A neural network approach to predict early neonatal sepsis

Fernando López-Martínez, Edward Rolando Núñez-Valdez, Jaime Lorduy Gomez, Vicente García-Díaz, A neural network approach to predict early neonatal sepsis, Computers & Electrical Engineering, Volume 76,2019, Pages 379-388, ISSN 0045-7906, https://doi.org/10.1016/j.compeleceng.2019.04.015.

A neural network approach to predict early neonatal sepsis

Fernando López-Martínez[a,b,1,4,*], Edward Rolando Nuñez-Valdez[b,3], Jaime Lorduy Gomez[c,2], Vicente García-Díaz[b,3]

[a]350 Engle Street,Englewood Hospital, Englewood,NJ 07631, USA
[b]C/ Federico Garc´ıa Lorca, University of Oviedo, 33007, Oviedo, Spain
[c]Centro Calle de la Soledad N 5 -70, Cartagena, Colombia

Abstract

The purpose of this study is to develop a non-invasive neural network classification model for early neonatal sepsis detection. Early neonatal sepsis is a public health issue and one of the leading causes of complications and deaths in neonatal intensive care units. The data used in this study is from Crecer's Hospital center in Cartagena-Colombia. An imbalanced dataset of 555 neonates with (66%) of negative cases and (34%) of positive cases was used for this study. The study results show a sensitivity of 80.32%, a specificity of 90.4%, precision on the positive predicted value of 83.1% in the test sample and a calculated area under the curve of 92.5% (95% Confidence Interval[91.4 - 93.06]). This neural network model can be used as a smart system's inference engine to support the detection of neonatal sepsis in neonatal intensive care units.

Keywords: Machine Learning, Artificial Neural Networks, sepsis neonatal, Medical Decision Support Systems, Smart Systems.

*Corresponding author
Email addresses: uo259897@uniovi.es (Fernando L´opez-Mart´ınez), nunezedward@uniovi.es (Edward Rolando Nu´n˜ez-Valdez), jaime.lorduy@curnvirtual.edu.co (Jaime Lorduy Gomez), garciavicente@uniovi.es (Vicente Garc´ıa-D´ıaz)
[1]Student. Department of Computer Science, Oviedo University. Phone: +1 5515870112
[2]Epidemiologist. Rafael Nun˜ez University. Phone: +57 3172672845
[3]Lecturer. Department of Computer Science, Oviedo University. +34 985103326
[4]Senior Systems Analyst, Englewood Hospital. Phone: +1 2018945663

1. Introduction

Neonatal sepsis is a severe problem for neonates, and it is estimated to affect more than 3 million newborns worldwide every year [1]. World Health Organization estimates that one in ten deaths associated with pregnancy and childbirth is due to maternal sepsis with over 95% of deaths due to maternal sepsis occurring in low and middle-income countries.

Early neonatal sepsis is defined as the presence of infection proven by a blood or cerebrospinal fluid (CSF) culture positive for bacteria and fungi, and viruses manifested in the first 72 hours of birth [2] in pre-term. Despite the advances in antibiotic therapy and the awareness of risk factors in neonatal intensive care units, neonatal sepsis continues to be a severe complication and cause of severe illness and deaths of hospitalized neonates. The amount of data routinely collected in electronic medical records (EMR) and bed-monitors allow us to obtain quality data to build useful predicting models and tools like smart systems to generate continuous risk-assessments for neonatal sepsis from big clinical data and machine learning to produce an earlier diagnosis and improve sepsis management in neonates. Early diagnosis has been shown to reduce delays in treatment, increase appropriate care and reduce mortality [3]. The development of smart systems that can gather a multitude of physiological measurements such as blood pressure, ECG, EEG, skin conductance, respiration and heart rate, plus all the information extracted from EMR systems, has a significant potential to perform data analysis and detect life-threatening symptoms in infants to prevent early neonatal sepsis.

We present a non-invasive prediction model that can be used as an inference engine of a smart system to provide decision support for health care providers at neonatal intensive care units to provide antibiotic administration when sepsis is detected.

The structure of this paper is as follows: Section 2 presents related work and literature research of several models that used neural networks for neonatal sepsis prediction. In section 3, we describe the development of the multi-layer

model, describing the population, data source, and validation of the model. Section 4 presents the statistical and clinical analysis of the outcomes. Section 5 discusses the results of the model and its limitations. Finally, section 6 covers conclusions and future work.

2. Related work

Several papers were reviewed that used artificial neural networks to predict sepsis neonatal. In every article, we analyzed the model building process, variable selection, ground truth, training and test datasets, overfitting avoidance, error estimate, and AUC (Area Under the Curve) information.

Subramani [4]. Presents several machine learning models including support vector machines (SVM), naive Bayes classifier (NB), tree augmented naive Bayes (TAN), averaged one-dependence estimators (AODE), K nearest neighbor (K-NN), decision tree classifier and regression trees (CART), random forest (RF), logistic regression (LR) and Lazy Bayesian Rules (LBR). The dataset used consisted of 299 infants evaluated for late-onset sepsis. Several feature selection algorithms were used to select highly predictive features including SVM - (forward, backward, forward-backward and recursive), HITON Markov blanket and HITON - parents and children algorithms. This study reported an AUC of 78% for naive Bayes.

Griffin [5]. Shows that the clinical diagnosis of neonatal sepsis is preceded by abnormal heart rate characteristics (HRC). This study reported an AUC of 82% for sepsis prediction using multivariable logistic regression. The dataset used consisted of 678 infants.

Honor´e [6]. Presented a shallow feed-forward Neural Network model with 30 hidden nodes that used an imbalanced dataset composed by heart frequency signals and SpO_2 signals after applying frame normalization and removed the mean of every signal. This study reported an AUC of 85% for sepsis prediction, but the author expressed that this model is based on the inaccurate modeling of a deficient number of training examples.

Calvert [7]. Develop high-performance early sepsis prediction technology for the general patient population. This model reported an average AUC of 92%. This model used nine vital sign variables, systolic blood pressure, pulse pressure, heart rate, temperature, respiration rate, white blood cell count, pH, blood oxygen saturation, and age.

Desautels [8]. Applied the InSight, machine learning classification model developed by Calvert, and use combinations of patient data such as vitals, peripheral capillary oxygen saturation, Glasgow Coma Score, and age. This model reported an average AUC of 88%.

Horng [9]. This study presents a model that include free text, vital signs, and demographic data to identify patients with sepsis at the emergency department. This model reported an average AUC of 85%.

Kam [10]. This study presents a detection model using deep learning methodologies and compares it with conventional regression methods. This model reported an average AUC of 92.9%, and it is the baseline for our study due to the similarity of the data used in their research and the neural network architecture.

Our study shows several improvements concerning the studies presented above, first of all, the type of risk variables included in our study such as sociodemographic, obstetric, neonatal and maternal infectious related variables were not used at the same time in the variable selection criteria in the other studies. Our classification results suggested that the combinatory use of these variables and the proposed ANN (artificial neural network) is a potentially useful clinical model to classify the neonatal sepsis. Second, our study shows a better calculated AUC than most of the work described before due to the performance of the ANN architecture even with a highly unbalanced dataset, weighting the samples from the underrepresented class shows better results than other proposed architectures. Third, our study shows different reporting metrics such as a confusion matrix and classification report that are not present in the related work in addition to the AUC, accuracy, and precision. Finally, our classification model used a non-invasive approach for the classification of neonatal sepsis and was found to have a better overall performance compared with other models.

3. Model Development and Validation

We developed a deep learning classification model in this study and will discuss and analyze how we implemented, trained and evaluated the model. The package Sklearn of Python programing language [11], Microsoft Cognitive Toolkit (CNTK) from Microsoft and AzureML were used to build the model.

3.1. Data Source

Our model used an anonymous dataset from a private institution in the city of Cartagena, Colombia from 2016 to 2017. Demographic, laboratory data, blood pressure, and body measures data were part of the dataset. The dataset includes cases of live newborns of ages inferior to 72 hours with a diagnosis of early neonatal sepsis by clinical criteria and laboratory blood cultures. Control cases were part of the dataset including all newborns healthy by clinical diagnosis and who have returned healthy for a follow up at 72 hours. This data was used under resolution 008430, title II, chapter 1, article 11 of the Colombian health department.

3.2. Study Population and Analysis

This retrospective study includes 186 cases and 368 controls based on a case-control relationship of 1:2 with a 95% of the trust factor, power of 80% and the application of Yates correction [12]. Bivariate analysis and logistic regression were performed to detect the variable associated with early sepsis, and the significant statistical association used was $p \leq 0.05$.

3.3. Input variables

This study considered nine sociodemographic, fourteen obstetric, nine neonatal and four maternal infectious related pathology variables. Table 1 shows the quantitative sociodemographic variables, table 2 shows the qualitative sociodemographic variables, table 3 shows the quantitative neonatal variables, table 4 shows the qualitative neonatal variables, table 5 shows the quantitative obstetric variables, table 6 shows the qualitative obstetric variables and table 7 shows the qualitative maternal infections of the cases and controls.

Table 1: Quantitative socio demographic variables in cases (186) and controls (369)

Quantitative Socio demographic variable	Cases				Controls				
	mean	median	SD	RIQ	mean	median	SD	RIQ	p-value
Age	23.93	23.5	4.99	20-26	24.22	23	6.19	19-28	0.793
Onset of sexual activity	16.06	16	0.945	15-17	15.6	16	0.971	15-16	0.0001

Table 2: Qualitative socio demographic variables in cases (186) and controls (369)

Qualitative Socio demographic variable	Categories	Cases		Controls		X2	p-value
		N	%	N	%		
Teen Mother	Yes	15	8.1	69	18.7	10.88	0.001
	No	171	91.9	300	81.3		
Health Regimen	Government	183	98.4	349	94.6	4.51	0.041
	Commercial	3	1.6	20	5.4		
Origin	Rural	42	22.6	5	1.4	71.87	0.00001
	Urban	144	77.4	364	98.6		
Marital Status	Married or in common law married	128	68.8	101	27.4	87.64	0.00001
	Single, divorced or widow	58	31.2	268	72.6		
Level of education	Elementary School	86	46.2	80	21.7	35.57	0.00001
	High School	100	53.8	289	78.3		
Start of Marital status life younger than 18 yo	Yes	178	95.7	357	96.7	0.39	0.531
	No	8	4.3	12	3.3		
Start of Marital status life younger than 16 yo	Yes	47	25.3	147	39.8	11.54	0.001
	No	139	74.7	222	60.2		

Table 3: Quantitative Neonatal variables in cases (186) and controls (369)

Quantitative Neonatal variable	Cases				Controls				
	mean	Median	SD	RIQ	mean	Median	SD	RIQ	p-value
New born weight in grams	2639.9	2768.5	546.5	2500-3020	3202.4	3224	412.1	2950-3500	0.0001
APGAR after 1 minute of birth	7.73	8.0	0.611	8.0	8.09	8.0	0.598	8.0	0.0001

3.4. Variables selection

In our study, a bivariate analysis chi-square with correction was performed to the qualitative variables to find statistically association between our independent variable and the possibility to develop early neonatal sepsis. For the quantitative variables, the Mann–Whitney U test was performed. From this statistical analysis it is essential to show that we didn't find significant statistical evidence for the variables Age, start of the marital status life younger than 18 years old, gender, APGAR (Appearance, Pulse, Grimace, Activity, and Respiration) value less than 7 after 1 minute and after 5 minutes, number of pregnancies, the type of birth. Prenatal control is not associated with the case of sepsis; however, assisting to 5 prenatal controls are associated with the pro-

Table 4: Qualitative Neonatal variables in cases (186) and controls (369)

Qualitative Neonatal variable	Categories	Cases		Controls		X2	p-value
		N	%	N	%		
Premature	Yes	100	53.8	25	6.8	156.4	0.0001
	No	86	46.2	344	93.2		
Gender	Male	109	58.6	202	54.7	0.748	1.672
	Female	77	41.4	167	45.3		
Less than 1500 grams	Yes	11	5.9	2	0.5	15.6	0.00001
	No	175	94.1	367	99.5		
Less than 2500 grams	Yes	44	23.7	9	2.4	64.44	0.00001
	No	142	76.3	360	97.6		
APGAR less than 7 after 1 minute of birth	Yes	2	1.1	3	0.8	0.095	0.999
	No	184	98.9	366	99.2		
APGAR less than 7 after 5 minutes	Yes	4	2.2	9	2.4	0.045	0.999
	No	182	97.8	360	97.6		
Respiratory distress	Yes	89	47.8	27	7.3	122.8	0.0001
	No	97	52.2	342	92.7		

Table 5: Quantitative Obstetric variables in cases (186) and controls (369)

Quantitative Obstetric variable	Cases				Controls				
	mean	Median	SD	RIQ	mean	Median	SD	RIQ	p-value
Gestational age at the time of birth	35.6	36.0	3.47	34-39	38.4	39.0	1.62	38-39	0.0001
Number of prenatal controls	4.08	5.0	1.83	3.75-5.0	4.32	5.0	1.83	4-5.0	0.002
Number of pregnacies	1.77	1.0	1.15	1.0-2.0	1.6	1.0	1.15	1-2.0	0.076
Number of births	1.04	1.0	1.03	0-1	0.7	1.0	1.03	0-1	0.0001
Numbers of C-sections	0.65	1.0	0.68	0-1	0.76	1.0	0.68	0-1	0.029

tection to avoid the appearance of early neonatal sepsis as shown in our previous study [2]. There was no evidence with the variables IUGR (Intrauterine growth restriction) background and multiple pregnancies. Twenty-seven (27) variables were selected as input variables for our artificial neural network architecture.

3.5. Neural Network Model

A neural network is a class of machine learning algorithms that receives input from one or more sources into layers of intermediate nodes which are connected (hidden layers [13]) to build a network that generates output values [14]. Figure 1 shows a sample of the structure of the artificial neural network architecture used in our study. The input nodes received categorical and numerical values

Table 6: Qualitative Obstetric variables in cases (186) and controls (369)

Qualitative Obstetric variable	Categories	Cases		Controls		X2	p-value
		N	%	N	%		
Type of birth	Vaginal	98	52.7	162	43.9	3.833	0.05
	C-Section	88	47.3	207	56.1		
IUGR Background	Yes	5	2.7	13	3.5	0.275	0.6
	No	181	97.3	356	96.5		
Assistance for prenatal control	Yes	165	88.7	318	86.2	0.702	0.402
	No	21	11.3	51	13.8		
Assistance for at least 4 prenatal control	Yes	140	75.3	301	81.6	3.01	0.083
	No	46	24.7	68	18.4		
Assistance for at least 5 prenatal control	Yes	105	56.5	254	68.8	8.301	0.004
	No	81	43.5	115	31.2		
Premature rupture of membrane with more than 18 hours	Yes	95	51.1	17	4.6	165.7	0.00001
	No	91	48.9	352	95.4		
Chorioamnionitis	Yes	23	12.4	3	0.8	36.96	0.00001
	No	163	87.6	366	99.2		
Premature membrane rupture with more than 6 hours	Yes	161	86.6	194	52.6	61.96	0.0001
	No	25	13.4	175	47.4		
Multiple Pregnacies	Yes	2	1.1	10	2.7	0.39	0.353
	No	184	98.9	359	97.3		

Table 7: Qualitative maternal infections variables in cases (186) and controls (369)

Qualitative Maternal infections variables	Categories	Cases		Controls		X2	p-value
		N	%	N	%		
Maternal Fever	Yes	67	36.0	40	10.8	50.38	0.0001
	No	119	64.0	329	89.2		
Yeast Infections	Yes	31	16.7	15	4.1	25.83	0.0001
	No	155	83.3	354	95.9		
Sexualy transmitted disease history	Yes	27	14.5	7	1.9	34.24	0.0001
	No	159	85.5	362	98.1		
Urinary Tract Infections	Yes	11	5.9	9	2.4	4.29	0.0381
	No	175	94.1	360	97.6		

that will be encoded and normalized with gaussian normalization [15] to improve the computation of the network and to prevent the data from being distant from each other.

This neural network extract features and learn high-level new features from the data, but to learn these new features, the design of the architecture is fundamental [16]. Three fully connected hidden layers in addition to the input layer and output layer were added to the network to generate our multi-layer architecture. The correct selection of the number of hidden layers and hidden

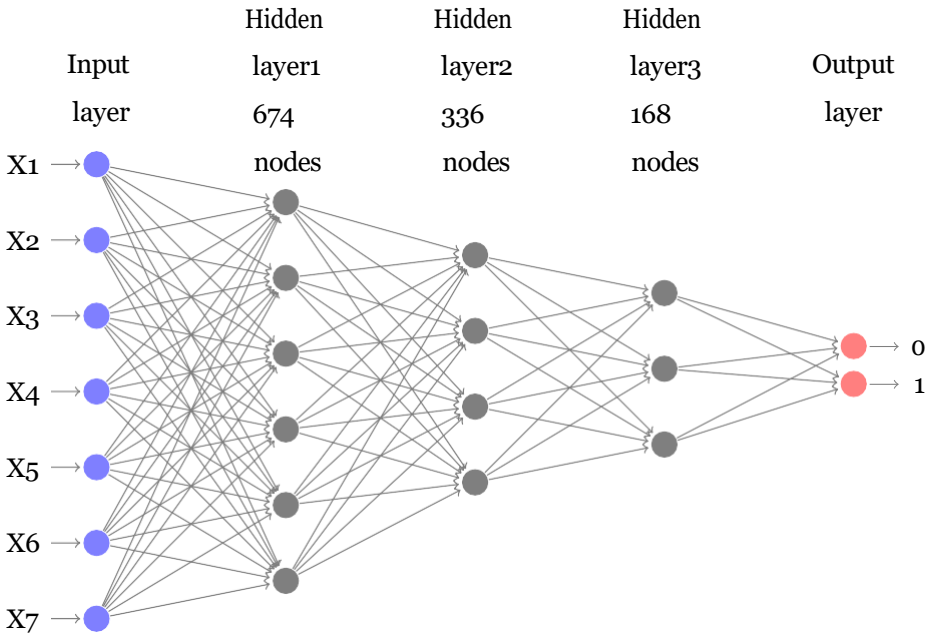

Figure 1: Multilayer Perceptron Architecture

nodes is an active area of research. For our study, we used a back propagation approach and three layers were enough so our network can generalize as a universal approximation for the mathematical function that can separate our input variables to solve the classification problem of arbitrary complexity and preventing overfitting.

We can define a simple neural network with the first layer taking an input matrix X with dimensions m number of samples in the dataset and n_x input size ($X.shape = (n_x, m)$), and produces a linear output (first hidden layer Z^1 with dimension n. Each feature in the input layer is connected with a node in the output layer by the weight which is represented by a matrix W with dimensions $n \times n_x$:

$$Z^1 = W \cdot X + b$$

Where b is a bias vector of dimension n and ($b.shape = (n, 1)$).

In the architecture of our network, each layer has its weight matrix W, its

bias vector b, and an output vector a with activation function f. Superscripts appended to these variables will identify the layers. For R inputs, Z^1 represents the number of neurons in the first layer and Z^2 neurons in the second layer. The output layer for our model is the layer 4 with three hidden layers 1,2 and 3. Then, the output of layer 1 is the input for layer 2, and the output of layer 2 is the input for layer 3, and the output of layer 3 is the input for the output layer.

if:

$$a^1 = f^1(W^1X + b^1)$$

$$a^2 = f^2(W^2a^1 + b^2)$$

$$a^3 = f^3(W^3a^2 + b^3)$$

$$a^4 = f^4(W^4a^3 + b^4)$$

then,

$$a^4 = predicted$$

The structure of our network can be expressed in a simple way as:

$$R \rightarrow Z^1 \rightarrow Z^2 \rightarrow Z^3 \rightarrow a^4$$

Each layer will be a dense layer with a specific input dimension, output dimensions, and activation function. Specifically, in our model, the first layer has an input dimension of 27 variables (the figure only shows 7), an output dimension of 674 nodes, and activation function being ReLU (Rectified Linear Unit activation function) [17]. The second layer having an input dimension of 674 nodes, an output dimension of 332 nodes, and activation function being also ReLU, the third layer has an input of 168 nodes, an output of 168 nodes and activation function being also ReLU. Furthermore, the output layer with no activation. The final output layer emits a vector of two values, and we used Softmax [18] to normalize the output of the model and to map the accumulated evidence or activations to a probability distribution over the classes.

Softmax is used in the model to interpret the output as a probability distribution over a set of classes C, where a vector z is mapped to a vector of probabilities p that sum up to 1 and can be interpreted as probabilities for mutually exclusive outcomes.

$$p_i = Softmax(z, i) = \frac{\exp(z_i)}{\sum_{k \in C} \exp(z_k)}$$

Where k goes over all the possible values of C, so k would go from 1 to C.

The vector of probabilities can be expressed at $z = x$ as:

$$\text{softmax}(x_i) = \left[\frac{\exp(x_i)}{\sum_k \exp(x_k)} \quad \frac{\exp(x_i)}{\sum_k \exp(x_k)} \quad \cdots \quad \frac{\exp(x_i)}{\sum_k \exp(x_k)} \right]^1$$

3.6. Network Architecture and hyper-parameters selection

Besides the logistic regression model used in a previous study, and explained in [2], decision trees was another data modeling technique considered to evaluate how the independent variables affect the responses of the dependent variable in the model for early diagnosis of sepsis neonatal. However, we decided in our study to use artificial neural networks due to useful capabilities, including online learning with incremental updates, better representation of non-linear interaction between the independent variables, and better performance than exhaustive search on noisy data.

For the architecture selection, we started with a simple two-layer perceptron and some hidden units matching the number of nodes and leaves of our initial decision tree approach. This first approach could not capture the significant non-linearity of the data. Then, we constructed a sequence of networks with increasing numbers of hidden layers and nodes. All these networks were trained on a different set of hyper-parameters and normalized data using Gaussian normalization, decreasing the training and testing errors in each network until we found the architecture with a better fit and better generalization for a final number of hidden layers and nodes. As a result, our best architecture presents 3 hidden layers, with dimensions of 674, 336 and 168 hidden nodes respectively. The minibatch size of 8 samples injects enough noise to each gradient update

allowing our network to achieve relatively fast convergence in combination with the selection of sgd with momentum of 0.9, and a learning rate of 0.01 where our network converges to something useful to improve its performance. Finally, Softmax in the output layer allows us to interpret the output as probabilities, and the cross-entropy loss was used to measure the error at this softmax layer. We have published the model architecture and dataset in the gallery of Microsoft Azure Machine Learning Studio for model replication [19].

3.7. Model Training

For training our model, we would like the generated probabilities to be as close as possible to the observed labels. We calculated the cost function which is the difference between the learned model versus the generated by the training set. We used cross-entropy with Softmax in our study, but the computation of the derivative of Softmax and the backward functions are beyond the scope of this article. However, the cross-entropy loss formula can be expressed as in 1.

$$crossEnt(Y, P) = -\sum_{k=1}^{C} Y(k) log(P(k)) \quad (1)$$

Where k goes over all the possible output classes, specifically, in our case, there are C output classes, so k would go from 1 to C. P are the Softmax output, and Y is the correct classification output. $P(k)$ is the probability of the class as predicted by the model and $Y(k)$ is the correct probability of the class provided by the ground-truth labels. If we make $Y(k) = 1$ and for all $k \neq y$ we have $Y(k) = 0$, and we express the formula in function of just P treating y as a constant and $P(y)$ as the $y - th$ element of P. The Softmax and cross-entropy loss can be expressed as in 2.

$$crossEnt(P) = -log(P_y) \quad (2)$$

Once defined the cost function, the model minimizes it using an optimization technique. In this study, we used stochastic gradient descent with momentum

[20]. The model starts with random initialization of the parameters and generates a new set of parameters after each evaluation. In our implementation, we used He initialization [21]. Which is similar to Xavier initialization except Xavier uses a different scaling factor for the weights W in layer l, and the author recommends for layers with ReLU activation. We used mini-batches to train our model, a small number of observations were loaded into the model to calculate the average of the loss to update the model parameters. Another critical parameter used in our training was the learning rate [22]. This learning rate is a factor that moderates how much we change the parameters in each iteration. Each iteration will work on eight samples, and we trained the model with 70% (388) of the dataset, and the number of mini-batches to train is defined by the number of samples to train divided by the mini batch size. Table 8 shows the parameters of the trainer architecture.

3.8. Model Evaluation

In our study, we evaluated the accuracy, sensitivity, specificity, and AUC in early neonatal sepsis detection when using an artificial neural network. The softmax layer finds the index of the highest value in the output array and compares it to the actual ground truth label. We evaluated the trained network on data that has not been used for training and corresponds to the ground truth of 30% (166) of the dataset. The resulting error is comparable to the training error, and this indicates that our model has a proper generalization error.

We generated some evaluation metrics to evaluate the classifier. Table 9, shows the confusion matrix with a classification results summary of the actual class label vs. the predicted ones. True positive value (49), True Negative value (95), False Negative (12) and False Positive value (10).

Classification report showed in table 10 shows the precision, sensitivity and the harmonic mean of precision and sensitivity. The sensitivity of the model moderately acceptable due to the imbalanced testing dataset, and there is still a high number of false negatives.

Table 8: Model Architecture parameters

Model Architecture parameters	
Parameter	**Value**
Input Dimension	27
Num Output classes	2
Num Hidden Layers	3
Hidden Layer1 Dimension	674
Activation Func Layer1	Relu
Hidden Layer2 Dimension	336
Activation Func Layer2	Relu
Hidden Layer3 Dimension	168
Activation Func Layer3	Relu
Minibatch size	8
Num samples to train	388
Num minibatches to train	48
Loss Function	cross entropy with softmax
Eval Error	Classification error
Learner for parameters	momentum sgd
Eval Metrics	Confusion Matrix, AUC

Table 9: Confusion Matrix

		Predicted	
		Non-Sepsis	**Sepsis**
True	**Non-Sepsis**	95	10
	Sepsis	12	49

4. Results and Analysis

Statistical and clinical analysis is essential to explain the results and usefulness of the model. In the next sections, we will review some of the results.

Table 10: Classification Report

Classification Report			
True Positive	**False Negative**	**Precision**	**Accuracy**
49	12	0.83	0.867
False Positive	**True Negative**	**Recall**	**f1-score**
10	95	0.803	0.817
Positive Label: 1		**Negative Label: 0**	

4.1. Statistical Analysis

An artificial neural network is a class of non-linear model used to identify the unknown non-linearity of the input variables. Our test sampling of 166[a] includes 105 (63,25%) negative examples and 61 (36,75%) positive examples. The model produces a sensitivity of 49/61 = 80.32% (positives that were correctly identified) and a specificity of 95/105 = 90.4% (negatives that are correctly identified).

The positive predicted value was 49/59 = 83% and the negative predicted value 95/107 =88.7%. The false negative rate of the model was 12/61 = 19.6% and a false positive rate of 10/105 = 9.5%. The probability of false alarm for the model was 1-0.904(specificity) = 9.6%. The miss rate and the fall out rate were not significant for our model due to the imbalance of the sampling data set. Figure 2 shows the area under the curve plotted with a true positive rate on the y-axis and false positive rate on the x-axis. Figure 3 shows the plot of the proportion of the true results of overall positives results versus the fraction of all correct results returned by the model.

4.2. Clinical Analysis

A sensitivity of 80.3% and a specificity of 90.4% shows that our model might be useful for detecting positive cases, and the true negative rate shows that our model is also efficient at identifying negative cases. The high precision value of 83.1% and the AUC of 92.5% confirm the adequacy of the model as a preliminary screening tool. The percentage of positive cases shows that our

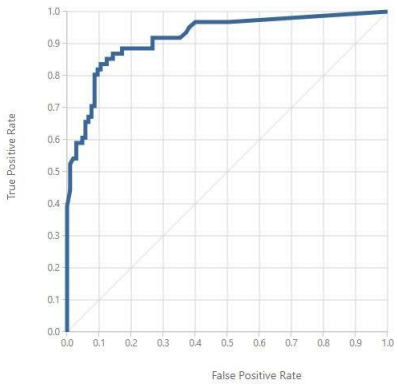

Figure 2: ROC Curve

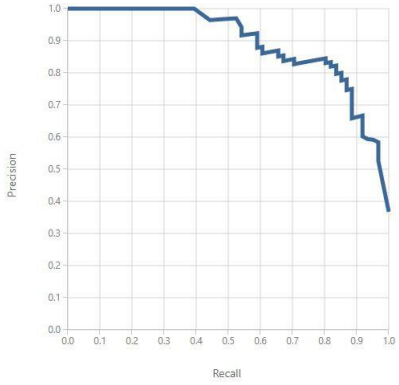

Figure 3: Precision/Recall

model works better than random guessing and the conditional probability of negative test results is considerably low. The accuracy of 86.74% shows that our model correctly identifies negative cases and positive cases based on the characteristics of the dataset and the small number of cases examined.

5. Discussion and Limitations

This neural network model can be used as a screening and decision support tool and not as a stand-alone decision-making inference engine for either a smart system or an expert system. Our model may not be directly applicable to other NICU (Newborn Intensive Care Unit) datasets, but we think that our study supports the potential for applying neural network architectures in assisting NICU care facilities in the management and treatment of neonatal sepsis. The AUC of our model was slightly higher than our previous developed logistic regression model [2] and showed improvement in early neonatal sepsis detection. Our model has limitations of data size, and the imbalanced dataset used caused a limited generalization.

In future studies, it is necessary to include a more complicated set of variables related to sepsis diagnosis, train the model with a more extensive and balanced dataset, and to explore different and more advanced neural network

architectures, transfer learning, recurrent neural networks [23] and fuzzy neural networks [24] to improve the detection early neonatal sepsis. Also, principal component analysis [25] can be used to reduce de dimensionality of the dataset. Finally, there is a need to perform cost-benefit studies concerning the implementation of this model and programs to detect and treat infections in pregnant women before 20 weeks of gestation, and this has shown to reduce low birth weight in preterm infants and premature birth.

6. Conclusions and Future work

Early identification of neonatal sepsis would allow physicians to implement treatments, determine proper antibiotic administration, and potentially reduce associated complications for neonates at Neonatal Intensive Care Unit. The use of data extracted from the electronic medical records allowed us to create a model with good performance and results when compared with others that used more complex data such as bio-signal data, laboratory results of blood cul- ture, electrocardiogram data and pulse oximeter data. Our study demonstrates the use of neural network models that learn features and make predictions for the detection of early neonatal sepsis. However, this study also indicates that such models have some limitations in setting the dependent variable, having a sufficient amount of data, and adequate explanatory power.

Mothers with premature rupture of membrane, maternal fever, and prema- ture newborn make an evident causal association for early neonatal sepsis. Level of education and marital status show significant evidence in the appearance of neonatal sepsis. Presence of maternal infectious pathology, such as vaginal infection was a determining factor to explain the cases of premature membrane rupture over the 18 hours.

Acknowledgments

This document presents an independent study funded by the company SYS-DEVELOPMENT LLC. The points of view expressed are those of the authorsand not necessarily those of Rafael Nuñez University, CRECER hospital or the department of health in Cartagena - Colombia. We thank Stephanye Carrillo González and Ruby Elena Muñoz Baldiris who provided insight and expertise that greatly assisted the study.

References

1. Fleischmann-Struzek C, Goldfarb DM, Schlattmann P, Schlapbach LJ, Reinhart K, Kissoon N. The global burden of paediatric and neonatal sepsis: a systematic review. 2018. URL: https://www.sciencedirect.com/science/article/pii/S2213260018300638?via{%}3Dihub. doi:10.1016/S2213-2600(18)30063-8.

2. Lorduy GJ, González SC, Muñoz Baldiris RE, D´ıaz-P´erez A, Perez I. Prognostic Factors of Early Neonatal Sepsis in the City of Cartagena Colombia. *Global Journal of Health Science* 2018;10(12):30. URL: http://www.ccsenet.org/journal/index.php/gjhs/article/view/0/37319. doi:10.5539/gjhs.v10n12p30.

3. Nguyen HB, Corbett SW, Steele R, Banta J, Clark RT, Hayes SR, Edwards J, Cho TW, Wittlake WA. Implementation of a bundle of quality indicators for the early management of severe sepsis and septic shock is associated with decreased mortality*. *Critical Care Medicine* 2007;35(4):1105–12. URL: http://www.ncbi.nlm.nih.gov/pubmed/17334251https://insights.ovid.com/crossref?an=00003246-200704000-00016. doi:10.1097/01.CCM.0000259463.33848.3D.

4. Mani S, Ozdas A, Aliferis C, Varol HA, Chen Q, Carnevale R, Chen Y, Romano-Keeler J, Nian H, Weitkamp JH. Medical decision support using machine learning for early detection of late-onset neonatal sepsis. *Journal of the American Medical Informatics Association : JAMIA* 2014;21(2):326–36. URL: http://www.ncbi.nlm.nih.gov/pubmed/24043317http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3932458. doi:10.1136/amiajnl-2013-001854.

5. Griffin MP, Lake DE, Moorman JR. Heart Rate Characteristics and Laboratory Tests in Neonatal Sepsis. *PEDIATRICS* 2005;115(4):937–41. URL: http://www.ncbi.nlm.nih.gov/pubmed/15805367http://

pediatrics.aappublications.org/cgi/doi/10.1542/peds.2004-1393. doi:10.1542/peds.2004-1393.

6. Honor\'e; A. Machine Learning for Neonatal Early Warning Signs. Master's thesis; KTH, Information Science and Engineering; 2017.

7. Calvert JS, Price DA, Chettipally UK, Barton CW, Feldman MD, Hoffman JL, Jay M, Das R. A computational approach to early sepsis detection. *Computers in Biology and Medicine* 2016;74:69–73. URL: https://www.sciencedirect.com/science/article/pii/S0010482516301123. doi:10.1016/J.COMPBIOMED.2016.05.003.

8. Desautels T, Calvert J, Hoffman J, Jay M, Kerem Y, Shieh L, Shimabukuro D, Chettipally U, Feldman MD, Barton C, Wales DJ, Das R. Prediction of Sepsis in the Intensive Care Unit With Minimal Electronic Health Record Data: A Machine Learning Approach. *JMIR medical informatics* 2016;4(3):e28. URL: http://www.ncbi.nlm.nih.gov/pubmed/27694098http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5065680. doi:10.2196/medinform.5909.

9. Horng S, Sontag DA, Halpern Y, Jernite Y, Shapiro NI, Nathanson LA. Creating an automated trigger for sepsis clinical decision support at emergency department triage using machine learning. *PLOS ONE* 2017;12(4):e0174708. URL: https://dx.plos.org/10.1371/journal.pone.0174708. doi:10.1371/journal.pone.0174708.

10. Kam HJ, Kim HY. Learning representations for the early detection of sepsis with deep neural networks. *Computers in Biology and Medicine* 2017;89:248–55. URL: https://www.sciencedirect.com/science/article/pii/S0010482517302743. doi:10.1016/J.COMPBIOMED.2017.08.015.

11. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J,

Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay É . Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 2012;12:2825–30. URL: http://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdfhttp://dl.acm.org/citation.cfm?id=2078195{%}5Cnhttp://arxiv.org/abs/1201.0490. doi:10.1007/s13398-014-0173-7.2. arXiv:1201.0490.

12. Lorduy Gómez J, González Carrillo S. ReviAsociación de factores obst´etricos y neonatales con casos de sepsis neonatal temprana. Cartagena, Colombiasta habanera de ciencias m´edicas.; vol. 17. La Habana: Instituto Superior de Ciencias M´edicas de La Habana; 2018. URL: http://www.revhabanera.sld.cu/index.php/rhab/article/view/2401.

13. Huang GB, Zhu QY, Siew CK, Ã GbH, Zhu QY, Siew CK, Huang GB, Zhu QY, Siew CK. Extreme learning machine: Theory and applications. *Neurocomputing* 2006;70(1-3):489–501. URL: https://www.sciencedirect.com/science/article/pii/S0925231206000385http://linkinghub.elsevier.com/retrieve/pii/S0925231206000385. doi:10.1016/j.neucom.2005.12.126. arXiv:1311.4555.

14. Ching T, Himmelstein DS, Beaulieu-Jones BK, Kalinin AA, Do BT, Way GP, Ferrero E, Agapow PM, Zietz M, Hoffman MM, Xie W, Rosen GL, Lengerich BJ, Israeli J, Lanchantin J, Woloszynek S, Carpenter AE, Shrikumar A, Xu J, Cofer EM, Lavender CA, Turaga SC, Alexandari AM, Lu Z, Harris DJ, Decaprio D, Qi Y, Kundaje A, Peng Y, Wiley LK, Segler MH, Boca SM, Swamidass SJ, Huang A, Gitter A, Greene CS. Opportunities and obstacles for deep learning in biology and medicine. *Journal of the Royal Society Interface* 2018;15(141). URL: http://www.ncbi.nlm.nih.gov/pubmed/29618526http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5938574. doi:10.1098/rsif.2017.0387. arXiv:142760.

15. Jain S, Shukla S, Wadhvani R. Dynamic selection of normalization

techniques using data complexity measures. *Expert Systems with Applications* 2018;106:252–62. URL: https://www.sciencedirect.com/science/article/pii/S095741741830232X. doi:10.1016/j.eswa.2018.04.008.

16. Abbasi R, Esmaeilpour M. Selecting Statistical Characteristics of Brain Signals to Detect Epileptic Seizures using Discrete Wavelet Transform and Perceptron Neural Network. *International Journal of Interactive Multimedia and Artificial Intelligence* 2017;4(5):33. URL: http://www.ijimai.org/journal/sites/default/files/files/2016/12/ijimai20174{_}5{_}6{_}pdf{_}12024.pdfhttp://www.ijimai.org/journal/node/1496. doi:10.9781/ijimai.2017.456.

17. Yarotsky D. Error bounds for approximations with deep ReLU networks. *Neural Networks* 2017;94:103–14. URL: https://www.sciencedirect.com/science/article/pii/S0893608017301545{#}b1. doi:10.1016/j.neunet.2017.07.002. arXiv:1610.01145.

18. Braathen A. Deep Learning. *Plant Factory* 2016;:3–5URL: https://mitpress.mit.edu/books/deep-learning. doi:10.1016/B978-0-12-801775-3.00001-9. arXiv:arXiv:1011.1669v3.

19. López-Mart´ınez F. Early neonatal sepsis. https://gallery.azure.ai/Experiment/Early-Neonatal-Sepsis; 2018.

20. Allen-Zhu Z, Orecchia L. Linear Coupling: An Ultimate Unification of Gradient and Mirror Descent. *Distill* 2014;2(4):e6. URL: http://distill.pub/2017/momentumhttp://arxiv.org/abs/1407.1537. doi:10.23915/distill.00006. arXiv:1407.1537.

21. He K, Zhang X, Ren S, Sun J. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. Tech. Rep.; 2015. URL: https://arxiv.org/pdf/1502.01852.pdf. doi:10.1109/ICCV.2015.123. arXiv:1502.01852.

22. Takase T, Oyama S, Kurihara M. Effective neural network training with adaptive learning rate based on training loss. *Neural Networks* 2018;101:68–78. URL: https://www.sciencedirect.com/science/article/pii/S0893608018300303. doi:10.1016/j.neunet.2018.01.016.

23. Jha S, Dey A, Kumar R, Kumar-Solanki V. A Novel Approach on Visual Question Answering by Parameter Prediction using Faster Region Based. *International Journal of Interactive Multimedia and Artificial Intelligence* 2018;URL: http://dx.doi.org/10.9781/ijimai.2018.08.004. doi:10.9781/ijimai.2018.08.004.

24. Esmaeilpour M, Reis A, Mohammadi A. Analyzing the EEG Signals in Order to Estimate the Depth of Anesthesia Using Wavelet and Fuzzy Neural Networks. *International Journal of Interactive Multimedia and Artificial Intelligence* ????;4:2–12. URL: http://www.ijimai.org/journal/sites/default/files/files/2016/11/ijimai20164{_}2{_}2{_}pdf{_}16114.pdf. doi:10.9781/ijimai.2016.422.

25. Percy SF, Piedrahita AM, Escobar RF, Crespo RG. Comparison of neural network topologies for the classification of frogs by their songs. *Soft Computing* 2017;21(23):7099–106. URL: http://link.springer.com/10.1007/s00500-016-2252-5. doi:10.1007/s00500-016-2252-5.

Fernando López-Martínez is currently a doctorate student at the University of Oviedo he has been working in healthcare for more than ten years researching artificial intelligence and healthcare informatics. Holds a Master of Computer Science, and a Bachelor of Computer Science. Worked as faculty professor in computer science, and currently Business Intelligence manager at Sanitas USA in New Jersey.

Edward Rolando Núñez-Valdez is a Software Engineer and has a doctorate in Computer Science. He is an assistant professor in the department of Computer Science at the University of Oviedo. He has participated in several

research projects and published several articles in international journals, conferences, and books. His research interests include recommendation systems, artificial intelligence, and distributed system.

Jaime Lorduy Gomez is working as an associate professor at Rafael Nunez University, with eleven years of experience in surgical instrumentation and academic programs evaluation. He is also a researcher at Surgical Argos group. Bacteriologists with specialization in clinical microbiology and a master's degree in epidemiology.

Vicente García-Díaz is a Software Engineer and has a doctorate in Computer Science. He is an associate professor in the department of Computer Science at the University of Oviedo. He has supervised 80 plus academic projects and published 80 plus research papers in journals, conferences, and books. His research interests include distributed systems, machine learning, and eLearning.