to different pregnancy-related circumstances [18,24].

A linear relationship between fetal weight (FW) and PW has been documented in humans, and the Fetal Weight/Placental Weight Ratio (FPW-R) has garnered increasing interest on obstetricians, since it can indicate some conditions associated with certain placental disorders. Deviations in FPW-R have been associated with growth restriction, perinatal death, non-reassuring fetal status, and low Apgar scores. Any deviation in these processes can compromise the placenta's ability to support healthy fetal growth, highlighting its role as a focus of clinical and scientific research aimed at preventing potential adverse pregnancy outcomes. These findings emphasize the importance of understanding fetoplacental biometric features in preventing complications and improving pregnancy results, thereby enabling more precise and effective medical interventions [24,31].

This study aimed to explore fetoplacental biometric data during the viability period of pregnancy, with the goal of understanding the obstetric outcome. This period was specifically chosen as it corresponds to the stage of pregnancy when the fetus has reached a level of development sufficient to survive

outside the womb after birth. The final purpose was to develop a web application that enables professionals to obtain real-time estimates of obstetric outcome. To achieve this, multinomial logistic regression (MLR) techniques were employed for predictive modeling.

The structure of this paper is organized into four sections. Section 2 provides a comprehensive overview of the dataset utilized, accompanied by the preprocessing steps applied, the study variables and the statistical analysis conducted. Section 3 presents the obtained results and offers a critical analysis of the findings, highlighting their implications and relevance in this context. Lastly, Sect. 4 summarizes the main conclusions drawn from the study and emphasizes potential directions for future research.

2 Material and Methods

2.1 Dataset

The data utilized in this study comprises fetoplacental biometric features collected at Embryofetal and Placental Pathology Laboratory, situated at CGC Genetics, in Unilabs, Porto, Portugal. Parameters were obtained from autopsy reports of placentas and fetuses collected over an eleven-year period, 2014 to 2024. These reports covered a wide range of gestational age (GA), from 12 to 41 weeks, and focused on the portuguese population. The initial sample comprised 7321 histopathological reports, which underwent rigorous preprocessing as part of a previous study conducted in 2019 by Nogueira *et al.* [28]. This procedure refined the dataset to 1951 samples. Subsequently, a study carried out in 2023 by Lemos included additional cases, adhering to the initially defined criteria, resulting in an expanded dataset of 2092 reports [23]. For the present study, further cases were added, culminating in a final dataset of 2278 samples used to determine the model that aims to predict the obstetric outcome.Considering the study's scope, 12 of the 25 available variables were selected based on their contextual significance and research objectives. These included "local", "Processnumber" (hospital process number), "MaternalAge" (mother's age), "GA" (gestational age), "Fetalweight" (fetal weight), "Fetalgender" (fetal gender), "Pregnancyoutcome" (pregnancy outcome), "Placentalweight" (placental weight), "Diameter1" (largest diameter of the placenta), "Diameter2" (smallest diameter of the placenta), "Placentalthickness" (placental thickness) and"ratio" (FPW-R).

2.2 Data Preprocessing

A comprehensive dataset preprocessing process emerges as the initial and crucial step in implementing a predictive model. This phase involves applying preparation and transformation techniques that precede statistical analyses, resulting in a dataset that is ready for analysis. In this study, all stages were performed using R, allowing for efficient data handling and transformation. The dataset underwent two more filtering steps to align with the criteria established for this work. In that regard, the study focused on the phase corresponding to the

period of viability. Consequently, only samples with GA $\geq$ 22 weeks were selected. Also to ensure the validity and reliability of the results, samples categorized as medical termination of pregnancy under the "Pregnancyoutcome" variable were excluded. After this, the "local" and "Processnumber" columns were removed to optimize subsequent analysis. This approach helps prevents potential biases and distortions in the interpreting of results.

Data Cleaning. One of the key tasks carried out was data cleaning, which involves the process of identifying and correcting missing or incorrect values in the raw data. The number of missing values was assessed using the `is.na` function, resulting in the identification of 18 missing entries (1.3% of the total dataset). Additionally, the `missmap` function from the *Amelia* package was employed to visualize missing data patterns. The presence of duplicate observations can introduce significant issues in statistical analysis and modeling and therefore, identifying and removing such data is paramount. The `duplicated` function was applied to detect these values, and the `sum` function was subsequently utilized to quantify the number of duplicate rows identified. A total of 5 duplicate rows were found and eliminated.

Correlation. Analyzing correlation effects between variables can significantly influence machine learning modeling in several ways, particularly in the context of feature selection. When two variables exhibit a high degree of correlation, it may be redundant to include them together in the model, as one of them may provide similar information to the other. Consequently, a correlation analysis was conducted among key variables, such as FW, PW, and FPW-R. Based on the results, FW and PW were excluded from the model, as their strong correlation could introduce unnecessary redundancy.

Data Transformation. Given the presence of numerical variables with divergent scales in the database, a direct comparison between them becomes challenging. To mitigate this disparity, a data normalization technique was applied, specifically Min-Max scaling, which rescales all numerical variables to a uniform range of 0 to 1. The values are adjusted according to the minimum and maximum observed value, using the transformation described in Eq. 1:

$$z = \frac{x - \min}{\max - \min} \tag{1}$$

Here, z is the normalized value, while x is the original value. The terms min and max represent the minimum and maximum observed values, respectively.

Proper encoding of categorical variables is crucial when applying MLR to maintain the accuracy and interpretability of the model. To achieve this, dummy variables are created to represent the distinct categories of a categorical variable. By creating only $k - 1$ dummies for k categories, we ensure that the variables are not linearly dependent on each other.

2.3 Study Variables: Predictors and Outcomes

After completing the data preprocessing and transformation steps, the resulting dataset is expected to be cleaner, more refined, free of missing values and standardized. Therefore, Table 1 presents the final dataset, outlining its attributes, the ranges of registered values for numerical variables, and the absolute frequency along with the corresponding codification for categorical variables. In this approach, the response variable ("outcome") was multicategorical, with the newborn as the reference category, while the remaining classes corresponded to intrauterine fetal death and neonatal death.

Table 1. Summary of variables, both numeric and categorical used in the study.

Numeric variable	Range/Codification	Frequency
MaternalAge	15.1 − 48.2 years	-
GA	22 - 41 weeks	-
Diameter1	5.5 - 32 cm	-
Diameter2	5 - 30 cm	-
Placentalthickness (PT)	0.8 − 13.5 cm	-
FPW-R	1.33 - 11.51	-
Categorical variable		
Fetalgender	0 - Female	496
	1 - Male	547
	2 - Unknown	277
Outcome	0 - Newborn	841
	1 - Intrauterine Fetal Death	453
	2 - Neonatal Death	26

2.4 Statistical Analysis

Machine learning algorithms are used to extract knowledge from diverse data and generate effective prediction and classification models [5]. Regression analysis is a statistical tool that examines the relationship between two or more variables and the logistic regression model is important to analyze categorical data [9]. MLR can be used to predict the categorical placement or the probability of belonging to a category on a dependent variable, considering several independent variables [21]. With this concept in mind, in this paper an MLR model was used to aid in the prediction of the obstetric outcome, taking into account the 7 independent variables and the response variable.

In a model that considers a categorical response variable with multiple categories (nominal), where Y is the response variable and J represents the different categories it can take, the probability of Y taking each of the J categories is

denoted by $\pi_j(x)$, where $\sum \pi_j(x) = 1$. The model described by Eq. 2 considers the logarithm of the odds ratio (OR) $\pi_j(x)/\pi_J(x)$, where $\pi_J(x)$ is the probability of the reference category. The model simultaneously describes the effects of the explanatory variables x on these J - 1 logits. The J - 1 equations determine the parameters for the logits with other pairs of response categories. In this context, in a MLR model, the logit can be expressed by the following Eq. 2 [1]:

$$\log \frac{\pi_j(x)}{\pi_J(x)} = \alpha_j + \beta_j x, \quad \text{for } j = 1, \ldots, J-1 \tag{2}$$

Here, $\pi_j(x)$ represents the probability of the response variable Y being assigned to category j, while $\pi_J(x)$ denotes the probability of Y belonging to the reference category J. The term α_j corresponds to the intercept for category j, and β_j is the vector of coefficients associated with the explanatory variables x.

Split the Dataset. A common practice in data analysis involves dividing the dataset into two distinct subsets: one for model training and the other for model evaluation [3,26]. In this approach, the training dataset was utilized to construct the model, while the testing dataset was reserved for assessing the model's predictive performance [26]. For this particular study, a split ratio of 0.7 was applied, which means that 70% of the data were allocated for training purposes, and the remaining 30% was designated for testing. Consequently, the resulting training dataset consisted of 589 samples from the newborn class, 318 samples from the intrauterine fetal death class, and 19 samples from the neonatal death class.

Balance the Dataset. The issue of unequal class distribution presents a notable challenge in many machine learning applications. This arises from the scarcity of training examples for the minority classes, which leads to models that tend to favor the prediction of the majority classes, due to the availability of abundant data for these classes [32]. To handle the significant class imbalance, the Synthetic Minority Oversampling Technique (SMOTE) was employed and evaluated in the training set. This approach generates synthetic samples through an oversampling approach, which was identified as the most effective method in the context of limited data, aiming to create samples in the minority classes to equalize the number of examples across different categories. It is important to highlight that these balancing techniques were applied exclusively to the training dataset, preventing biased or overestimated model performance. After applying SMOTE, the training set was expanded to 1496 samples: 589 newborn, 318 intrauterine fetal death, and 589 neonatal death instances.

Feature Selection. Variable selection is a critical and challenging phase in model construction, often considered the most intricate aspect of the process. This step involves removing redundant or irrelevant features, in order to select those that are best suited to optimize the model's performance. Additionally, it is crucial to prevent the inclusion of noise variables that may degrade the model's effectiveness [8]. However, identifying the most appropriate algorithm for this task is not straightforward as different algorithms apply distinct criteria to determine the most representative variables. Feature selection methods can

be categorized into two main approaches: wrapper, and embedded techniques [7].

- *Wrapper Method.* Incorporate a predictive model during training and evaluate variable subsets to optimize feature selection [7]. The effectiveness of these methods depends on the choice of classification algorithms and search strategies [15]. Commonly employed techniques include Sequential Forward Selection (SFS), an iterative approach that progressively adds variables in a step-by-step process, Sequential Backward Elimination (SBE), which removes variables iteratively starting from the full model, and bidirectional search, which combines the addition and removal of features in each iteration, beginning with an initial set of features [4,8,25,33].

In this modeling approach, the `multinom` function from the *nnet* package in R was employed to fit the initial model, specifically designed for MLR, which is well-suited for classification tasks involving response variables with more than two categories. Following this step, variable selection in MLR can be performed using one of three sequential search techniques, leveraging the *MASS* package and utilizing the fitted model as input.

- *Embedded Method.* The core principle of embedded techniques lies in introducing a penalization of the model coefficients, forcing some of them to approach zero or be exactly zero, gradually reducing their importance. This can lead to model simplification without the need to eliminate variables. Among the key regularization techniques are *Ridge*, Least Absolute Shrinkage and Selection Operator (*LASSO*), and *Elastic net* [2]. *Ridge* regularization applies an L_2 penalty, shrinking coefficient magnitudes proportionally toward zero without eliminating them. *LASSO* regularization, using a L_1 penalty, reduces some coefficients exactly to zero, selecting only the most relevant variables [13,29]. *Elastic net* combines *Ridge* (L_2) and *LASSO* (L_1) regularization, making it particularly advantageous in scenarios involving highly correlated variables [13].

In this approach, the *glmnet* and *caret* packages were employed to implement *Ridge*, *LASSO*, and *Elastic net* regularization techniques efficiently [13]. Within the *glmnet* framework in R, the parameter alpha controls the type of penalty applied in regression models. Specifically, alpha is set to 0 for *Ridge* regularization, equals 1 for *LASSO*, and takes values between 0 and 1 to represent *Elastic net*. This method fits models with different penalty values (lambda) and uses lambda as the key hyperparameter for regularizing model coefficients [29]. For tuning regularization, 10-fold cross-validation was applied to determine the optimal lambda value [10].

Assessment Metrics. To compare the performance of different predictive models in forecasting obstetric outcomes using the dataset variables, several assessment metrics were used to select the best model:

- *Accuracy:* It is defined as the proportion of correctly classified cases relative to the total number of instances in the dataset. Thus, the precision formula considers the rates of True Positive (TP) and True Negative (TN) [16,19].

- ***F1-Score:*** It provides a balanced relationship between precision and recall, where their relative contributions are equally weighted in the F1-Score. This formula takes into account both false positives (FP) and false negatives (FN) [16,19].
- ***Precision:*** It measures the proportion of instances that were correctly classified as positive relative to the total number of instances predicted as belonging to the positive class [19]. Precision indicates how much we can trust the model when it predicts a case as positive [16].
- ***Recall:*** Also known as sensitivity, represents the proportion of actual positive instances that were correctly identified by the model. The formula involves TP divided by the total number of units classified as positive [16,19].
- ***AUC-PR/AUC-ROC:*** Metrics such as Area Under a Receiver Operating Characteristic Curve (AUC-ROC) and Area Under a Precision-Recall Curve (AUC-PR) can be adapted for multiclass problems through two methodologies: One versus One, and One versus All [12,14]. The ROC curve summarizes the test's ability to discriminate between positive and negative classes across all possible threshold values, balancing trade-offs between sensitivity and specificity [34]. The PR curve can evaluate various types of classifiers across the entire spectrum of thresholds, which illustrates a trade-off between precision and recall. AUC-PR is particularly useful in datasets with class imbalance [14].

3 Results and Discussion

To better understand the distribution and behavior of these variables across obstetric outcomes, a detailed statistical analysis was conducted for each category. Additionally, Kruskal-Wallis and Chi-squared tests were applied to assess

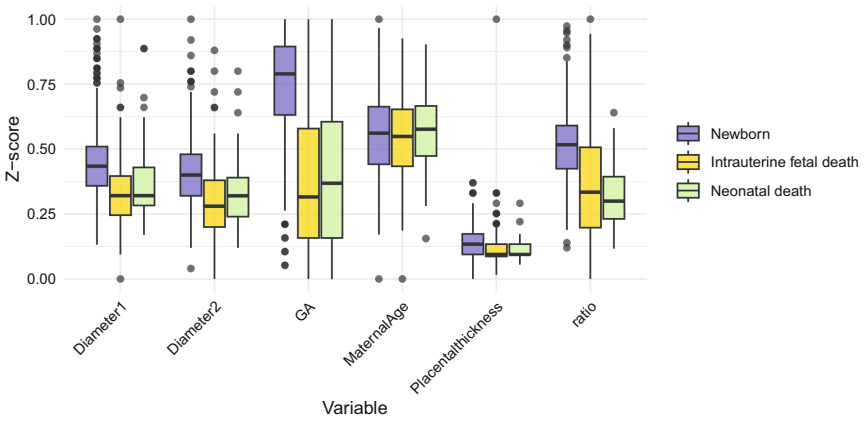

Fig. 1. Boxplots illustrating the distribution of numerical variables values across the three obstetric outcomes, highlighting the median, interquartile ranges, and potential outliers within each group.

differences among groups, with p-values below 0.05 indicating statistical significance. Variables such as GA, PT, placental diameters, FPW-R, and fetal gender showed significant differences among obstetric outcome classes.

Beyond statistical analysis, Fig. 1 provides a visual representation of the distribution of normalized numerical variables among the three outcome categories. The observed differences in GA, placental diameters, PT, and FPW-R also suggest potential trends that may be crucial for understanding their impact on obstetric results. In contrast, maternal age exhibits no marked differences, indicating a less pronounced impact when considered in isolation.

3.1 Model Selection

Despite the previously mentioned class imbalance, an MLR model was implemented using the unbalanced dataset, achieving an accuracy of 83%. While the model performs well for the majority class (newborn), it struggles significantly with the minority classes, particularly neonatal death. This issue stems from the unequal class distribution, resulting in a bias towards the majority class. Therefore, implementing a balancing strategy was crucial to addressing this class imbalance to improve the model's generalization across all categories.

Based on the above, after applying the SMOTE technique, the MLR model was trained, and predictions were made on the test data. Compared to the unbalanced training set, the balancing technique led to overall performance improvements, as shown in Table 2. It yielded superior results in most cases, particularly in precision, recall, AUC-PR, and AUC-ROC metrics for the neonatal death category. These findings highlight SMOTE as an effective technique for improving model performance when applied to the training dataset.

Table 2. Performance metrics of the MLR model for the unbalanced dataset and the SMOTE technique.

Class	Accuracy	F1	Precision	Recall	AUC-PR	AUC-ROC
	Unbalanced dataset					
Intrauterine fetal death	0.830	0.756	0.733	0.780	0.781	0.884
Neonatal death	0.830	NaN	0.000	NaN	0.020	0.571
	SMOTE					
Intrauterine fetal death	0.617	0.259	0.163	0.629	0.645	0.812
Neonatal death	0.617	0.037	0.286	0.020	0.022	0.600

Variable selection methods analyzed and compared in detail included the wrapper and embedded approaches. In this type of regression, one category of the response variable was designated as the reference, while the remaining categories were compared against it. In this framework, the coefficients associated with the reference category were set to zero, whereas the coefficients for the other classes

represented the logarithm of the OR of belonging to each category relative to the reference.

In this context of imbalanced datasets, the AUC-ROC metric was particularly relevant, as it plays a crucial role in discriminating between classes and should be considered. To compare various MLR models with different subsets of variables, a reference model (designated as model 0) was used, which included all seven available variables without any selection technique.

In relation to the wrapper method, as observed in the analysis of Table 3, the models developed through sequential search did not show improvements in predictive performance compared to the reference model. Specifically, none of the three methods demonstrated any improvement, yielding the same or even lower estimates than those obtained with the model without variable selection. It was worth noting that none of these methods reduced the number of attributes used in the process. Therefore, the application of these methods did not bring noticeable benefits to the model.

Table 3. Performance metrics obtained after applying wrapper methods for variable selection in MLR.

Model	Accuracy	F1	Precision	Recall	AUC-PR	AUC-ROC	Variables
0	Method without variable selection						
Intrauterine fetal death	0.617	0.259	0.163	0.629	0.645	0.812	7
Neonatal death	0.617	0.037	0.286	0.020	0.022	0.600	7
1	**Forward**						
Intrauterine fetal death	0.612	0.249	0.156	0.618	0.645	0.811	7
Neonatal death	0.612	0.036	0.286	0.019	0.022	0.601	7
2	**Backward**						
Intrauterine fetal death	0.612	0.249	0.156	0.618	0.645	0.811	7
Neonatal death	0.612	0.036	0.286	0.019	0.022	0.601	7
3	**Bidirectional**						
Intrauterine fetal death	0.612	0.249	0.156	0.618	0.645	0.811	7
Neonatal death	0.612	0.036	0.286	0.019	0.022	0.601	7

Concerning the embedded method, about the results obtained for the different regularization techniques using the dataset obtained through SMOTE, as shown in Table 4. It was noteworthy that the *Ridge* regularization model resulted in a total number of variables (7), which was expected given its more generalized nature aimed at avoiding overfitting. The *LASSO* regularization approach using the minimum lambda (lambda min) showed a reduction in the number of variables, with estimates like those of the reference model without variable selection. *LASSO* regularization with the one standard error lambda (lambda 1se) also significantly reduced the number of variables and showed improvements in some metrics compared to the reference model.

For *Elastic net* regularization, different alpha values were tested. This was an important aspect since low alpha values retain more variables, offering a more comprehensive model, though potentially more complex. In contrast, higher alpha values result in a potentially simpler model with fewer variables, which can better handle multicollinearity and focus on the most relevant variables. Models with alpha values of 0.1 and 0.9 showed better results compared to others, with higher values for AUC-PR and AUC-ROC metrics. Although model 8 demonstrated similar results to model 7, the latter was chosen due to its superior performance in evaluation metrics, particularly in AUC-PR and AUC-ROC values. This was a decisive criterion for the decision.

Table 4. Performance metrics obtained after applying regularization techniques for variable selection in MLR.

series Model	series Accuracy	series F1	series Precision	series Recall	series AUC-PR	series AUC-ROC	series Variables
0	**Method without variable selection**						
Intrauterine fetal death	0.617	0.259	0.163	0.629	0.645	0.812	7
Neonatal death	0.617	0.037	0.286	0.020	0.022	0.600	7
4	**Ridge Regularization**						
Intrauterine fetal death	0.594	0.085	0.044	0.857	0.720	0.846	7
Neonatal death	0.594	0.062	0.571	0.033	0.023	0.622	7
5	**LASSO Regularization (lambda min)**						
Intrauterine fetal death	0.609	0.229	0.141	0.613	0.652	0.816	2
Neonatal death	0.609	0.036	0.286	0.019	0.021	0.597	6
6	**LASSO Regularization (lambda 1se)**						
Intrauterine fetal death	0.596	0.170	0.096	0.722	0.683	0.840	1
Neonatal death	0.596	0.063	0.571	0.034	0.022	0.603	4
7	**Elastic net Regularization (alpha = 0.1)**						
Intrauterine fetal death	0.602	0.098	0.052	0.875	0.715	0.841	4
Neonatal death	0.602	0.064	0.571	0.034	0.023	0.620	7
8	**Elastic net Regularization (alpha = 0.9)**						
Intrauterine fetal death	0.612	0.205	0.119	0.762	0.680	0.838	2
Neonatal death	0.612	0.067	0.571	0.035	0.022	0.601	5

In this regard, the chosen model showed a relatively reasonable AUC-ROC value, with scores of 0.841 for intrauterine fetal death and 0.620 for neonatal death. On the other hand, the AUC-PR metric yielded values of 0.715 and 0.023 for intrauterine fetal death and neonatal death, respectively. This means that the predictive model performed reasonably well in identifying true positives for the first class, demonstrating high sensitivity. However, it showed low precision. Furthermore, the model faced significant difficulty in classifying cases of neonatal death, revealing both poor sensitivity and a limited ability to correctly distinguish this specific class.

The results could be attributed to the low proportion of the minority classes, which negatively impacts the model's performance. Out of the 1 320 samples available in the records, only 24.09% correspond to intrauterine fetal death, and 1.44% pertain to neonatal death. In turn, the similarity in performance metrics between wrapper methods and regularization techniques (Tables 3 and 4) suggests that both strategies extract a similar level of predictive signal from

the variables. This may indicate that the most informative features are consistently selected across methods, or that the overall signal is weak and redundant, thereby reducing the effectiveness of more complex selection strategies. While regularization approaches (e.g., *LASSO* and *Elastic net*) simplified the models by selecting fewer variables, this did not consistently translate into improved predictive performance, possibly due to inherent data limitations such as low signal-to-noise ratio and class imbalance.

3.2 Variable Importance and Estimated Coefficients

Intrauterine Fetal Death. The *Elastic net* method with alpha = 0.1 in the MLR classifier selected four key variables. As illustrated in Fig. 2, fetal female sex category was found to be the most contributive to intrauterine fetal death, exhibiting the highest positive coefficient. This suggests a strong association between female fetuses and an increased risk of intrauterine fetal death compared to newborns, potentially due to biological differences between the genders where hormonal and developmental factors could make female fetuses more susceptible to adverse conditions. Conversely, Diameter2, Diameter1, and GA play a protective role, reducing the likelihood of this outcome.

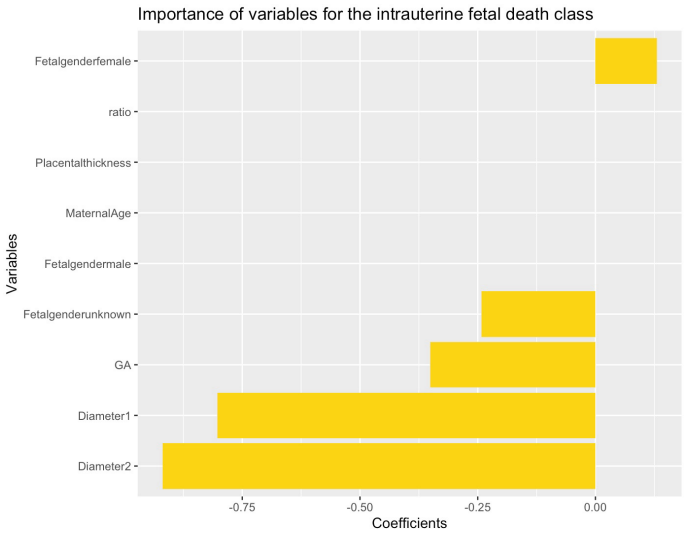

Fig. 2. Graph representing the importance of variables based on their respective coefficients for the intrauterine fetal death class, obtained through variable selection using the *Elastic net* regularization method with alpha=0.1.

Table 5 reveals that female infants have 1.140 times higher odds of intrauterine fetal death. In contrast, each additional unit of GA decreases this probability by 29.6. Similarly, increases in Diameter1 and Diameter2 lower the risk by 55.2%

and 60.1%, respectively. These findings emphasize the differential impact of these predictors on intrauterine fetal death, offering valuable insights into obstetric risk assessment.

Table 5. Demonstration of the estimated coefficients for the independent variables in the selected model and the respective OR for the intrauterine fetal death category.

series Variables	series Estimated coefficients	series OR
Intercept	0.582	1.789
GA	−0.351	0.704
Fetalgender (female)	0.131	1.140
Fetalgender (unknown)	−0.242	0.785
Diameter1	−0.803	0.448
Diameter2	−0.920	0.399

Neonatal Death. Seven original variables were identified for this category, as illustrated in Fig. 3. The most contributory factors for neonatal death include Diameter2, male sex, and maternal age. Among these, Diameter2 emerges as the most influential predictor, with higher values strongly associated with an increased likelihood of neonatal death compared to newborn. On the other hand, FPW-R and GA exhibit the strongest negative coefficients, suggesting that greater values significantly reduce the probability of this type of death. These results highlight potential fetal development concerns linked to Diameter2, the increased vulnerability of male infants during critical neonatal stages, and the increased risk of obstetric complications associated with advanced maternal age.

The results presented in Table 6 show that each additional year of maternal age increases the probability of neonatal death by 21.2%. Similarly, an OR of 1.241 suggests that male sex is associated with a 24.1% higher probability of neonatal death. An increase of one unit in Diameter1 and Diameter2 raises the risk of neonatal death by 1.098 and 1.666 times, respectively, compared to the newborn category. In contrast, a one-unit increase in GA reduces the probability of neonatal death by approximately 60%, while higher PT values reduce the risk by 0.524 times relative to the reference class. Furthermore, an increase in FPW-R dramatically decreases the probability of neonatal death by 84.4%, highlighting its strong protective effect. These observations provide valuable insights into both risk and protective factors for neonatal mortality, offering critical information for potential clinical interventions.

3.3 Web Application

After selecting the best MLR model (model 7) for predicting obstetric outcome, a web application was developed using the *shiny* package. Aligned with the study's

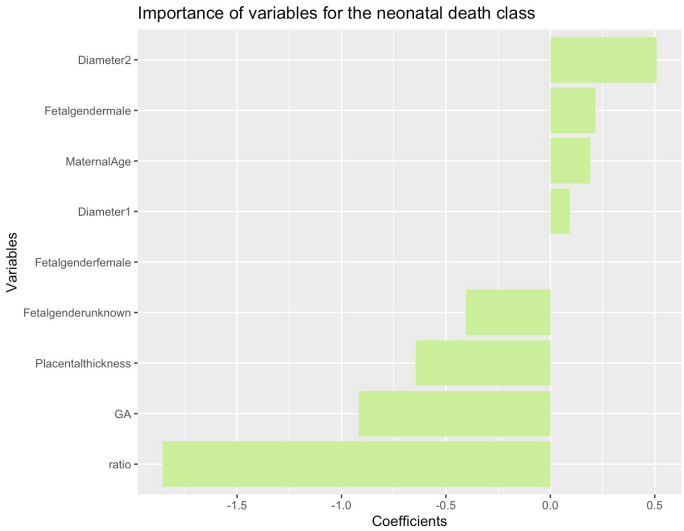

Fig. 3. Graph illustrating the importance of variables based on the coefficients obtained through *Elastic net* regularization with alpha = 0.1, in which the variables were selected to predict the neonatal death class.

objectives, this tool enables healthcare professionals to estimate fetal outcomes based on fetoplacental biometric features.

The "ObstetricOutcome" application has carefully designed to provide an intuitive and accessible experience, catering to the needs of healthcare professionals and other users. Its interface has been crafted to be user-friendly and easy to navigate, allowing for efficient interaction with the available tools and resources. The application enables real-time access to obstetric outcome predictions by filling in the required fields in the designated form, estimating the probability of each possible outcome. The platform is structured into four main sections - Home, About, Form, and Results - accessible through a navigation bar, each offering distinct functionalities. It is available at the following link: https://obstetricoutcome.shinyapps.io/mo_uminho/.

Hypothetical Case Study. To illustrate the application developed for predicting obstetric outcomes, consider the case of a 20-year-old pregnant woman in her third trimester, at 31 weeks of gestation. During a consultation, she presented the following clinical data: Diameter1 of 10 cm, Diameter2 of 8 cm, PT of 2 cm, FW of 440 g, PW of 120 g, and a female fetal gender. This information about the pregnant woman, the fetus, and the pregnancy was entered into the corresponding fields of the form available in the application. After submitting the data, the application generated a prediction of neonatal death (Fig. 4). Based on this result, the medical team can assess the implications of the outcome, emphasizing the importance of regular monitoring and early interventions, if necessary. This

Table 6. Presentation of the estimated coefficients for the predictor variables in the selected model, along with the corresponding OR for the neonatal death category.

series Variables	series Estimated coefficients	series OR
Intercept	1.111	3.037
MaternalAge	0.192	1.212
GA	−0.917	0.400
Fetalgender (male)	0.216	1.241
Fetalgender (unknown)	−0.406	0.666
Diameter1	0.094	1.098
Diameter2	0.511	1.666
Placentalthickness	−0.646	0.524
FPW-R	−1.857	0.156

approach allows better planning and the implementation of measures aimed at improving the health and well-being of both the mother and the newborn.

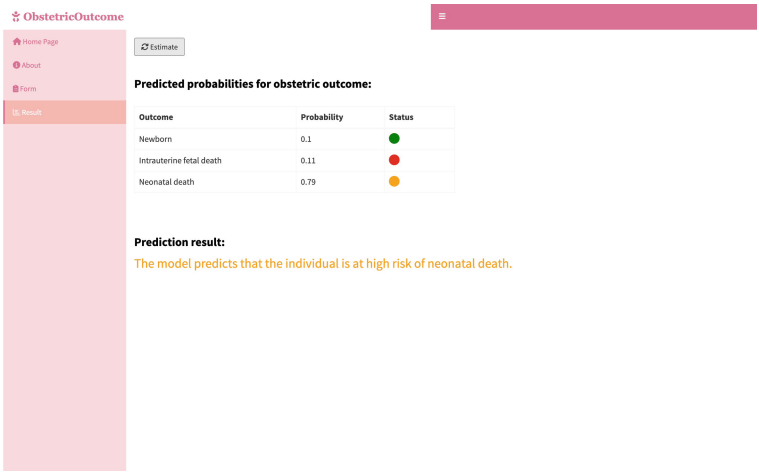

Fig. 4. Example of the obstetric outcome prediction result based on the data entered by the user in the hypothetical case.

4 Conclusion

Throughout this study, a comprehensive approach was employed to predict three different obstetric outcomes based on placental and fetal parameters. To carry

out this process, MLR models were implemented to identify the most contributive variables in determining fetal health across different gestational scenarios. Furthermore, the use of R packages such as *nnet* and *glmnet* demonstrated the versatility of the R environment in developing robust predictive solutions.

Although the classifier demonstrates reasonable predictive capability, its performance is constrained by the limited number of cases in certain outcome categories. It is hypothesized that a substantial increase in records for minority classes, particularly for neonatal death, would improve model's reliability. This could improve classification precision, strengthen predictive robustness, and enhance overall model effectiveness.

From a future perspective, expanding the dataset with additional maternal and fetoplacental biometric records could significantly improve the model's ability. Moreover, exploring alternative feature selection techniques and diverse machine learning algorithms, such as Random Forest or Support Vector Machines, could be an alternative to provide valuable insights for optimizing predictive outcomes.

In conclusion, this research could significantly contribute to the medical community in the field of obstetrics, offering support to assist healthcare professionals in making critical real-time decisions. They have great potential to facilitate the early detection of complications related to maternal and fetal health, while also transforming care into a more proactive and personalized approach. This could lead to more effective interventions, promoting a healthy and safe pregnancy.

References

1. Agresti, A.: Categorical Data Analysis, vol. 792. John Wiley & Sons (2012)
2. Altelbany, S.: Evaluation of ridge, elastic net and lasso regression methods in precedence of multicollinearity problem: a simulation study. J. Appl. Econ. Bus. Studies **5**(1), 131–142 (2021)
3. Baldi, P., Brunak, S.: Bioinformatics: The Machine Learning Approach. MIT Press (2001)
4. Bashir, S., Khattak, I.U., Khan, A., Khan, F.H., Gani, A., Shiraz, M.: A novel feature selection method for classification of medical data using filters, wrappers, and embedded approaches. Complexity **2022**(1), 8190814 (2022)
5. Birba, D.E.: A comparative study of data splitting algorithms for machine learning model selection (2020)
6. Bonds, D.R., Gabbe, S.G., Kumar, S., Taylor, T.: Fetal weight/placental weight ratio and perinatal outcome. Am. J. Obstet. Gynecol. **149**(2), 195–200 (1984)
7. Chen, C.W., Tsai, Y.H., Chang, F.R., Lin, W.C.: Ensemble feature selection in medical datasets: combining filter, wrapper, and embedded feature selection results. Expert. Syst. **37**(5), e12553 (2020)
8. Chowdhury, M.Z.I., Turin, T.C.: Variable selection strategies and its importance in clinical prediction modelling. Family Med. Comm. Health **8**(1) (2020)
9. El-Habil, A.M.: An application on multinomial logistic regression model. Pakistan J. Stat. Oper. Res. **8**(2), 271–291 (2012)

10. Friedman, J., Hastie, T., Tibshirani, R.: Regularization paths for generalized linear models via coordinate descent. J. Stat. Softw. **33**(1), 1 (2010)
11. Gaccioli, F., Lager, S., Sovio, U., Charnock-Jones, D.S., Smith, G.C.: The pregnancy outcome prediction (pop) study: investigating the relationship between serial prenatal ultrasonography, biomarkers, placental phenotype and adverse pregnancy outcomes. Placenta **59**, S17–S25 (2017)
12. Galar, M., Fernández, A., Barrenechea, E., Bustince, H., Herrera, F.: An overview of ensemble methods for binary classifiers in multi-class problems: experimental study on one-vs-one and one-vs-all schemes. Pattern Recogn. **44**(8), 1761–1776 (2011)
13. García-Nieto, P.J., Garcia-Gonzalo, E., Paredes-Sánchez, J.P.: Prediction of the critical temperature of a superconductor by using the woa/mars, ridge, lasso and elastic-net machine learning techniques. Neural Comput. Appl. **33**, 17131–17145 (2021)
14. Gong, M.: A novel performance measure for machine learning classification. Int. J. Manag. Inf. Technol. (IJMIT) **13** (2021)
15. González, J., Ortega, J., Damas, M., Martín-Smith, P., Gan, J.Q.: A new multi-objective wrapper method for feature selection-accuracy and stability analysis for BCI. Neurocomputing **333**, 407–418 (2019)
16. Grandini, M., Bagli, E., Visani, G.: Metrics for multi-class classification: an overview. arXiv preprint arXiv:2008.05756 (2020)
17. Grobman, W.A.: Obstetric patient safety: an overview. Am. J. Perinatol. **29**(01), 03–06 (2012)
18. Haavaldsen, C., Samuelsen, S.O., Eskild, A.: Fetal death and placental weight/birthweight ratio: a population study. Acta Obstet. Gynecol. Scand. **92**(5), 583–590 (2013)
19. Hossin, M., Sulaiman, M.N.: A review on evaluation metrics for data classification evaluations. Int. J. Data Mining Knowl. Manag. Process **5**(2), 1 (2015)
20. Jeppegaard, M., Rasmussen, S.C., Anhøj, J., Krebs, L.: Winter, spring, summer or fall: temporal patterns in placenta-mediated pregnancy complications–an exploratory analysis. Arch. Gynecol. Obstet. **309**(5), 1991–1998 (2024)
21. Kwak, C., Clayton-Matthews, A.: Multinomial logistic regression. Nurs. Res. **51**(6), 404–410 (2002)
22. Lemos, D., Braga, A.C., Nogueira, R.: Nonlinear regression on growth curves for placental parameters in R. In: International Conference on Optimization, Learning Algorithms and Applications, pp. 575–590. Springer (2023)
23. Lemos, D.F.M.: Regressão não linear em curvas de crescimento para parâmetros placentares em R. Master's thesis, Universidade do Minho (Portugal) (2023)
24. Matsuda, Y., et al.: Impact of placental weight and fetal/placental weight ratio z score on fetal growth and the perinatal outcome. Int. J. Med. Sci. **15**(5), 484 (2018)
25. Mlambo, W., Cheruiyot, W.K., Kimwele, M.W.: N,"a survey and comparative study of filter and wrapper feature selection techniques". Int. J. Eng. Sci. **5**(8), 57–67 (2016)
26. Muraina, I.: Ideal dataset splitting ratios in machine learning algorithms: general concerns for data scientists and data analysts. In: 7th International Mardin Artuklu Scientific Research Conference, pp. 496–504 (2022)
27. Njeze, N.R., Ogbochukwu, J.O., Chinawa, J.M.: Correlation of ultrasound placental diameter and thickness with gestational age. Pak. J. Med. Sci. **36**(5), 1058 (2020)

28. Nogueira, R., et al.: Placental biometric parameters: the usefulness of placental weight ratio and birth/placen-tal weight ratio percentile curves for singleton gestations as a function of gestational age. J. Clin. Anat. Pathol. **4**, 1–15 (2019)
29. Ogutu, J.O., Schulz-Streeck, T., Piepho, H.P.: Genomic selection using regularized linear regression models: ridge regression, lasso, elastic net and their extensions. In: BMC Proceedings, vol. 6, pp. 1–6. Springer (2012)
30. Pettker, C.M., Grobman, W.A.: Obstetric safety and quality. Obstet. Gynecol. **126**(1), 196–206 (2015)
31. Salavati, N., Gordijn, S.J., Sovio, U., Zill-E-Huma, R., Gebril, A., Charnock-Jones, D., Scherjon, S.A., Smith, G.C.: Birth weight to placenta weight ratio and its relationship to ultrasonic measurements, maternal and neonatal morbidity: a prospective cohort study of nulliparous women. Placenta **63**, 45–52 (2018)
32. Susan, S., Kumar, A.: The balancing trick: optimized sampling of imbalanced datasets–a brief survey of the recent state of the art. Eng. Rep. **3**(4), e12298 (2021)
33. Wah, Y.B., Ibrahim, N., Hamid, H.A., Abdul-Rahman, S., Fong, S.: Feature selection methods: case of filter and wrapper approaches for maximising classification accuracy. Pertanika J. Sci. Technol. **26**(1) (2018)
34. Zou, K.H., O'Malley, A.J., Mauri, L.: Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. Circulation **115**(5), 654–657 (2007)