Shiny App to Predict the Risk of Death in Very Low Birth Weight Newborns Through a New Classifier

Claudia Rodrigues[1(✉)], Ana Rita Antunes[2], and Ana Cristina Braga[2]

[1] School of Engineering, University of Minho, Campus Gualtar, Braga, Portugal
[2] ALGORITMI Centre, University of Minho, 4800-058 Guimarães, Portugal
id9069@alunos.uminho.pt, acb@dps.uminho.pt

Abstract. Neonatal mortality is a major health concern worldwide, and one of the groups of babies that has contributed most to this outcome is the very low birth weight (VLBW) newborns. Knowing the factors that contribute to the death of these babies and having an idea of the probability of their survival can help health professionals to make important decisions. The objective of this study was to develop a classifier that predicts whether a particular newborn weighing less than 1500 g will survive or die. For this, the tools R and RStudio were used, as well as a database provided by the National Registry of Newborns of Very Low Weight from Portugal. Construction of logistic regression models were performed, and a shiny web application was created to facilitate the use of the final model by health professionals. The application is available in https://rodrigues.shinyapps.io/VLBWnewborn/. Multivariate analysis showed that gestational age, length at birth, prenatal corticosteroids, sex, average of the three Apgar scores (1st, 5th and 10th min), insufflator resuscitation, major congenital malformation, diagnosis of necrotizing enterocolitis and administration of Ibuprofen for treatment of persistent ductus arteriosus are factors responsible for neonatal mortality. This model predicts with 0.926 of certainty the real state of a VLBW newborn. Its area under the ROC curve was 0.891 and 0.797 for internal and external validation, respectively. The fact that it presents better predictive results than the CRIB and SNAPPE II indexes, when using test data, makes it a possible alternative index to be used.

Keywords: Logistic regression · ROC curve · Shiny · Classifier

1 Introduction

In the 21st century, neonatal mortality is still one of the main health concerns worldwide, where 45% of all deaths occurring in children with less than five

This work has been supported by FCT - Fundação para a Ciência e Tecnologia within the R&D Units Project Scope: UIDB/00319/2020. The authors express their gratitude to the Portuguese National Registry for supplying the dataset used in this study.

© Springer Nature Switzerland AG 2021
O. Gervasi et al. (Eds.): ICCSA 2021, LNCS 12952, pp. 593–608, 2021.
https://doi.org/10.1007/978-3-030-86973-1_42

years occur within the neonatal period [12]. Newborns with very low birth weight (VLBW) are one of the groups of babies that have most contributed significantly to neonatal mortality. A VLBW newborn is defined as one with a birth weight less than 1500 g and/or is born before 32 weeks of gestation [11].

The LBW has been identified as one of the most common causes of mortality in newborns. It is estimated that neonatal mortality worldwide is 20 times more likely to occur in newborns with low weight than in newborns with normal weight [14]. According to Cutland et al. (2017), newborns with low weight are more likely to present several morbidities, such as neurologic disability, increased risk of chronic diseases and exhibit also greater probability of occurrence of intracranial hemorrhage, respiratory distress, blindness, gastrointestinal disorders, among others [3]. These conditions make these babies more vulnerable to death.

However, more recent studies have been considering other variables, in addition to low weight, as possible mortality risk factors, such as male gender, congenital malformations, among others [19]. Understanding what are the factors that contribute to the mortality of VLBW newborns, among the variables that are routinely used by doctors, can help health professionals to make the best decisions in planning perinatal care and during each stage of neonatal intensive care.

Thus, severity scoring systems have been developed in order to assess mortality risk of these babies. Among the existing indexes, the most used to more objectively analyze the perspective of survival and the quality of life of very low birth weight newborns are both the CRIB (Clinical Risk Index for Babies) and the SNAPPE (Score for Neonatal Acute Physiology-Perinatal Extension) [6].

According to Medlock et al. (2011), a typical approach to estimate the risk of death of VLBW newborn and to identify the main mortality risk factors involves the use of predictive models such as logistic regression models [10]. However, before adopting the predictive model in health care units, it is necessary to determine its predictive capacity, in order to prevent it from misclassifying a particular newborn. It is usual to demonstrate the performance of a clinical classifier by analyzing the ROC curve (Receiver Operating Characteristic).

The ROC curve is defined as a plot which its axes are represented by the True Positive Fraction of a classifier (sensitivity), on the y-axis, and by the False Positive Fraction (1-specificity), on the x-axis, and where each point is generated by a different threshold value [7]. In addition, the ROC analysis also allows the visualization of several classifiers, in order to select the best predictive model based on its performance measures. One of the most used indicators in this methodology, which allows evaluate the performance of the models, is the Area Under the Curve (AUC) [5].

Nowadays, while many health professionals use analysis tools, often commercial type, to analyze ROC curves, most of them have limited analytical resources [7]. Alternatively, there are free software packages, such as R, that offer advanced analysis features, but they are more difficult to use as they require the use of a command-based user interface. This can be challenging for users who have no

knowledge of the R language. In this sense, web applications based on the R language have been developed in which the end-user does not need to write any line of code. This is possible by using the shiny package available in RStudio. In a study carried out by Ribeiro (2017), a shiny application that predict the admission status of a very low birth weight newborn was constructed [15]. In this case, the factors gestational age, weight, length, administration or not of prenatal corticosteroids, the gender of the newborn, type of delivery, delivery reason, Apgar score at 1st and 10st min and the existence or not of congenital malformation were chosen to be part of the logistic regression model implemented in the application.

The objective of this article is to present a logistic regression model that can serve as a classifier that predicts whether a very low birth weight newborn will survive or die.

2 Methods

2.1 Binary Logistic Regression

Logistic regression was originally applied in the field of epidemiology, and today it is also used in other fields such as biomedical sciences, finance, criminology, ecology, engineering, biology, among others [8].

This technique is commonly used in a perspective of prediction and analyzing relationships between variables. When considering a binary logistic regression, the dependent variable can take two possible values denoted by 0 and 1, where 1 represents the occurrence of the event and 0 the absence.

Using a binary dependent variable and p independent predictors, $\mathbf{x}$, the logistic regression model, in terms of the event probability, $\pi(\mathbf{x})$, can be written as represented in Eq. 1:

$$\pi(\mathbf{x}) = \frac{e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\ldots+\beta_p x_p}}{1+e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\ldots+\beta_p x_p}} \tag{1}$$

where β_0, β_1, β_2, ..., β_p are coefficients of the *logit* model [8]. The *logit* is the transformation of $\pi(\mathbf{x})$, that is, the linearization of the multivariate logistic model, and is defined in terms of $\pi(\mathbf{x})$, as shown below in Eq. 2:

$$g(\mathbf{x}) = \ln\left[\frac{\pi(\mathbf{x})}{1-\pi(\mathbf{x})}\right] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p \tag{2}$$

Variable Selection Method. For the construction of the logistic regression model, the researcher must evaluate the association of each independent variable with the binary outcome. For that, some researchers opt for selection techniques like stepwise procedures. In stepwise regression, the selection procedure is automatically performed by statistical packages, in which examples of variable selection criteria are the Akaike information criterion (AIC) and the Bayesian information criterion (BIC).

This method of selecting variables can be used in three different directions, Forward, when we start with an empty model and successively add covariates, Backward, when we start with the full model and remove covariates and Both, when it is used both directions.

2.2 Diagnostic Measures of the Prediction Model

ROC Curve. The ROC curve is a graphical approach where the x-axis and y-axis represent false positive fraction, FPF (1-specificity) and true positive fraction, TPF (sensitivity), respectively, where its performance is given by the area under the curve. The analysis of these curves is considered a useful tool to visualize, organize and select classifiers according to their performance. Classifiers with AUC equal to 1 have perfect discrimination capacity, while those with an AUC of 0.5, their discrimination capacity is null [4]. For a clinical context Fan et al. (2006) indicates that ROC curves with an AUC value of 0.75 or less are not clinically useful but an AUC of 0.97 has a very high clinical value.

Sensitivity and Specificity. A good predictive model has a high value of TPF (sensitivity) and a low value of FPF (1-specificity), being considered ideal when the value of TPF is equal to one and the value of FPF is equal to zero [1]. In contrast, when the TPF and FPF values are the same, this means that the classifier is not conclusive about the prediction.

Akaike Information Criterion. AIC is a criterion widely used in the selection of the best model among a set of candidate models [21]. The selection of the best model pass by choosing the one with the lowest AIC value. The AIC criterion is defined by the following equation:

$$AIC = -2(\ln(L)) + 2k \tag{3}$$

where k represents the number of independent parameters and L represents the value obtained from the maximum likelihood function of the model.

Bayesian Information Criterion. BIC is another criterion used in the selection of the best model, and the one with the lowest BIC value will be chosen [20]. This criterion is represented by the following expression:

$$BIC = -2\log(L) + k\log(n) \tag{4}$$

where L is the likelihood function evaluated at the maximum likelihood estimate (MLE), k is the number of parameters in the model and n is the total number of observations.

Pseudo-R^2. In logistic regression, the adequacy of a model can also be assessed by the pseudo-R^2. Although there are many different statistics of R^2, none of them has been accepted as a standard.

One of the most used pseudo-R^2 in the context of logistic regression is the McFadden index. This index expresses the percentage variation between the model's Likelihood Value, which considers only the constant, and the Likelihood Value, which incorporates the explanatory variables, as is indicated in the following expression:

$$R^2logit = \frac{[-2LLnull - (-2LLmodel)]}{(-2LLnull)} \tag{5}$$

where LL represents the logarithm of the likelihood functions. According to Hu & Lo (2007) models that present McFadden-R^2 values between 0.2 and 0.4 are considered well-adjusted models [9].

Other pseudo R^2 that are used in different studies are Cox & Snell R^2 and Nagelkerke R^2, among others. The Nagelkerke index corresponds to an improved version of Cox & Snell since it restricts its value so that it does not exceed the value of 1.

2.3 Shiny Application

Today, there are several commercial software, such as IBM® SPSS® Statistics, MedCalc, Stata® and free software, such as R, which are used in order to guide various health professionals in their decision making. However, each of these statistical analysis' tools offers different resources and some are more limited than others. Among these, R has proved to be one of the most complete, thanks to the thousands of packages it has.

One of the packages that has attracted the interest of several specialists is the shiny package. With this package it is possible to build web applications that are easy to use, quick to analyze data (such as on [2] and [13]) and they do not require great knowledge of programming language.

Regarding its structure, shiny is basically divided into two components, the user interface (UI) and the server. The user interface object is the set of code that defines the aesthetics of the application and is responsible for showing to the user the inputs to perform a certain task [18]. The server function corresponds to the code segment that decides how the inputs entered by the user will be used and how the result will be presented in the output, that is, it will be responsible for the calculations that the application will perform. As it is a tool that makes use of "reactive programming", it ensures that changes in the input are immediately reflected in the output, thus enabling the construction of a highly interactive tool [17].

To make the Shiny application accessible to other users, it is possible to put it on host platform like "Shinyapps.io" and GitHub [16].

3 Results and Discussion

All the statistical analysis performed on the data, as well as the construction of the algorithm and the web application, were made using the RStudio (version 1.3.959) and R (version 4.0.1) tools.

3.1 Database Characterization

The dataset used in this study is part of the Portuguese National Registry of Very Low Weight Newborns which include the data collected by neonatal intensive care units in the Portuguese territory between the years 2010 and 2012. All newborns registered in this database had a weight of less than 1500 g and/or gestational age of less than 32 weeks.

The original database included 3496 records and 105 variables. However, for this study, only a sample of 2306 newborns were used, in which all twins and all newborns who died in the delivery room were excluded. Of these 2306 VLBW newborns, a total of 263 (11.41%) babies ended up dying.

The age of mothers of these babies ranged from 14 to 47 years with mean of 30.25 (±6.15) years. The median age of mothers was 31. Considering the admission status of newborns, the average age of mothers for survivors and dead is 30.30 and 29.88 years, respectively. Most of these mothers had pathologies during pregnancy (66.10%) and cesarean section was the most frequent type of delivery (69.12%). Regarding the reason for the birth, spontaneous delivery was the most prevalent among mothers (44.93%), followed by fetal pathology (28.83%), voluntary terminations of pregnancy (VTP) (25.90%) and maternal pathology (0.35%).

Regarding care and treatments performed during the prenatal period, 93.40% of mothers received care before the birth of their babies and only 4.09% of mothers were subjected to medically assisted conception. Regarding the administration of corticosteroids before delivery, 12.91% of mothers did not receive corticosteroids, 25.96% of them delivered within less than 24 h after the first dose of corticosteroids or more than a week after the last dose, and finally, most mothers (61.13%) had their birth in more than 24 h and less than a week, after the administration of at least one dose of corticosteroids.

The babies in this database have an average gestational age of 207.80 (± 18.13) weeks and a birth weight with an average of 1196 (±356.97) g. The birth weight values are ranged from 370 to 2810 g and the median is 1210 g. In this sample of newborns, 25% of them have a birth weight less than or equal to 930 g and 25% a weight greater than 1440 g. Regarding the variation in the weights of newborns according to the state of admission, survivor and dead, it was seen that, in general, survivors have a higher birth weight than newborns who end up dying. In this sample, it was also observed that it is more common for survivors to have a higher gestational age value than newborns who end up dying where 50% of the survivors had a gestational age of 212 days or less and 50% of the dead had a gestational age of 186 days or less. In addition, in this database there are no dead newborns with gestational ages higher than 268 days. Besides that,

the average length at birth of newborns is 37.16 (±3.62) cm and the average head circumference at birth is 26.57 (±2.66) cm.

In this study it was also observed that the sample of 1207 (52.34%) of male babies has a slightly higher mortality rate (6.7%) compared to the sample of 1099 (47.66%) of the female gender (4.7%).

Analyzing the index that is equivalent to the average of the values of three Apgar indexes (1st min, 5th min, 10th min), it was found that its average is 7.94 and its median 8. The lowest value present in the study sample was the value 1, with three newborns having this score, and the maximum value of the index was 10 with 419 newborns. Within this scale, the value 9 was the most frequent, with 599 newborns receiving this score. Babies who were given a score of 1,2 and 4 on the Apgar index, a large part of them correspond to newborns who ended up dying with values of 66.7%, 66.7% and 51.7%, respectively. Babies that have a higher value on the Apgar index, a large part of them ended up surviving, and the score of 10 on the Apgar index is the one with the lowest number of dead newborns (2.4%). Regarding surviving newborns, 50% of them had an Apgar index equal to or less than 8, while 50% of dead newborns had a value equal to or less than 6.

Analyzing the values of the CRIB and SNAPPE II in relation to newborns for this database, it was found that the CRIB index has an average score of 2.74, a median of 1 and a final score between 0 and 20 and the SNAPPE II index has an average score of 21.64, a median of 15 and a final score ranging from 0 to 131.

Of the available resuscitation methods, the oxygen resuscitation process was the most used (60.72%). Cardiac compression and adrenaline resuscitation were the methods by which health professionals least opted, with only 4.27% and 3.70% of babies using such procedures, respectively. Among the 8 most common diseases observed in this population, the disease respiratory distress syndrome (RDS) was the most detected in these newborns (75.24%) and the least detected was gastrointestinal (GI) perforation (2.6%).

Finally, the average of hospitalization time of these babies was 49.12 (±41.00) days, and the median was 41 days. The minimum value recorded for the hospitalization time was 0 days and the maximum value was 524 days, that is, about 75 weeks.

3.2 Database Treatment

Since the main objective of this study was develop a model capable of predicting the risk of death of a very low birth weight newborn, it was necessary, in a first phase, to treat the database in order to exclude all the variables that should not be included into the model as predictor variables, as well as treat the missing values present in the database.

As mentioned before, all the data related to newborn twins and newborns who ended up dying in the delivery room were excluded from the database. In addition, 56 variables were also excluded in the first phase due to their lack of relevance to the case study. All the records with strange values and denominations have been recoded to missing values.

For this database it was decided to create a new variable "AverageApgar" that corresponds to the average of the indexes of Apgar at the 1st, 5th and 10th min.

In order to build the predictive model, it was necessary to apply a method for treating missing values in the database used in this study. In this case, the deletion method was chosen to deal with the 1.3% of missing values present in the database. Thus, the database became composed of 1535 cases and 49 variables.

3.3 Logistic Regression Model Development

To develop and validate a classifier for mortality risk, the 1535 VLBW newborns were divided into two data groups using the *caret* package (version 6.0–86), present in R, with a split of about 70:30. In this case, 70% of the data worked as the training set to train the model and the remaining 30% worked as the test set to validate the developed model. The training set consisted of 1075 records and the test set consisted of 460 records.

Before proceeding with the construction of the logistic regression model, it was necessary to choose the best set of independent variables to be introduced in the predictive model, eliminating those that did not present variability. For that, statistical and exploratory techniques were used to see which variables that a priori would not contribute mathematically to the model and could be eliminated. It started by performing a variable-to-variable univariate procedure, using logistic regression, to eliminate those that were not significant, that is, that had a p-value above 0.25. The choice of this value was based on Hosmer & Lemeshow in which they proposed this value as a criterion for choosing the variables to be implemented in the multivariate analysis [8]. Then, an exploratory analysis was carried out through simple frequency tables for all 49 variables, in order to observe the distribution of their data at each level. All the variables that showed a significant difference in the distribution of data were eliminated from the database. Finally, cross tables were built for each variable with the dependent variable, "Death", in order to analyze the correspondence of surviving and dead newborns for each variable. At the end of this analysis, a set of 18 independent variables was obtained, the possible candidates to be implemented in the model to be developed (Table 1).

In this study, different multiple logistic regression models were built. For that, the training set was used to develop different logistic regression equations having mortality as the outcome variable. All the candidate variables were included in a complete multivariate model, that is, in the first model built in this study. From this model, others were built through the application of the stepwise technique on it and by choosing variables manually that had a p-value less than a value established by us, in this case 0.050 for some models and 0.058 for others.

The model chosen for the classifier was obtained by applying the analysis of variance (ANOVA) on the complete model. This model consists of 9 independent variables corresponding to the variables that were indicated in the ANOVA analysis with a p-value less than 0.05. The statistical data related to the chosen model are shown in Table 2.

Table 1. Summary of the characteristics of the independent variables candidates to be included in the LR model.

Variable	Codes/Levels	Coefficient	p-value*
GestationalAge	–	−0.053	<0.001
WeightBirth	–	−0.002	<0.001
LengthBirth	–	−0.234	<0.001
HeadCircumferenceBirth	–	−0.313	<0.001
PrenatalCorticosteroids2	Partial	−1.352	<0.001
PrenatalCorticosteroids3	Complete	−1.006	0.001
PathologiesPregnancy1	Yes	−0.315	0.222
TypeDelivery2	Caesarean	−0.577	0.023
ReasonDelivery2	Maternal Pathology	2.478	0.015
ReasonDelivery3	Fetal Pathology	−0.501	0.112
ReasonDelivery4	VTP	−0.405	0.200
Sex2	Female	−1.301	<0.001
AverageApgar	–	−0.550	<0.001
OxygenResuscitation1	Yes	1.597	<0.001
InsufflatorResuscitation1	Yes	0.494	0.049
EtIntubationResuscitation1	Yes	1.900	<0.001
MajorCongenitalMalformation1	Yes	1.241	<0.001
DiagRds1	Yes	1.404	0.001
DiagPda1	Yes	1.252	<0.001
DiagNec1	Yes	2.004	<0.001
PdaTerapeutic1	Yes	0.616	0.040

*The p-value corresponds to Wald's statistic

Table 2. Output obtained for the final logistic regression model.

Variable	Coef.	Std. Error	z	p-value
Constant	11.400	1.970	5.811	6.20e−09**
GestationalAge	−0.019	0.012	−1.646	9.97e−02
LengthBirth	−0.149	0.054	−2.747	6.01e−03**
PrenatalCorticosteroids2	−2.053	0.485	−4.232	2.32e−05**
PrenatalCorticosteroids3	−0.981	0.379	−2.590	9.60e−03**
Sex2	−1.384	0.345	−4.016	5.93e−05**
AverageApgar	−0.482	0.092	−5.254	1.49e−07**
InsufflatorResuscitation1	−0.600	0.309	−1.938	5.26e−02
MajorCongenitalMalformation1	1.590	0.459	3.462	5.36e−04**
DiagNec1	1.930	0.353	5.481	4.22e−08**
PdaTerapeutic1	−0.332	0.365	−0.911	0.362

**p-values less than the 5% significance level, that is, they correspond to variables in the model that are statistically significant.

Results of the multiple logistic regression analysis (Table 2) revealed the impact of multiple factors on the probability of death in VLBW newborn. Gestational age, length at birth, non-administration of prenatal corticosteroids, the gender of the newborn, the average of the three Apgar scores (1st, 5th and 10th min), the non-occurrence of initial resuscitation with an insufflator, the presence of some major congenital malformation, presence of necrotizing enterocolitis (NEC), and the absence of Ibuprofen to treat ductus arteriosus (PDA) to the newborn were the risk factors for death in these babies.

3.4 Logistic Regression Model Validation

The predictive models were validated with the test set, which consisted in 460 VLBW newborn records. Some metrics chosen to help to select the most parsimonious model among of the different models developed were Akaike information criterion (AIC) and Bayesian information criterion (BIC). Pseudo-R^2 as McFadden R^2, Cox & Snell R^2 and Nagelkerke R^2 were also chosen to assess the adequacy of a model. Furthermore, model discrimination was assessed using the area under the curve (AUC), sensitivity, specificity and predictive values. All these metrics were considered to compare the different predictive models developed. The one with the best goodness of fit, performance capacity value and a smaller number of variables was chosen to be the possible classifier (Table 3).

Table 3. Summary of model validation measures.

Model #	1	2	3	4	5	6	7	8	9	10
#var	18	11	8	9	7	9	8	9	6	8
Quality of adjustment										
AIC	378.9	367.5	388	371	396	373.5	388	371	394.2	372.4
BIC	488.5	432.2	437.8	425.8	440.9	428.3	437.8	425.8	434.1	422.2
McFadden	0.346	0.334	0.282	0.319	0.262	0.314	0.282	0.319	0.262	0.312
Cox e Snell	0.152	0.147	0.128	0.141	0.117	0.139	0.126	0.141	0.117	0.138
Nagelkerke	0.401	0.388	0.332	0.372	0.310	0.367	0.332	0.372	0.309	0.365
Predictive ability										
Acc	0.926	0.924	0.924	0.924	0.924	0.926	0.924	0.924	0.924	0.924
Se	0.932	0.932	0.926	0.934	0.926	0.930	0.926	0.930	0.926	0.930
Sp	0.500	0.429	0	0.400	0	0.500	0	0.500	0	0.400
AUCi	0.903	0.901	0.878	0.895	0.870	0.891	0.878	0.895	0.868	0.893
AUCe	0.790	0.799	0.718	0.790	0.695	0.797	0.718	0.790	0.698	0.793

The final model (# 6) seems to be parsimonious and seems to fit well with the data, showing values of 373.498, 428.279 for AIC and BIC, respectively, as well as 0.314, 0.139 and 0.367 for McFadden R^2, Cox & Snell R^2 and Nagelkerke

R^2, respectively. In addition, the model also seems to have a good discriminating power, being able to predict with 0.926 certainty the admission status of a VLBW newborn. The prediction model presented an AUC of 0.891 when it was evaluated on the training sample and an AUC of 0.797 when evaluated in the test sample.

Figure 1 shows the ROC curves obtained in the internal (A) and external (B) validation processes of the chosen model. The best cutoff value presented by the model, when performed external validation, was 0.626. This value is generated by the combination of the best values of sensitivity and specificity, and to obtain them, the *optimalCutoff* function present in R was used. In this case, the values obtained for sensitivity and specificity were 0.930 and 0.500, respectively.

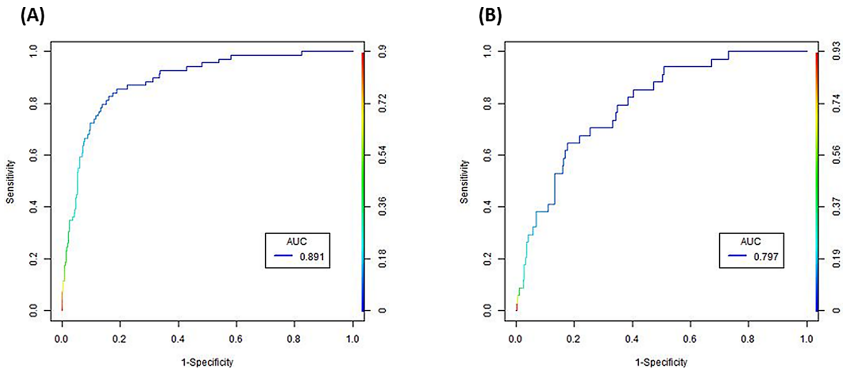

Fig. 1. ROC curves of the final model obtained in the internal validation (A) and in the external validation (B).

3.5 Comparison of the CRIB and SNAPPE II with the Developed Classifier

A comparative analysis was made between the classifier developed in this project and the mortality indexes CRIB and SNAPPE II, in order to assess whether the model developed could in fact be an alternative tool to be used in hospital units. In this case, a discriminative evaluation study was carried out, whose purpose was to estimate the predictive capacity of each indicator in predicting between two states, survival or death, of VLBW newborns. For this purpose, comparisons of ROC curves and their AUC values were made, using the *caTools* package (version 1.18.0) present in R.

Since, in the sample presented above, there are some newborns without simultaneously information referring to the CRIB and SNAPPE II indexes, for this part, a new set of test data was selected, which includes only newborns who meet these criteria. In this case, a sample of 350 records was used.

Figure 2 shows the three ROC curves for the CRIB and SNAPPE II indexes, as well as the classifier developed in this work (Model). When using the *caTools* package using the *coIAUC* function, it was also possible to obtain the values of its

AUC. The AUC values presented by the built classifier, the CRIB indicator and the SNAPPE II indicator were 0.909, 0.830 and 0.794, respectively. By analyzing the ROC curves and the AUC values, it can be concluded that the developed classifier has a better predictive capacity than the other two indexes.

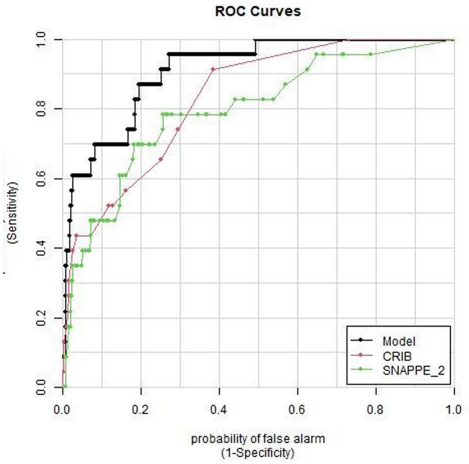

Fig. 2. Graphical representation of the ROC curves of the three indexes under study, obtained through *caTools*.

3.6 Shiny Application

Once confirmed a good predictive capacity of the developed model (final model) which will work as a classifier for the admission status of a VLBW newborn, a web application with shiny was planned, VLBWNB.

The idea of building this application arose in order to facilitate the estimation, in real time, of the probability of risk of death in these newborns, thus helping health professionals to make the best decisions in planning prenatal and postnatal care, so they can act in time. For that, the logistic regression model chosen was implemented in the built application. The application developed in this work is available for use in https://rodrigues.shinyapps.io/VLBWnewborn/.

Programming Code. For the creation of this application, we started by creating two R scripts in RStudio, the ui.R, which corresponds to the user interface object, and the server.R relating to the server function. Both files were saved in the same folder. Although shiny makes use of a template by default of Bootstrap and does not need the use of web development languages such as HTML, CSS and JavaScript for the building of the VLBWNB application, it was preferred to use the *shinythemes* package (version 1.1.2), as well as CSS, in order to customize the application in terms of its design. The code for its development can be found at https://github.com/Claudia-Rodrigues/Risk-of-death-VLBWNB.git.

The Shiny application developed for this work has an environment organized in 3 tabs, each with its own purpose. The functionalities of each section of the application will be explained as follow.

Home Page. As show in Fig. 3, the first tab of the application's navigation bar corresponds to the application's home page, which shows only the application title.



Fig. 3. Home page of the developed application.

About. As shown in Fig. 4, in the "About" tab we see explained the purpose of the application, the authors responsible for its development, the tools used for its construction, as well as, a brief description of the model/algorithm implemented on it. In the part of the model description, the accuracy of the model is presented, the AUC values for internal and external validation and the variables that are part of the model, accompanied by a short description of which values and metrics can be used for each one.

Prevision. In the "Prevision" tab of the Shiny application, the prevision x, if a particular very low birth weight newborn will survive or die is calculated in real time. Figure 5 shows the form, which is part of the "Prevision" tab, with the 9 inputs to be filled in by the user. These inputs correspond to the explanatory variables of the model implemented in the application.

After filling the form with the data of the newborn in question and pressing the calculate button, the application will indicate the probable state of admission of the newborn, as well as, the percentage of probability of dying. It should be noted that the use of this application for the calculation of the predicted risk of death of a VLBW newborn is limited to babies who present values within the

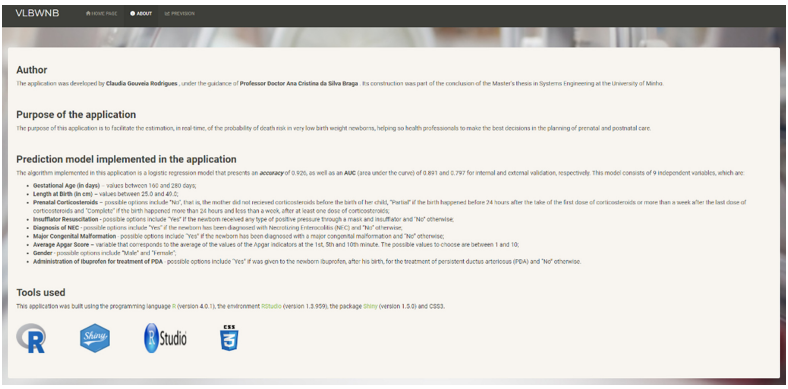

Fig. 4. Second tab of the VLBWNB application where the purpose of the application, the authors, the description of the implemented model and tools used in its construction are described.

limit presented by the application, for the 9 characteristics to be filled in the form. The possible values to be entered in each input of the form are shown in the "About" tab of the application. In addition, it should also be noted that this application is configured so that values equal to or greater than 63%, will indicate that the baby will die.

Figure 5 shows an example of an output provided by the "Prevision" tab, where the result is displayed in the right side of the application.

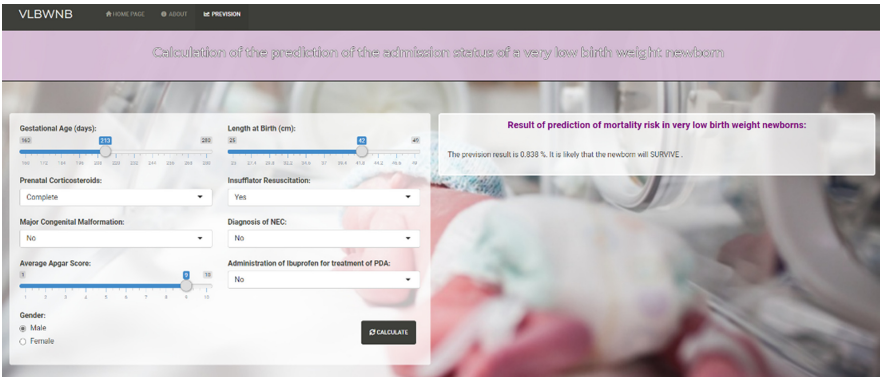

Fig. 5. Example of the output provided in the "Prevision" tab.

4 Conclusion

As mentioned before, the value of the mortality rate in VLB weight newborns is one of the main public health problems that has caused the most concerns worldwide. In this sense, it was intended with this work to develop a predictive model that was able to predict the risk of death in these kind of newborns. This model could help health professionals to understand which ones have a higher risk of death, and thus help them in making decisions about how they should act.

The model that served as basis for the construction of the classifier is a logistic regression model capable to predict with 0.926 certainty the admission status of a very low birth weight newborn. Through this model it was found that some of the factors that may contribute to the death of these babies are gestational age, length at birth, non-administration of prenatal corticosteroids, the gender of the newborn, the average of the three Apgar scores (1st, 5th and 10th min), the non-occurrence of initial resuscitation with an insufflator, the presence of some major congenital malformation, the presence of necrotizing enterocolitis (NEC), and the absence of Ibuprofen in treatment of ductus arteriosus (PDA) to the newborn.

This model presents a good fit to the data and a good predictive capacity.

Since it presented better predictive results than the CRIB and SNAPPE II indexes, which are widely used today, it leads us to believe that this classifier could be an alternative tool to be used in hospital centers. In addition, the VLBWNB web application that implements this classifier can be a more appealing and friendly tool to be used since in a few minutes it returns the probable result of a newborn's admission status. However, it should be noted that this application has limitations regard to the data that can be used to calculate the prediction of the newborn's admission status.

In a perspective of future work, it was possible to build another method of machine learning, such as, for example, Random Forest, Support Vector Machine, Artificial Neural Networks, among others.

References

1. Alonzo, T.A., Pepe, M.S.: Development and evaluation of classifiers. Topics in Biostatistics, pp. 89–116 (2007)
2. Antunes, A.R., Braga, A.C.: Shiny app to predict agricultural tire dimensions. In: Gervasi, O., Murgante, B., Misra, S., Garau, C., Blečić, I., Taniar, D., Apduhan, B.O., Rocha, A.M.A.C., Tarantino, E., Torre, C.M., Karaca, Y. (eds.) ICCSA 2020. LNCS, vol. 12251, pp. 247–260. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-58808-3_19
3. Cutland, C.L., et al.: Low birth weight: case definition & guidelines for data collection, analysis, and presentation of maternal immunization safety data. Vaccine **35**(48Part A), 6492 (2017)
4. Fan, J., Upadhye, S., Worster, A.: Understanding receiver operating characteristic (roc) curves. Canadian J. Emergency Med. **8**(1), 19–20 (2006)

5. Fawcett, T.: An introduction to roc analysis. Pattern Recogn. Lett. **27**(8), 861–874 (2006)
6. Gagliardi, L., Cavazza, A., Brunelli, A., Battaglioli, M., Merazzi, D., Tandoi, F., Cella, D., Perotti, G., Pelti, M., Stucchi, I., et al.: Assessing mortality risk in very low birthweight infants: a comparison of crib, crib-ii, and snappe-ii. Archives of Disease in Childhood-Fetal and Neonatal Edition **89**(5), F419–F422 (2004)
7. Goksuluk, D., Korkmaz, S., Zararsiz, G., Karaagaoglu, A.E.: easyroc: an interactive web-tool for roc curve analysis using r language environment. R J. **8**(2), 213 (2016)
8. Hosmer, D.W., Lemeshow, S., Cook, E.: Applied Logistic Regression, 2nd edn. Wiley, New York (2000)
9. Hu, Z., Lo, C.: Modeling urban growth in Atlanta using logistic regression. Comput. Environ. Urban Syst. **31**(6), 667–688 (2007)
10. Medlock, S., Ravelli, A.C., Tamminga, P., Mol, B.W., Abu-Hanna, A.: Prediction of mortality in very premature infants: a systematic review of prediction models. PloS One **6**(9), e23441 (2011)
11. Mourão, M.F., Braga, A.C., Oliveira, P.N.: Accommodating maternal age in crib scale: quantifying the effect on the classification. In: International Conference on Computational Science and Its Applications, pp. 566–579. Springer (2014)
12. Nayeri, F., Emami, Z., Mohammadzadeh, Y., Shariat, M., Sagheb, S., Sahebi, L.: Mortality and morbidity patterns of very low birth weight newborns in eastern mediterranean region: a meta-analysis study. J. Pediatrics Rev. **7**(2), 67–76 (2019)
13. Quintas, J.P., Machado e Costa, F., Braga, A.C.: ROSY application for selecting R packages that perform ROC analysis. In: Gervasi, O., Murgante, B., Misra, S., Garau, C., Blečić, I., Taniar, D., Apduhan, B.O., Rocha, A.M.A.C., Tarantino, E., Torre, C.M., Karaca, Y. (eds.) ICCSA 2020. LNCS, vol. 12251, pp. 199–213. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-58808-3_16
14. Raja, M., K, S.K., Deneshkumar, V.: Regression modeling for maternal determinants of low birth weight. Int. J. Stat. Syst. **12**(3), pp. 585–591 (2017)
15. Ribeiro, D.S.: Previsão do risco de morte de recém-nascidos prematuros de muito baixo peso. Master's thesis, University of Aveiro (2017)
16. RStudio: Share your apps. https://shiny.rstudio.com/tutorial/written-tutorial/lesson7/ (2020). Accessed 18 Mar 2021
17. Seal, A., Wild, D.J.: Netpredictor: R and shiny package to perform drug-target network analysis and prediction of missing links. BMC Bioinform. **19**(1), 1–10 (2018)
18. Sivaprakasam, B., Sadagopan, P.: Development of an interactive web application "shiny app for frequency analysis on homo sapiens genome (safa-hsg)." Interdisciplinary Sciences: Computational Life Sciences **11**(4), 723–729 (2019)
19. Vincer, M.J., Armson, B.A., Allen, V.M., Allen, A.C., Stinson, D.A., Whyte, R., Dodds, L.: An algorithm for predicting neonatal mortality in threatened very preterm birth. J. Obstetrics Gynaecology Canada **37**(11), 958–965 (2015)
20. Xi, R., Hadjipanayis, A.G., et al.: Copy number variation detection in whole-genome sequencing data using the bayesian information criterion. Proc. Natl. Acad. Sci. **108**(46), E1128–E1136 (2011)
21. Yanagihara, H., Kamo, K.I., Imori, S., Satoh, K.: Bias-corrected aic for selecting variables in multinomial logistic regression models. Linear Algebra Appl. **436**(11), pp. 4329–4341 (2012)