

Predicting Neonatal Condition at Birth through Ensemble Learning Methods in Pregnancy Care

Mário W. L. Moreira^{1,2}, Joel J. P. C. Rodrigues^{1,3,4}, Guilherme A. B. Marcondes³,
Augusto J. Venâncio Neto⁵, Vasco Furtado⁴

¹Instituto de Telecomunicações, Universidade da Beira Interior (UBI)
Covilhã – Portugal

²Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE)
Aracati, CE – Brazil

³National Institute of Telecommunications (Inatel)
Santa Rita do Sapucaí, MG – Brazil

⁴University of Fortaleza (UNIFOR)
Fortaleza, CE – Brazil

⁵Department of Informatics and Applied Mathematics
Federal University of Rio Grande do Norte (UFRN) – Natal, RN – Brazil

{mario.moreira, joeljrc}@ieee.org, guilherme@inatel.br,

augusto@dimap.ufrn.br, vasco@unifor.br

Abstract. *Prematurity represents the determinant cause of infant mortality. This serious public health problem is directly related to the assistance provided during pregnancy and childbirth. Hence, this paper proposes the use of leading machine learning techniques capable of supporting health experts in pattern recognition in the prediction of high-risk situations for the fetus. The proposed model creates an ensemble of nearest-neighbor classifiers using the random subspace algorithm, reaching an overall accuracy of 0.937 and area under the curve of 0.721, in predicting the Apgar score, and 0.829 and 0.669 in predicting if the newborn will be small for gestational age, respectively. These results show the model effectiveness in reducing severe pregnancy related-problems.*

1. Introduction

According to the United Nations (UN), the infant mortality rate is an indicator of the quality of health services, basic sanitation, and education. This rate expresses the number of children, in a given place, who died before completing five years of life for every 1,000 live births [You et al. 2015]. Among the principal causes of infant mortality are the lack of assistance and education for pregnant women, absence of medical supervision, deficiency in hospital care, malnutrition, deficits in environmental sanitation services, among others [Wang et al. 2014]. According to UN reports, the world infant mortality rate is 45 deaths per 1,000 live births. These data are continually declining. However, it is critical to regard that this reduction does not occur in the same way in all countries. In economically developed nations, the infant mortality rate is considerably low, with some averaging less than three deaths per 1,000 born. Given this scenario, the UN has included

reducing child mortality among one of the eight Millennium Development Goals (MDGs) [Gaffey et al. 2015]. For the goal to be achieved, developed countries must contribute to the structuring of the nations that face this critical social problem through the construction of hospitals, training of medical staff, family education, subsidies for adequate feeding, environmental sanitation, among others. A significant barrier to reducing infant deaths is the lack of adequate medical monitoring of high-risk pregnancies. In this sense, information and communication technologies (ICTs) can contribute to a better physician-patient communication anywhere and anytime, thus improving pregnant women's healthcare.

Concerning to growth of use by health organizations of electronic health (e-health) systems, the flow of medical information has also increased [Nguyen et al. 2014]. Similarly, interconnected devices based on the Internet of Things (IoT) paradigms generate extensive information databases [Al-Fuqaha et al. 2015]. Platforms based on the Big Data concept can store these vast amounts of data securely and reliably. In the same agility, these platforms can find information in the midst of thousands in a few seconds. The acquisition, analysis, and understanding of data are definitive factors in the decision-making process in health [Andreu-Perez et al. 2015]. Pattern recognition is an essential issue for success in monitoring high-risk situations. Therefore, to contribute to the MDGs in reducing the number of maternal and fetal deaths, it is necessary to use novel techniques of data analytics capable of inferring from the risk factors and symptoms presented by the pregnant woman the severity of a pregnancy-related condition in real time.

There are some diseases that, when pre-existing, can aggravate the health state of the woman during the gestation. The most prevalent are high blood pressure, diabetes, heart disease, kidney disease, epilepsy and autoimmune diseases [English et al. 2015]. Maternal history of fetal death, repetitive late miscarriages, poor uterine formations, twin pregnancy, premature delivery, or fetal malformation constitute a high-risk pregnancy. Preeclampsia and eclampsia (increased blood pressure), gestational diabetes, hemorrhages, infectious diseases that can cause fetal malformation and rupture of the amniotic sac with the premature fetus [Abalos et al. 2014]. The problem of a high-risk pregnancy without adequate prenatal care may range from mild neonatal infections, premature birth to maternal and fetal deaths. The principal preventive measure is adequate prenatal monitoring. The protocol adopted by the World Health Organization (WHO) aims at the early identification of the primary complications, indicating to health professionals the more unfavorable outcomes [Vogel et al. 2014]. Thus, the main contributions of this paper are the following:

- Prediction of fetus-related problems using a real pregnancy database;
- Performance assessment of several ensemble algorithms;
- Classification mechanism for the primary indicators of fetus-related problems: hypertensive disorders, delivery method, Apgar score, and intrauterine growth restriction problems;
- Use of the 10-fold cross-validation method for proving the feasibility of the proposed methods.

The remainder of this paper is organized as follows. Section 2 elaborates on related work about the topic focusing on smart decision support systems (DSS) in its application on healthcare. Section 3 describes the use of ensemble learning techniques for pattern recognition in predicting fetus-related problems in pregnancies that developed hypertensive disorders in pregnancy. The performance evaluation, comparison of various

methods, and the analysis of the results of the proposed approaches is presented in Section 4. Finally, the Section 5 concludes the paper and suggests further works.

2. Related Work

The combination of multiple classifiers (ensembles) to generate a consolidated classifier has been a promising area of research in recent years. Several studies have provided theoretical results and empirical evidence on the efficacy of such approaches to solve complex classification problems. Besides, recent state-of-the-art has shown that an aggregate method obtained from diversified classifiers is commonly better than its components, *e.g.*, [Guan et al. 2014, Krawczyk et al. 2017].

In [Huda et al. 2016], the authors proposed to solve issues related to brain tumor diagnosis in unbalanced medical data through aggregation of attribute selection and ensemble-based classification. This approach reduced the disequilibrium of a healthcare data set using a hybrid wrapper-filter feature selection for universal optimization together with ensemble classification. The results showed that the proposed method outperformed the conventional techniques regarding brain tumor classification. Ramírez *et al.* presented a performance comparison of the most ensemble learning methods used in recent literature, to know, the boosting and bagging algorithms [Ramírez et al. 2016]. This study was applied to magnetic resonance image (MRI) analysis and classification and the early detection of the Alzheimer's disease. These methods were applied on an attribute extraction process, estimating first-order statistics from cerebral gray matter segmented MRI using decision tree ensembles. The results showed that the bagging technique outperformed the boosting algorithm concerning classification error and convergence rate. Rathore *et al.* proposed the use of data mining techniques to classify several types of cancer data [Rathore et al. 2014]. This study concluded that ensemble classification methods presented the best results for almost all data bases, principally in predicting the breast cancer patient survivability.

Regarding the prediction of fetus related-problems, Cnattingius and Villamor studied how the maternal overweight and obesity can influence the infant mortality and stillbirth [Cnattingius and Villamor 2016]. This study investigated associations between change in the maternal body-mass index (BMI) during early pregnancy and the related risks to stillbirth and neonatal, post-neonatal, and infant mortality in a posterior gestation using a binomial regression method. Iliodromiti *et al.* discussed the Apgar score as an important indicator for an early predicting of neonatal related-problems [Iliodromiti et al. 2014]. This research established an inherent relationship between the Apgar score at the newborn first 5 minutes of life and the risk of neonatal and infant mortality, considering specific causes. For the data analysis, this study considered the binomial log-linear modeling and adjustment for confounders. Results showed the importance of the gestational age during childbirth and that a low Apgar score at first minutes of life is strongly associated with the neonatal death risk.

Therefore, this paper proposes the use of novel machine learning (ML) techniques for analysis of data related to gestation, considering the related works presented above. As predictive classes, an Apgar score below 7 in the first 5 minutes of life of the newborn was considered as a critical parameter for severity, as well as the fetal development during pregnancy. These two critical indicators are discussed in the next section.

3. Methods Based on Ensemble Learning for Prediction of Critical Indicators for Fetal-related Problems

This section discusses the conditions that can lead the newborn to develop serious health problems, as well as the ensemble-based approaches for pattern recognition from symptoms and risk factors presented by the pregnant woman during pregnancy.

The Apgar test is the first test done by the newborn to assess its adjustment level to the life outside the womb. This test consists of the evaluation of five vital signs of the newborn. In the first minute to determine the baby tolerance degree to the childbirth, in the fifth minute to assess the baby's degree of adaptation to the extra-uterine life, and in the tenth minute after childbirth. Each of the signals is assigned a score of 0 to 2 according to the observed condition [Iliodromiti et al. 2014]. This simple method allows determining if the newborn needs support to breathe or if it has a heart problem. The Apgar score measures the newborn's vitality through a score of 5 parameters, namely, heart rate, breathing, muscle tone, reflex irritability, and skin color. The closer to 10 the score, the better the baby's vitality and the better its adaptation to extra-uterine life. The score of seven, eight or nine is considered normal, representing a sign that the newborn is in perfect health. A score of ten, especially at the first minute, is very unusual since almost all newborns lose a point due to blue hands and feet, a characteristic after childbirth. A score below 7 represents a sign that the newborn needs medical attention. The lower the score, the more support it needs to adjust to extra-uterine life. A lower score could be due to childbirth complications, a cesarean section, or the presence of fluid in the newborn's respiratory tract. A newborn presenting a low Apgar score needs support for breath for the heart rate to normalize to expected values. Normally, a low score in the first minute evolves to a standard five-minute score. It is important to consider that the Apgar score is commonly used as a parameter and not as a definition of prognosis, *i.e.*, if the baby presents a low score, it does not necessarily mean that it will present future problems and vice versa.

Another serious risk for the newborn is related to restricted intrauterine growth. A full-term, preterm, or post-term newborn whose birth weight is less than 90% of newborns of the same gestational age (below the 10th percentile) is considered small for gestational age (SGA) [Cetin et al. 2017]. There are several causes for this condition. In many cases, newborns are small only because of genetic factors such as having small parents. In other cases, the placenta can have malfunctioned. Therefore, the fetus did not receive adequate nutrients, which hindered its growth. A malfunctioning placenta can occur if the mother presents high blood pressure, preeclampsia, kidney disease, or diabetes. An acquired viral infection before birth can also be responsible. The fetus growth can also have been impaired if the mother smoked, consumed alcohol, or used illegal drugs during pregnancy. Newborn SGA non-premature infants do not have complications related to the immaturity of organ systems that preterm newborns of similar size present. They present, however, a higher risk of the following problems, namely meconium aspiration syndrome, excess of red blood cells, hypoglycemia, difficulty in regulating body temperature, as well as an impaired immune system. Hence, Figure 1 summarizes the attributes considered in this work for the prediction of these two conditions that are directly related to infant mortality.

In health, recent studies have shown that methods based on ensemble classifiers have a better performance than artificial intelligence techniques [Weng et al. 2016]. An

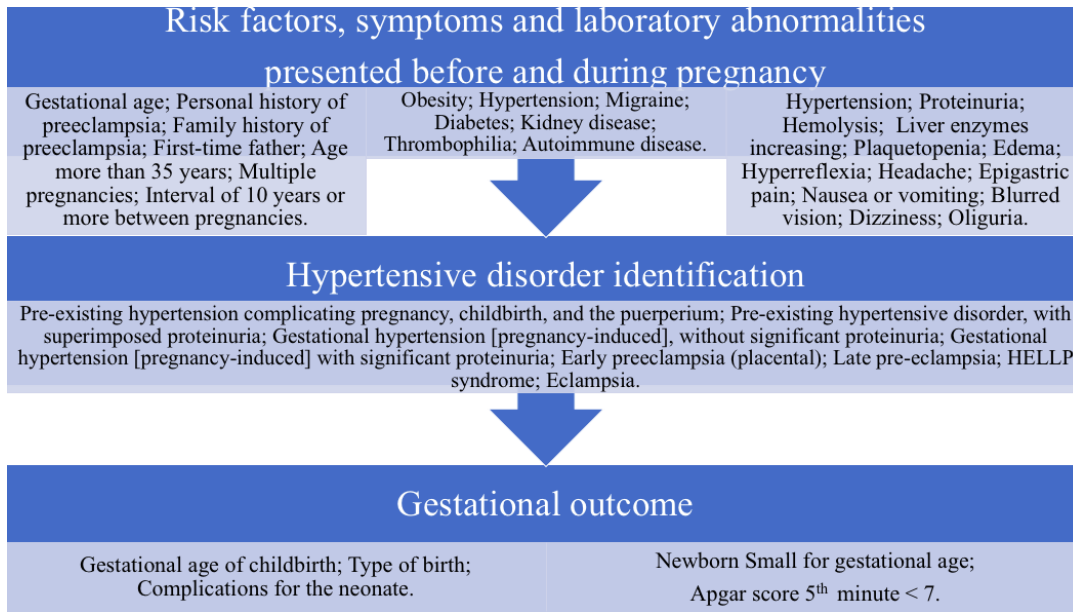


Figure 1. Attributes considered in this study for the prediction of problems that can lead both the pregnant woman and fetus to death.

ensemble classifier consists of a set of individually trained classifiers, called base classifiers, whose decisions are combined. The aggregating bootstrap algorithm generates a sampled bootstrap data set of the original data. The dataset generates a set of models using a simple learning algorithm by combining votes for classification. Its use is particularly attractive when the available information is of limited size. In the bagging method, classifiers are trained independently by different training sets through the boot method. To construct them it is necessary to assemble k identical training sets and replicate this training data at random to build k independent networks by re-sampling with replacement. The method then aggregates the k networks through an appropriate merge method, such as the majority of votes. Algorithm 1 shows the summary of the bagging code proposed by this work.

Algorithm 1 Bagging algorithm pseudo-code

- 1: Input: Dataset $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$:
 - 2: Number of learning rounds T .
 - 3: Process: For $t = 1, 2, \dots, T$:
 - 4: (i) Constitute sets bootstrap of S_t data selecting m random examples of the training set with substitution and (ii) Let h_t be the training base result of the algorithm based on S_t
 - 5: End.
 - 6: Output: Combined classifier: $H(x) = majority(h_1(x), \dots, h_T(x))$
-

In the boosting algorithm, similar to bagging, each classifier is trained using a different training set. The main difference concerning bagging method is that the re-sampled datasets are explicitly constructed to generate mutual learning, and the vote importance is weighted based on the performance of each model rather than the attribution of the same weight to all votes. Essentially, this procedure allows increasing the performance of an ar-

bitrary threshold by merely adding weaker learners. Given the usefulness of this finding, the boosting method is considered one of the most significant findings in ML. Algorithm 2 presents the pseudo-code of the boosting method.

Algorithm 2 Boosting algorithm pseudo-code

- 1: Input: Dataset $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$:
 - 2: Algorithm of learning base L ; Number of learning rounds T .
 - 3: Process: $D_1(i) = 1/m$. %Initializes the distribution of weights.
 - 4: For $t = 1, 2, \dots, T$: $h_t = L(D, D_t)$;
 - 5: Train the learning base h_t for D using the D_t distribution
 - 6: $\epsilon_t = Pr_{i \sim D_t} [h_t(x_i) \neq y_i]$;
 - 7: Measures the error of h_t
 - 8: $\alpha_t = \frac{1}{2} \ln \frac{1-\epsilon_t}{\epsilon_t}$
 - 9: Determines the weight of h_t
 - 10: $D_{t+1}(i) = \frac{D_t(i)}{Z_t} \begin{cases} \exp(-\alpha_t) & \text{if } h_t(x_i) = y_i \\ \exp(\alpha_t) & \text{if } h_t(x_i) \neq y_i \end{cases}$
 - 11: Updates the distribution
 - 12: $= \frac{D_t(i) \exp(-\alpha_t y_i h_t(x_i))}{Z_t}$ %Normalization factor allowing D_{t+1} to be a distribution
 - 13: End.
 - 14: Output: $H(x) = \text{sign} \left(\sum_{t=1}^T \alpha_t h_t(x) \right)$
-

The next section presents the performance evaluation of the proposed ensemble classifiers.

4. Performance Evaluation and Results Analysis

For this study, approval was obtained from the Research Ethics Committee of the Maternity School Assis Chateaubriand of the Federal University of Ceará, Fortaleza, CE, Brazil, under the certificate of presentation for ethical appreciation, number 66929317.0.0000.5050, and assent received with protocol number 2.036.062. This research considered 205 parturient women diagnosed with a hypertensive disorder during pregnancy.

This study considered for the performance evaluation of the ensemble classifiers the confusion matrix and the 10-fold cross-validation method [Deng et al. 2016, Wong 2015]. Table 1 presents the results of the proposed ensemble learning methods regarding the main evaluation indicators of the confusion matrix.

Table 1. Performance evaluation of ensemble learning classifiers concerning the indicators of the confusion matrix for the prediction of the Apgar score in the first 5 minutes of the newborn.

| Algorithm | ACC | TPR | FPR | AUC |
|------------------|--------|--------|--------|--------|
| Boosted Trees | 0.932 | 0.000 | 0.000* | 0.524 |
| Bagged Trees | 0.932 | 0.142 | 0.011 | 0.767 |
| Subspace KNN | 0.937* | 0.143 | 0.005 | 0.721 |
| RUSBoosted Trees | 0.790 | 0.786* | 0.209 | 0.816* |

Accuracy (ACC) represents the proportion of correct predictions without taking into account what is positive and what is negative. The true positive (TPR) rates measure

the ratio of true positives regarding all positive predictions. The false positive rate (FPR) represents the proportion of instances incorrectly classified as true, *i.e.* false alarms. The lower the FPR, the better the classifier is. The Boosted Trees classifier has achieved excellent FPR, however, a very low TPR, *i.e.*, this approach is superior in predicting when the condition is nonexistent. However, it cannot classify cases correctly when it is present. Its high accuracy rate is due to the fact that most of the observed instances present a normal Apgar score for the first 5 living minutes. Table 2 shows the performance evaluation of these same algorithms in relation to fetal growth, *i.e.*, whether it will be (or not) SGA.

Table 2. Performance evaluation of ensemble learning classifiers concerning the indicators of the confusion matrix for the prediction of intrauterine growth restriction.

| Algorithm | ACC | TPR | FPR | AUC |
|------------------|--------|--------|--------|--------|
| Boosted Trees | 0.829 | 0.156 | 0.046 | 0.621 |
| Bagged Trees | 0.854* | 0.125 | 0.012* | 0.598 |
| Subspace KNN | 0.829 | 0.063 | 0.029 | 0.669* |
| RUSBoosted Trees | 0.639 | 0.594* | 0.353 | 0.633 |

The presented results show that the algorithm based on the boosting method, *i.e.*, the Boosted Trees algorithm, presents an excellent accuracy, but a low TPR, which indicates randomness regarding prediction. Concerning FPR, this algorithm has excellent performance. The algorithm based on the bagging method (Bagged Trees) performs similarly. The Subspace KNN classifier performs well on almost all indicators however it also presents a low TPR. Differently, from the previous methods, this algorithm presents a better area under the curve (AUC) for both problems related to the fetus. This work considered K equal to 10 nearest neighbors. The RUSBoosted algorithm presents the most reasonable results with excellent TPR and AUC. However, this method has a high FPR, *i.e.*, it presents a high false alarm rate. Based on these results, the Subspace KNN method is considered by this research the most suitable classifier to predict problems that may lead to the fetus to developing severe complications.

The ensemble algorithms were also analyzed by the receiver operating characteristic (ROC) curve, which enabled the determination of the cut grade by the mathematical combination of higher specificity (1 - FPR) and sensitivity (TPR). Verification of accuracy, sensitivity, and specificity allow a better classification of problems related to fetal development by the definition of AUC. The AUC has five grades for the evaluation of predictive algorithms, namely excellent (0.9 to 1), good (0.8 to 0.9), reasonable (0.7 to 0.8), poor (from 0.6 to 0.7), and non-discriminatory (0.5 to 0.6). The Subspace KNN algorithm reached an AUC of 0.721 and 0.669 for predicting the Apgar score and intrauterine growth restriction, respectively, in newborns whose mothers suffered from some hypertensive disorder during pregnancy. Considering the difficulty of predicting complex conditions, these results are considered reasonable for this study. Figures 2 and 3 present the ROC curve for the ensemble learning method named Subspace KNN concerning both the Apgar score and SGA classes.

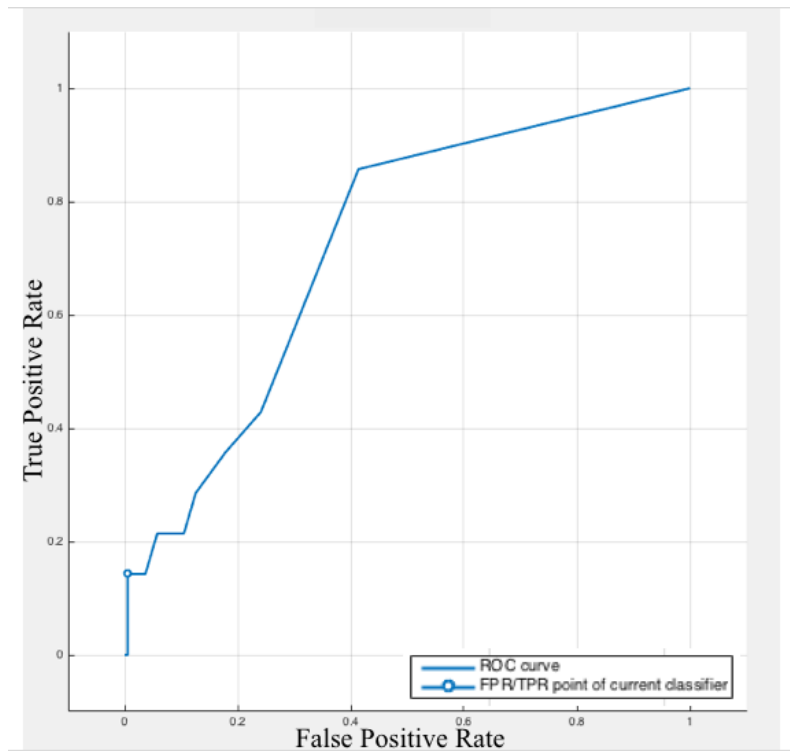


Figure 2. ROC curves for the ensemble learning method named Subspace KNN concerning Apgar score class.

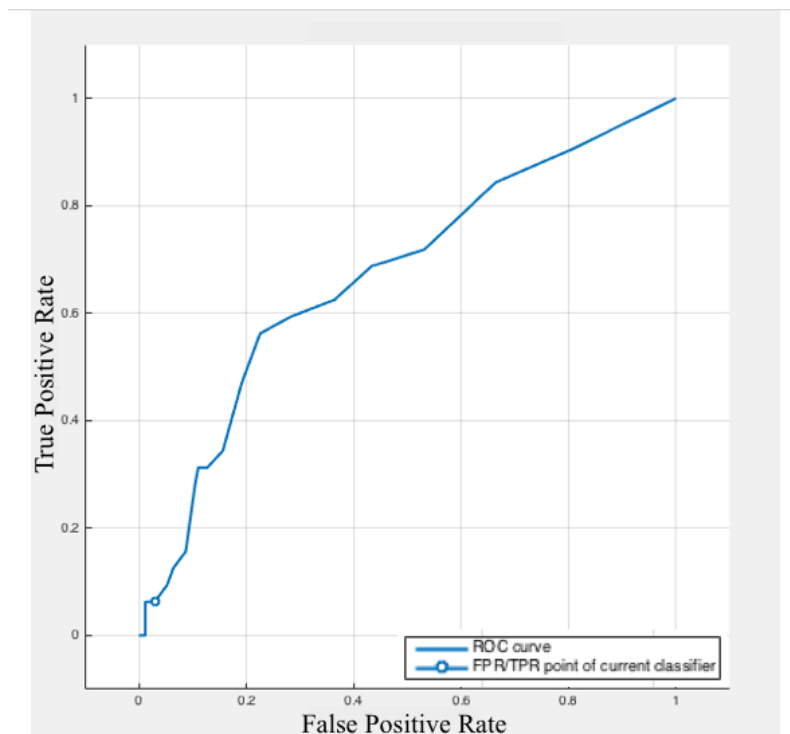


Figure 3. ROC curves for the ensemble learning method named Subspace KNN concerning SGA class.

5. Conclusion and Future Work

This study suggests that the optimal model of prediction of unfavorable neonatal outcome cannot yet be determined, but the prediction of the Apgar score and intrauterine growth restriction seem to be good predictors of death and severe morbidity in the neonatal period.

This paper aimed to present novel evolutionary approaches, called ensemble learning methods, for the classification of an unbalanced pregnancy database. The performance evaluation considered the AUC and other indicators of the well-known confusion matrix with the objective of evaluating a set of hypertensive disorder cases of the training base so that these samples generate classifiers with high accuracy and diversity. Finally, the methods resulting from the process compose an ensemble of classifiers to predict new examples by majority vote. The results obtained in the experiments showed that the resulting ensemble, called Subspace KNN, showed high predictive accuracy.

Further works propose the combination of other types of classifiers to decrease randomization in the classification, *i.e.*, increase the TPR. The study of other problems related to the fetus is also strongly suggested. This research is part of an effort to find a smart mechanism for e-health environments that is both robust and accurate enough for predicting problems that can lead both mother and baby to severe complications, including death, thus contributing to the reduction of the mortality and morbidity rates.

Acknowledgements

This work was supported by National Funding from the FCT - Fundação para a Ciência e a Tecnologia through the UID/EEA/50008/2013 Project; by Finep, with resources from Funttel, Grant No. 01.14.0231.00, under the Centro de Referência em Radiocomunicações - CRR project of the Instituto Nacional de Telecomunicações (Inatel), Brazil; by the Brazilian National Council for Research and Development (CNPq) via Grant No. 309335/2017-5; and by Ciência sem Fronteiras of CNPq, Brazil, through process number 207706/2014-0.

References

- Abalos, E., Cuesta, C., Carroli, G., et al. (2014). Pre-eclampsia, eclampsia and adverse maternal and perinatal outcomes: a secondary analysis of the World Health Organization multicountry survey on maternal and newborn health. *BJOG: An International Journal of Obstetrics & Gynaecology*, 121(1):14–24.
- Al-Fuqaha, A., Guizani, M., Mohammadi, M., Aledhari, M., and Ayyash, M. (2015). Internet of things: A survey on enabling technologies, protocols, and applications. *IEEE Communications Surveys & Tutorials*, 17(4):2347–2376.
- Andreu-Perez, J., Poon, C. C., Merrifield, R. D., Wong, S. T., and Yang, G.-Z. (2015). Big data for health. *IEEE Journal of Biomedical and Health Informatics*, 19(4):1193–1208.
- Cetin, I., Mazzocco, M. I., Giardini, V., et al. (2017). PlGF in a clinical setting of pregnancies at risk of preeclampsia and/or intrauterine growth restriction. *The Journal of Maternal-Fetal & Neonatal Medicine*, 30(2):144–149.

- Cnattingius, S. and Villamor, E. (2016). Weight change between successive pregnancies and risks of stillbirth and infant mortality: a nationwide cohort study. *The Lancet*, 387:558–565.
- Deng, X., Liu, Q., Deng, Y., and Mahadevan, S. (2016). An improved method to construct basic probability assignment based on the confusion matrix for classification problem. *Information Sciences*, 340:250–261.
- English, F. A., Kenny, L. C., and McCarthy, F. P. (2015). Risk factors and effective management of preeclampsia. *Integrated Blood Pressure Control*, 8:7–12.
- Gaffey, M. F., Das, J. K., and Bhutta, Z. A. (2015). Millennium Development Goals 4 and 5: Past and future progress. In *Seminars in Fetal and Neonatal Medicine*, volume 20, pages 285–292. Elsevier.
- Guan, D., Yuan, W., Lee, Y.-K., Najeebullah, K., and Rasel, M. K. (2014). A review of ensemble learning based feature selection. *IETE Technical Review*, 31(3):190–198.
- Huda, S., Yearwood, J., Jelinek, H. F., Hassan, M. M., Fortino, G., and Buckland, M. (2016). A hybrid feature selection with ensemble classification for imbalanced health-care data: A case study for brain tumor diagnosis. *IEEE Access*, 4:9145–9154.
- Iliodromiti, S., Mackay, D. F., Smith, G. C., Pell, J. P., and Nelson, S. M. (2014). Apgar score and the risk of cause-specific infant mortality: a population-based cohort study. *The Lancet*, 384:1749–1755.
- Krawczyk, B., Minku, L. L., Gama, J., Stefanowski, J., and Woźniak, M. (2017). Ensemble learning for data stream analysis: A survey. *Information Fusion*, 37:132–156.
- Nguyen, L., Bellucci, E., and Nguyen, L. T. (2014). Electronic health records implementation: an evaluation of information system impact and contingency factors. *International Journal of Medical Informatics*, 83(11):779–796.
- Ramírez, J., Górriz, J. M., Ortiz, A., Padilla, P., Martínez-Murcia, F. J., et al. (2016). Ensemble tree learning techniques for magnetic resonance image analysis. In *Innovation in Medicine and Healthcare 2015, September 11-12, Kyoto, Japan*, pages 395–404. Springer.
- Rathore, N., Tomar, D., and Agarwal, S. (2014). Predicting the survivability of breast cancer patients using ensemble approach. In *International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT), February 07-08, Ghaziabad, India*, pages 459–464. IEEE.
- Vogel, J., Souza, J., Mori, R., et al. (2014). Maternal complications and perinatal mortality: findings of the World Health Organization multicountry survey on maternal and newborn health. *BJOG: An International Journal of Obstetrics & Gynaecology*, 121(1):76–88.
- Wang, H., Liddell, C. A., Coates, M. M., et al. (2014). Global, regional, and national levels of neonatal, infant, and under-5 mortality during 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet*, 384:957–979.
- Weng, C.-H., Huang, T. C.-K., and Han, R.-P. (2016). Disease prediction with different types of neural network classifiers. *Telematics and Informatics*, 33(2):277–292.

- Wong, T.-T. (2015). Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. *Pattern Recognition*, 48(9):2839–2846.
- You, D., Hug, L., Ejdemyr, S., et al. (2015). Global, regional, and national levels and trends in under-5 mortality between 1990 and 2015, with scenario-based projections to 2030: a systematic analysis by the UN Inter-agency Group for Child Mortality Estimation. *The Lancet*, 386:2275–2286.