



Expert system supporting an early prediction of the bronchopulmonary dysplasia

Marcin Ochab*, Wiesław Wajs

AGH University of Science and Technology, al. Mickiewicza 30, 30-059 Kraków, Poland

ARTICLE INFO

Article history:

Received 30 October 2014

Accepted 22 August 2015

Keywords:

Machine learning

Feature selection

Chronic lung disease

Prediction

Bronchopulmonary dysplasia

Expert system

Support vector machine

Logit regression

Prematurity

Low-birth-weight infant

ABSTRACT

This work presents a decision support system which uses machine learning to support early prediction of bronchopulmonary dysplasia (BPD) for extremely premature infants after their first week of life. For that purpose a knowledge database was created based on the historical data gathered including data on 109 patients with birth weight less than or equal to 1500 g. The core of the database consists of support vector machine and logit regression classification results calculated specifically for that system, and obtained by considering 2^{14} different combinations of 14 risk factors. Based on the results obtained and user demands, the system recommends the best methods and the most suitable parameter subset among those currently available to the user. The program is also able to estimate the accuracy, sensitivity and specificity together with their standard deviations. The user is also given information on which additional parameter it is worth adding to his measurement system most and what an increase in prediction efficiency it is expected to trigger. The BPD can be predicted by the system with the accuracy reaching up to 83.25% in the best-case scenario, i.e. higher than for most of the models presented in the literature. This work presents a set of examples illustrating the difficulties in obtaining one single model that can be widely used, and thus explaining why an expert system approach is much more useful in day-to-day clinical practice. In addition, the work discusses the significance of the parameters used and the impact of a chosen method on the sensitivity and specificity.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Bronchopulmonary dysplasia (BPD) is a chronic pulmonary disorder affecting premature infants [1], which results in significant morbidity and mortality: almost a third of infants with birth weight lower than 1000 g [2] are affected. This chronic lung disease is most common among children with low birth weight and those who received prolonged mechanical ventilation to treat the respiratory distress syndrome. Due to the fact that the disease is poorly understood, many research projects are focused on identifying its risk factors. It is known that steroids applied before the eighth day of life can prevent BPD development; however, the risks of such treatment may outweigh the benefits [3]. Since illness cannot be diagnosed until the 28th day of life [4], it is very important to predict such a result by the end of the first week, which would enable early intervention and an increased likelihood of preventing the disease [3]. Therefore, an intensive work has been done to define a classifier, based on static parameters

(gathered just after birth) and dynamic ones (collected during the first week of life), which would be able to predict the diagnosis. The literature [5–15] reports several prediction models of BPD used in the research. However, none of them could be used in common clinical practice due to a variety of reasons. For instance, when for some reason there was no technical capability to measure one of the parameters just after birth or the mensuration was delayed, no models based on that variable can be used. Similarly, when using roentgenographic scoring systems – an image of sufficient quality which could be compared with a database, especially by a machine, is extremely hard to obtain, because of ceaseless child movements. Therefore, it is essential to propose an expert system which could advise on which model and method to choose in a given situation.

2. Background

As already mentioned, there are numerous papers devoted to BPD, its risk factors and prediction [16–21]. The most common factors are derived from the analysis of static data among which gestational age and birth weight are mainly considered [9,12,13,15,18,21]. The other factors covered by the literature are administration of surfactant

* Corresponding author at: ul. Jastrzębia 4, 35-207 Rzeszów, Poland. Tel.: +48 504024195 fax: +48 178630252.

E-mail addresses: mj.ochab@labor.rzeszow.pl (M. Ochab), wwa@agh.edu.pl (W. Wajs).

[3,17–19,22], presence of patent ductus arteriosus (PDA) [3,9,14,16–18,20,22–24], or respiratory support [14,21]. Some of the papers introduced sex [3,18,19,21] or even race and ethnicity [21] as factors which seem to be promising. Unfortunately, due to the Polish social structure proposing the second as a parameter in our system would require a very big set of data which we do not have. An analysis of dynamical data can be found in the literature much less frequently, because it requires a constant acquisition of data during the first week of life. Most common parameters acquired that way are arterial blood gas variables like fraction of inspired oxygen (FiO_2) [6,7,9,12,13,20,22] or alveolar–arterial ratio (AA) [25,26] (which is respiratory distress degree measure Eq. (B.1)); blood gas levels like oxygen saturation of arterial hemoglobin (SpO_2) [27] and its standard deviation, mean value, etc. [26], or even time series analysis [28]; heartbeat and its derivatives [26]. One can find several papers regarding BPD prediction with analysis of radiological images [8,10,11]. Unfortunately, during our investigation we found that those to which we had access and which were taken before fourth day of life give ambiguous results. Most likely, the infants' lungs were not developed enough in that stage of life. Therefore, we were not able to use lung images in presented system.

The vast majority of studies make use of logit regression (LOGIT) and the best of such models are able to achieve accuracy in the range 73–82%; some authors use neural networks [26] with accuracy over 80%. A few mention that use of support vector machine (SVM) [29] could give interesting results, but nobody has really investigated that method in the context of BPD prediction. That is why we compared SVM with LOGIT classifiers in our previous papers [30,31]. Generally, SVM models have proved to be more unstable than LOGIT and should be used with a particular care. However, we proved that in certain situations choosing a proper SVM model even from a limited group of randomly constructed ones may lead to better results.

In the lack of a generally accepted model, a multitude of the ones proposed in the literature and considering our previous scientific experience in this respect, we propose to use an expert system. Such a system would assist doctors in deciding which parameters to measure, and which method to use in certain circumstances instead of searching for a single universal method of BPD forecasting. Since we did not find any mention of such a system in the literature we decided to construct one.

3. Materials and methods

3.1. Data collection

Thanks to the Neonatal Intensive Care Unit of The Department of Pediatrics at the Jagiellonian University Medical College, we were able to collect data with our own software. It includes 109 patients born prematurely with birth weight less than or equal to 1500 g, admitted no later than on the second day of life. For 46 of them BPD has been diagnosed after the fourth week of life. Data has been normalized to $[-1,1]$.

In the proposed expert system we consider 14 different features which are used in BPD prediction:

- binary parameters such as
 - presence of patent ductus arteriosus (*pda*),
 - use of a respirator (*respimv*) during the first week of life,
 - administration of surfactant (*surfact*) in the same period;
- real-valued (range in parentheses) such as
 - birth weight (*bweight*) (550–1500 g),
 - gestational age (*gage*) (22–34 weeks),
 - alveolar–arterial ratio (*aa*) (0.05–1) measured during patient admission, which depends on FiO_2 Eq. (B.1),

- the percentage of time during the first week for which the oxygen saturation of hemoglobin was less than 85% (*low85*) (0.03–12.45%) or higher than 94% (*high94*) (14.56–99.02%) [27],
- average number of heartbeats per minute (*bpmmean*) (124.69–161.42 bpm),
- mean and standard deviation of oxygen saturation (*spo2mean*, *spo2dev*) (89.89–98.99% and 1.19–7.98, respectively) and their trends (first day to first week ratio: *bpmmean_tr*, *spo2mean_tr*, *spo2dev_tr*) (0.8–1.18, 0.96–1.07 and 0.51–2.36, respectively).

3.2. Prediction methods

As a prediction methods in our system we used SVM and LOGIT. A brief description of these two algorithms, with the equations and an explanation of the parameters, is provided in Appendix A. All computations were performed in the Matlab R2013a environment. To obtain probability of positive diagnosis, in LOGIT calculations we used functions *glmfit* and *glmval*, whereas in SVM we used LIBSVM library (version 3.17) [32] in C-SVC mode with a sigmoid kernel function, Eq. (1). As mentioned in the previous paper [31], the analysis of several arbitrary tested models revealed that for the specified problem C-SVC method is more effective than nu-SVC (which simply means that the acceptable range of penalty parameter c , Eq. (A.6), is from zero to infinity, rather than between $[0,1]$). It has also been investigated that the sigmoid kernel function gives better results and is much faster in finding the separating hyperplane than the radial basis function, Eq. (2):

$$\text{Sigmoid: } K(X_i, X_j) = \tanh(\gamma X_i^T X_j + r), \quad (1)$$

$$\text{RBF: } K(X_i, X_j) = e^{-\gamma \|X_i - X_j\|^2}, \quad \gamma > 0, \quad (2)$$

where γ , r are kernel parameters.

Accuracy (ACC) defined as below was considered as a preliminary result measure. The sensitivity (TPR) and specificity (SPC) were also obtained the same way:

$$ACC_i = \frac{TP + TN}{TP + TN + FP + FN}, \quad (3)$$

$$TPR_i = \frac{TP}{TP + FN}, \quad (4)$$

$$SPC_i = \frac{TN}{TN + FP}, \quad (5)$$

$$ACC = \frac{1}{n} \sum_{i=1}^n ACC_i, \quad (6)$$

$$ACC_{dev} = \sqrt{\frac{\sum_{i=1}^n (ACC_i - ACC)^2}{n - 1}}, \quad (7)$$

where TP is True Positives, FP is False Positives, FN is False Negatives, TN is True Negatives, i is the Jackknife iteration, $n=30$ is the number of iterations.

3.3. Knowledge database construction procedure

Due to the fact that calculations were made on such a small set of data (109 patients only), a special attention was required in order to find a model that is the most likely to be independent of the specific learning data. Therefore, we decided to use a method similar to Jackknife [33]: for each feature combination and method, calculations were repeated 30 times, each time randomly excluding 10 samples of data and using a cross-validation procedure on the rest of it (each patient was treated as a test sample while the remaining data were a learning set). This way the standard deviation and the mean value of accuracy, sensitivity and specificity were obtained, giving an estimate of the model 'sensitivity' to the data structure and its limited size. Each time we refer in this paper to ACC, TPR or SPC values we mean these averages computed as presented in Eq. (6). Additionally, in the SVM computations the sigmoid parameter γ ($\gamma = 2^{-15}, 2^{-14}, \dots, 2^3$, Eq. (1)) and penalty parameter c ($c = 2^{-5}, 2^{-4}, \dots, 2^{15}$, Eq. (A.6)) were optimized for each model (as extensively discussed in [31], standard procedure of cross-validation and parameters optimization for LIBSVM which is *grid.py* script was not used, because it is not suitable for that problem and gives false results). The coefficient r in Eq. (1) was constant and arbitrarily set to 0 as suggested in SVM library tutorial. Examples of optimized surfaces are presented in Appendix C. In LOGIT, after cross-validation – at the stage of ACC calculation, instead of optimizing parameters –, a threshold value of probability is found, which maximizes the prediction accuracy. The full algorithm for SVM is presented below (Algorithm 1):

Algorithm 1. SVM models evaluation.

```

for each feature_combination do
  for  $\gamma = 2^{-15}, 2^{-14}, \dots, 2^3 \times c = 2^{-5}, 2^{-4}, \dots, 2^{15}$  do
    for 1...30 do
      patients=delete_10_random(all_patients);
      for each test_patient of patients do
        learn_patients=patients – test_patient;
        predict(test_patient, learn_patients);
      end for
      calculate ACC, TPR, SPC;
    end for
  for ACC, TPR, SPC do
    calculate mean_value and dev;
  end for
  end for
  best_results[feature_combination]=best(meanACC);
end for

```

Algorithm 2. Simplified expert system algorithm in pseudo SQL.

```

select MODEL where FEATURES in AVAIL_FEATURES order by
  ACC group by METHOD;
print best result for each method;
foreach METHOD if (DEV > 3.5)
  select MODEL where FEATURES in AVAIL_FEATURES and
    METHOD and DEV < 3.5 order by ACC;
  print alternative model for given method with lower DEV;
end for
foreach not AVAIL_FEATURES as ADD_FEATURE
  select EXT_MODEL where FEATURES in AVAIL_FEATURE-
    S+ADD_FEATURE order by ACC;
end for
select ADD_FEATURE FROM not AVAIL_FEATURES where max
  (ACC) in EXT_MODEL;
print propose additional feature which give max improvement;

```

```

select FEATURE from AVAIL_FEATURES where FEATURE not in
  (MODEL or EXT_MODEL);
print features which can be omitted;

```

Some randomly chosen models, for which high accuracy was reported, were tested again, this time excluding the data of 30 random patients (instead of 10). This way we could confirm that by increasing the number of learning data we also increase the accuracy of prediction and so overfitting does not occur (more comprehensive analysis is presented in [31]).

3.4. Logic structure of the expert system

Having the knowledge database, authors worked out a computer program which is able to propose a prediction method. The method suggestion is based on the results of the previous calculations and on the set of parameters that are available for a given case. The system was developed as a web application using MySQL and PHP. Simplified algorithm of the system in pseudo-SQL is presented in Algorithm 2. At first the system searches for the historical results using simple queries that achieved the best ACC for each method and which have the features selected by the user. If deviation of the model thus found is too high, the user is notified. At the same time, using the same method, the system is looking for another result that meets the deviation requirements. In the next stage the program tries to add to the previous query a single explanatory variable not selected by the user and checks if there occurs an increase in ACC (see Section 3.6.1). If there does, the best one is presented to the user. At last, the application checks if there are any parameters selected by the user that did not appear in the results of all the previous queries. One should consider omitting all such parameters. Sample system output is presented in Example 1.

3.5. Capabilities of the expert system

First, the user selects from among available feature measurements; then, the system looks for highest ACC model which could be used for prognosis. It recommends a parameter set and a method of calculation (SVM or LOGIT). In addition, the application presents the best possible model when using an alternative method. For each recommended forecast model, the expected ACC, TPR, SPC values are presented. What is important, also a standard deviation is presented, which is a measure of how much different (better or worse) results should the user expect in a real prediction of that case. If the deviation is too high (default limit is arbitrarily set to 3.5, which seems to be reasonable in most cases) a warning is shown and another, alternative model is proposed; a custom threshold value can also be defined at the beginning of the search procedure. Further, the user learns which parameters are not worth an effort to measure, because they are not increasing the accuracy of prognosis. Furthermore, the system is able to show which feature should be additionally measured if possible and by how much ACC, TPR and SPC could benefit from that.

3.6. Experiments

3.6.1. Sequential model expansion in the expert system

We performed a series of experiments to analyze and present the system responses in a transparent way.

At first, one parameter only was selected and the system was asked which feature should be chosen to construct the best two-parameter model. Next, in the second trial, the procedure was

Table 1
Results of the experiments performed (Section 3.6.1).

| Starting parameter | Explanatory variable proposed by the expert system in the n -th experiment which maximizes prediction ACC | | | | | | | | | | | | | |
|---------------------------------|---|--|---|--|---|--|--|--|---|--|---|--|---|--|
| | 1st | | 2nd | | 3rd | | 4th | | 5th | | 6th | | 7th | |
| $bweight$ | $\Rightarrow respimv$ 74.05% 76.1% | | | | | | | | | | | | | |
| $gage$ | $\Rightarrow surfact$ 72.85% 80.3% | | $\Rightarrow low85$ 80.8% 81.29% | | $\Rightarrow bpmmean$ 80.8% 81.39% | | $\Rightarrow pda$ 80.8% 82.2% | | $\Rightarrow bweight$ 80.8% 82.67% | | | | | |
| $respimv$ | $\Rightarrow bweight$ 74.05% 76.1% | | | | | | | | | | | | | |
| aa | $\Rightarrow gage$ - 75.55% | | $\Rightarrow surfact$ 74.76% 80.3% | | $\Rightarrow low85$ 80.8% 81.29% | | $\Rightarrow bpmmean$ 80.8% 81.39% | | $\Rightarrow pda$ 80.8% 82.2% | | $\Rightarrow respimv$ 83.04% 82.2% | | | |
| pda | $\Rightarrow gage$ - 78.34% | | $\Rightarrow surfact$ 78.08% 80.37% | | $\Rightarrow low85$ 80.8% 81.29% | | $\Rightarrow bpmmean$ 80.8% 82.2% | | $\Rightarrow bweight$ 80.8% 82.67% | | | | | |
| $surfact$ | $\Rightarrow gage$ 72.85% 80.3% | | $\Rightarrow low85$ 80.8% 81.29% | | $\Rightarrow bpmmean$ 80.8% 81.39% | | $\Rightarrow pda$ 80.8% 82.2% | | $\Rightarrow bweight$ 80.8% 82.67% | | | | | |
| $spo2mean$ | $\Rightarrow bpmmean$ 73.18% 76.8% | | $\Rightarrow gage$ 74.31% 77.77% | | $\Rightarrow surfact$ 80.57% 80.46% | | $\Rightarrow pda$ 80.57% 81.82% | | $\Rightarrow low85$ 81.04% 82.2% | | $\Rightarrow bweight$ 81.04% 82.67% | | | |
| $spo2dev$ | $\Rightarrow gage$ - 76.85% | | $\Rightarrow surfact$ 79.97% 80.3% | | $\Rightarrow low85$ 80.8% 81.29% | | $\Rightarrow bpmmean$ 80.8% 81.39% | | $\Rightarrow pda$ 80.8% 82.2% | | $\Rightarrow bweight$ 80.8% 82.67% | | | |
| $low85$ | $\Rightarrow gage$ - 78.75% | | $\Rightarrow surfact$ 80.8% 81.29% | | $\Rightarrow bpmmean$ 80.8% 81.39% | | $\Rightarrow pda$ 80.8% 82.2% | | $\Rightarrow bweight$ 80.8% 82.67% | | | | | |
| $high94$ | $\Rightarrow gage$ - 76.11% | | $\Rightarrow surfact$ 80.06% 80.3% | | $\Rightarrow low85$ 80.8% 81.29% | | $\Rightarrow bpmmean$ 80.8% 81.39% | | $\Rightarrow pda$ 80.8% 82.2% | | $\Rightarrow bweight$ 80.8% 82.67% | | | |
| $bpmmean$ | $\Rightarrow spo2mean$ 73.18% 76.8% | | $\Rightarrow gage$ 74.31% 77.77% | | $\Rightarrow surfact$ 80.57% 80.46% | | $\Rightarrow pda$ 80.57% 81.82% | | $\Rightarrow low85$ 81.04% 82.2% | | $\Rightarrow bweight$ 81.04% 82.67% | | | |
| $bpmmean_{tr}$ | $\Rightarrow gage$ 72.99% 74.41% | | $\Rightarrow surfact$ 80.45% 80.3% | | $\Rightarrow low85$ 80.8% 81.29% | | $\Rightarrow bpmmean$ 80.8% 81.39% | | $\Rightarrow pda$ 81.68% 82.59% | | $\Rightarrow bweight$ 81.68% 82.67% | | | |
| $spo2dev_{tr}$ | $\Rightarrow \del{high94}$ 72.59% 74.62% | | $\Rightarrow gage$ 72.76% 76.11% | | $\Rightarrow surfact$ 80.5% 80.3% | | $\Rightarrow low85$ 80.8% 81.29% | | $\Rightarrow bpmmean$ 81.17% 81.45% | | $\Rightarrow pda$ 81.69% 82.2% | | $\Rightarrow bweight$ 81.69% 82.67% | |
| $spo2mean_{tr}$ | $\Rightarrow \del{spo2dev}$ 70.11% 74.51% | | $\Rightarrow gage$ 76.01% 76.85% | | $\Rightarrow surfact$ 80.17% 80.3% | | $\Rightarrow low85$ 80.8% 81.29% | | $\Rightarrow bpmmean$ 80.8% 81.39% | | $\Rightarrow pda$ 81.29% 82.2% | | $\Rightarrow bweight$ 81.29% 82.67% | |
| | SVM LOGIT | | SVM LOGIT | | SVM LOGIT | | SVM LOGIT | | SVM LOGIT | | SVM LOGIT | | SVM LOGIT | |
| | ACC | | ACC | | ACC | | ACC | | ACC | | ACC | | ACC | |

Table presents results of all experiments generated by the expert system. Each column contains a result of the n th experiment. The aim of the experiments was to find the best additional feature to the model defined by parameters in previous columns of the selected row. The accuracy of each new model is presented below the winning explanatory variable: on the left side are SVM results, LOGIT on the right (result of the better method is in bold). Parameters which were marked as insignificant are crossed out (incidentally, both the parameters are always marked as insignificant in the third and subsequent experiments). Some rows are quite short because the system looks only for a single parameter which increases the result, but sometimes only adding two or more items at once would be successful.

repeated on the previously proposed model and, based on that, a three-parameter model could be built. Further experiments were continued in the same way, until the system was no longer able to identify an explanatory variable which would improve the prediction accuracy (in practice there was no more than seven attempts).

The above-mentioned series of experiments was repeated 14 times, each time starting from a different explanatory variable. Below we present the full algorithm (Algorithm 3):

Algorithm 3. Experiments: upgrading models.

```

for each variable of all_explanatory_variables do
  model = variable;
  repeat
    new_variable = system  $\rightarrow$  upgrade_model(model);
    model  $\rightarrow$  add(new_variable);
    Print model, method, ACC, ...;
  until new_variable != NULL;
end for

```

Table 2The best 40 models with the highest ACC and $ACC_{dev} < 4$ in the constructed database.

| Params count | BWEIGHT | GAGE | RESPIMV | AA | PDA | SURFACT | SPO2MEAN | SPO2DEV | LOW85 | HIGH94 | BPMMEAN | BPMMEAN_TR | SPO2DEV_TR | SPO2MEAN_TR | ACC dev | ACC (%) | TPR (%) | SPC (%) | Method |
|--------------|---------|------|---------|----|-----|---------|----------|---------|-------|--------|---------|------------|------------|-------------|---------|---------|---------|---------|--------|
| mean value | | | | | | | | | | | | | | | | | | | |
| 9 | | • | | • | • | | • | | • | • | • | • | | • | 3.63 | 83.29 | 79.08 | 86.4 | LIBSVM |
| 8 | | • | | • | • | | | | • | • | • | • | | | 2.16 | 83.25 | 78.82 | 86.49 | LIBSVM |
| 8 | • | • | • | | • | | | • | | • | • | • | | • | 3.88 | 83.23 | 76.54 | 88.07 | LIBSVM |
| 9 | • | • | • | | • | | • | • | | • | • | | • | | 2.98 | 83.11 | 75.07 | 88.92 | LIBSVM |
| 7 | | • | • | | • | • | | | • | | • | • | | | 2.6 | 83.04 | 82.63 | 83.28 | LIBSVM |
| 8 | | • | | • | • | | • | | • | • | • | | • | | 3.39 | 82.96 | 78.12 | 86.5 | LIBSVM |
| 8 | | • | | • | • | | • | | • | | | • | | • | 2.39 | 82.88 | 79.07 | 85.64 | LIBSVM |
| 6 | | • | | • | • | | | • | | | • | • | | | 1.11 | 82.79 | 84.2 | 81.73 | LOGIT |
| 9 | | • | | • | • | | • | • | | • | • | • | | • | 2.3 | 82.78 | 77.84 | 86.45 | LIBSVM |
| 6 | • | • | | | • | • | | | • | | • | • | | | 1.19 | 82.67 | 87.5 | 79.09 | LOGIT |
| 8 | | • | | • | • | | | | • | • | | • | | • | 2.69 | 82.63 | 78.59 | 85.54 | LIBSVM |
| 11 | • | • | | • | • | | • | • | • | • | • | • | | • | 3.12 | 82.6 | 78.65 | 85.55 | LIBSVM |
| 5 | | • | | | • | | | | • | | • | • | | | 1.49 | 82.59 | 84.65 | 81.07 | LOGIT |
| 5 | | • | | | • | | | • | | | • | • | | | 1.32 | 82.56 | 85.25 | 80.54 | LOGIT |
| 6 | | • | | | • | | | | • | | • | • | | | 1.43 | 82.53 | 83.82 | 81.54 | LOGIT |
| 7 | | • | | • | • | | | | • | • | • | | | • | 2.7 | 82.52 | 76.91 | 86.6 | LIBSVM |
| 6 | | • | | | • | | | • | | | • | • | | • | 1.64 | 82.52 | 84.87 | 80.81 | LOGIT |
| 7 | | • | | • | • | | | | • | • | • | • | | | 2.11 | 82.42 | 78.69 | 85.07 | LIBSVM |
| 8 | | • | | • | • | | • | | • | • | • | | | • | 3.44 | 82.41 | 76.2 | 86.96 | LIBSVM |
| 8 | | • | | • | • | | • | | • | • | • | • | | | 1.52 | 82.4 | 78.65 | 85.12 | LIBSVM |
| 6 | | • | | | • | • | | • | | • | • | • | • | • | 3.25 | 82.32 | 80.03 | 83.97 | LIBSVM |
| 10 | | • | | • | • | | • | • | • | • | • | • | | • | 2.44 | 82.31 | 77.86 | 85.56 | LIBSVM |
| 8 | | • | | • | • | | • | | | • | • | • | | • | 2.41 | 82.28 | 77.78 | 85.63 | LIBSVM |
| 7 | | • | | | • | | • | | • | | • | • | | • | 1.41 | 82.27 | 83.35 | 81.47 | LOGIT |
| 7 | | • | | • | • | | | • | • | | • | • | | | 1.73 | 82.24 | 83.18 | 81.51 | LOGIT |
| 8 | | • | | • | • | | | • | | • | • | • | | • | 2.37 | 82.23 | 79.19 | 84.49 | LIBSVM |
| 7 | | • | | • | • | • | | | • | | • | | • | | 3.75 | 82.22 | 78.11 | 85.1 | LIBSVM |
| 8 | | • | | | • | | • | | • | • | • | • | | • | 1.85 | 82.21 | 79.62 | 84.07 | LIBSVM |
| 5 | | • | | | • | • | | | • | | • | | • | | 1.27 | 82.2 | 86.06 | 79.38 | LOGIT |
| 6 | | • | | | • | • | | | • | | • | | | • | 1.76 | 82.2 | 86.31 | 79.25 | LOGIT |
| 7 | | • | | • | • | | • | | • | | • | • | | • | 1.99 | 82.19 | 75.2 | 87.17 | LIBSVM |
| 6 | | • | | | • | | | • | | • | • | • | | | 1.79 | 82.18 | 84.76 | 80.28 | LOGIT |
| 6 | | • | | • | • | | | | • | | • | • | | | 1.69 | 82.17 | 83.66 | 81.08 | LOGIT |
| 7 | | • | | | • | | | • | | • | • | • | | | 1.08 | 82.11 | 82.69 | 81.63 | LOGIT |
| 8 | | • | | • | • | | • | • | | • | • | | • | • | 2.24 | 82.08 | 75.32 | 87.01 | LIBSVM |
| 9 | | • | | • | • | | • | • | • | • | • | • | | • | 2.25 | 82.04 | 77.62 | 85.27 | LIBSVM |
| 8 | | • | • | | • | • | | • | | • | | | | • | 2.93 | 82.04 | 80.13 | 83.4 | LIBSVM |
| 5 | | • | | | • | | | | | • | | | • | | 1.96 | 82.04 | 82.97 | 81.35 | LOGIT |
| 9 | | • | | | • | | • | | • | • | • | • | • | • | 2.55 | 82.01 | 78.33 | 84.71 | LIBSVM |
| 4 | | • | | | • | | | | • | | • | | | | 1.73 | 82.01 | 84.38 | 80.34 | LOGIT |

3.6.2. Features selection using basic Matlab methods

To prove that the expert system enables the user a better result of BPD prognosis than using standard methods, we have performed the following experiment. We used the widely known *sequentialfs* Matlab function, which simply sequentially selects features until no improvement of a prediction is observed. Choices are made based on a user defined classifier:

Example 1. Sample output of the expert system.

You have selected parameters: *bweight*, *aa*, *spo2mean*, *low85*, *bpmmean*, *bpmmean_tr*.

The best result you can achieve is LIBSVM with about 77.47% of accuracy with *dev*=3.44

(*TPR*=65.42%, *SPC*=86.27%).

Suggested parameters for this method are: *bweight*, *spo2mean*, *bpmmean*.

Using alternative method LOGIT you can get up to 76.8% of accuracy with *dev*=1.37

(*TPR*=78.73%, *SPC*=75.33%).

Suggested parameters for this method are: *spo2mean*, *bpmmean*.

You should consider measuring additional parameter *gage* – it can increase the accuracy by 3.28% as follows:

Assuming *dev* < 3.5 best result you can achieve is LIBSVM with about 80.75% of accuracy with *dev*=1.34

(*TPR*=71.59%, *SPC*=87.43%).

Suggested parameters for this method are: *bweight*, *gage*, *low85*, *bpmmean*, *bpmmean_tr*.

Assuming *dev* < 3.5 using alternative method LOGIT you can get up to 80.19% of accuracy with *dev*=1.43

(*TPR*=73.27%, *SPC*=85.07%).

Suggested parameters for this method are: *gage*, *low85*, *bpmmean*, *bpmmean_tr*.

In Your situation parameter *aa* can be omitted – using it doesn't improve accuracy.

- using Logit, we performed 20 repetitions of Monte-Carlo for 10-fold cross-validation without stratification, which provided the best and quite stable result,
- using SVM, we found that even for 50 M-C repetitions we get different results almost each time. A solution to the problem proved to be a backward selection, i.e. inclusion of all parameters and then removing features sequentially, until the accuracy increases. Moreover, using the method mentioned, we were not able to optimize γ and c parameters in the selection stage, so we had to use arbitrary chosen values $c=0$, $g = -1$. We knew from our previous research that these give quite reasonable results in most cases. The parameter optimization was performed only after the feature selection procedure has been completed.

4. Results

While analyzing Table 1, which presents results of experiments mentioned in Section 3.6.1, we are able to extract much information about BPD prediction:

- for models with a limited number of explanatory variables (less than 7) in almost all cases LOGIT gives higher ACC than SVM (and standard deviation of accuracy is about twice lower),
- four parameters (*aa*, *bpmmean_tr*, *spo2dev_tr*, *spo2mean_tr*) never increase the result,
- *high94* and *spo2dev* are used once and later marked as insignificant,

- *spo2mean* is used only once, *respimv* twice, and both seem to have no significance,
- *bweight* is present 11 times, but each time at the end; that means it provides some information which, however, is not essential,
- *pda* seems to be quite important: 11 occurrences at places from 5 to 7,
- one of the most important features is *low85*: 11 occurrences at places from 3 to 6, *bpmmean*: 11 occurrences at places from 2 to 6,
- *surfact* is the second most important item: 10 occurrences at places 3 and 4,
- *gage* is the most important parameter: 11 occurrences always at a place 2 or 3.

Summarizing, LOGIT classifier is able to provide accuracy of up to 82.67% with only 6–7 explanatory variables and in most cases with standard deviation of ACC lower than 2.0. What might be surprising, we can easily obtain more than 81%, using as few as three features.

On the other hand, the analysis of Table 2, which presents 40 best models found in database, shows that better results can be achieved with the SVM classifier. When all features are available, SVM is able to achieve accuracy as high as 83.29% (*TPR* = 79.08%, *SPC* = 86.4%), which is a very promising result. Although in that case standard deviation of ACC was high (*dev*=3.63), the system would propose an alternative SVM model with *ACC* = 83.25% (*TPR* = 78.82%, *SPC* = 86.49%, using *aa*, *pda*, *low85*, *high94*, *bpmmean*, *bpmmean_tr* and *spo2dev_tr*). It is a slightly worse result with a significantly lower deviation, which is only 2.16. In the same situation, using LOGIT, a user can expect ACC of at most 82.79% (*TPR* = 84.2%, *SPC* = 81.73%) with *dev*=1.11. The application would also remark that parameters *bweight*, *respimv* and *surfact* do not increase the accuracy in that case and can be omitted no matter the method (even though *bweight* and *surfact* were quite important in the above analysis; this inconsistency is quite interesting and it is caused by a different approach of classifiers). Thus SVM method is very useful when the doctor has a wide scope of parameters to choose from or, in certain situations, when he has no influence on the choice of measured parameters. In most cases, use of SVM makes sense for more than seven parameters.

The analysis of the presented results reveals that in certain situations the right selection of a method and features is not trivial. Using common parameter selection methods in experiment 3.6.2 with Logit function, we get a model consisting of *gage*, *surfact*, *pda*, *bpmmean* and *high94*. That five-element classifier can provide ACC of 81.57% (*TPR* = 84.03%, *SPC* = 79.78%). It is indeed a good result, but even in experiment 3.6.1 we were able to find a one with *ACC* = 82.2%, not mentioning that the best Logit result for five-parameter model found by the expert system was 82.59% (Table 2). The difference is even larger for SVM, where the features selected are *bweight*, *gage*, *aa*, *surfact*, *spo2mean*, *spo2dev*, *high94*, *bpmmean* and *spo2mean_tr*. In such a configuration, SVM can provide an accuracy of only 78.27% (*TPR* = 68, 56%, *SPC* = 85, 50%), while the presented expert system is able to obtain ACC of up to 83.29% for same number of features. Moreover, as mentioned in Section 3.6.2, only using a backward selection we were able to get any reasonable model candidate. Situation was probably caused by very poor SVM results for small models – adding features one at a time only, when improvements were very low could probably lead the algorithm to a dead end. It can be assumed that using some more complicated (and much more computationally complex) evolutionary algorithm could give better results.

5. Discussion

Generally it could be argued that LOGIT gives very good results, using models with a high number of explanatory variables and SVM is pointless – especially looking at the astounding result of 81.29% accuracy using only three features. However, our experience shows that, in contrast to SVM, the more variables we use, the worse the results we get with LOGIT. Thus, we can assume that if we extend our expert system to include more parameters (radiological images, sex, race and ethnicity, etc.), the difference in accuracy will highly increase to the advantage of SVM. In such a situation, using LOGIT we will have to choose only a small number of the most significant variables, while those which contain some information, but not essential, will be dropped. Contrary to SVM, which is able to profit even from less significant information.

Moreover, in certain situations we can take advantage of the inconsistency of the methods used to assess significance of parameters that we have mentioned. Let us assume that the only knowledge of the case is *bweight*, *aa* and *SpO₂* measurement variables (*spo2mean*, *spo2mean_tr*, *spo2dev*, *spo2dev_tr*). If we have LOGIT only at our disposal, the best result achievable would be generated by a model that uses only *bweight*, *aa* and *spo2mean*: 75.57% of accuracy with *dev*=1.38 (TPR=73.23%, SPC=77.27%). LOGIT is not able to use the information contained in the rest of available parameters to increase the prediction accuracy. However, using our system, we would be advised to use SVM with the following set of features: *bweight*, *aa*, *spo2dev*, *spo2dev_tr*. It might be quite surprising because, as mentioned, SVM generally does not provide good results for small models. Nonetheless, in that situation we would get an accuracy of 78.38% with *dev*=2.49 (TPR=73.82%, SPC=81.57%), almost 3% more than with LOGIT.

Analyzing the constructed database it has been also noticed that in almost all cases LOGIT prediction provides higher sensitivity than specificity – while using SVM exactly the opposite can be observed. Although in common medical practice sensitivity is mostly more important, in the context of BPD it is not so obvious. Applying a treatment could be extremely risky in some cases, and can proceed only having a very high confidence that it is really needed. In such a situation knowledge of a method with the highest specificity is essential [34].

Moreover, thanks to mentioned experiment (Section 3.6.2) authors proved that the usage of simple features selection methods gives unsatisfactory results especially when considering SVM. The solution could be the usage of more sophisticated methods, which however, might be challenging due to high computational complexity and constantly changing circumstances and doctors' demands. For practical reasons it is much easier to use an expert system like the one presented in this paper, which can instantly provide the best possible solution in certain situation and inform what are the reasonable alternatives. No matter the availability of measurements and data, the user can easily adjust the search procedure with desired parameters like expected standard deviation of prediction result, sensitivity or specificity preferences and features being considered. At the same time doctor can be certain that all possible combinations are taken into account and those worth consideration will be presented and compared. Constructing a database for such an expert system is indeed computationally demanding, but needs to be performed only once. Already done, it can be used straight away in common medical practice regardless of the various conditions – further required computations are limited to a few seconds and do not depend on a number of learning cases.

6. Conclusions

6.1. Significance of the parameters

It was confirmed that one of the most important risk factors mentioned in the literature [3,4,12–15,18,20–24,35] is the *gage*. The situation is similar for *pda*, which is reported on [3,4,9,14,16–18,20,22–24,31] as crucial. In that context the use of surfactant (*surfact*) seems to be more important than so far believed [30,31], especially when logit regression is used, which seems logical since it contains some unrevealed knowledge of the doctor who has administered it. Interestingly the *low85* and *bpmmean* parameters, which are quite uncommon in the literature [26,27], turned out to be very important. The last feature which should be mentioned is *bweight*, which is common in the literature and increases prediction accuracy, but it does not seem to be crucial [20,30,31].

6.2. Expert system approach

Authors have found some new models which achieved higher accuracy than the ones referred to in the literature [5–15] (differences are between 1 and 10%), but the real value of research presented herein is the possibility to predict BPD in real-life situations. This paper presents some of the limitations encountered by doctors on a daily basis and because of which it is very hard to use a single model which could be deemed the best one. Therefore, the solution is to use an expert system, which will provide a user with the best method in a given situation and simply help make the right decision. The authors are aware of the fact that the system needs to be extended to include more prediction methods and explanatory variables such as already mentioned: radiological images scores, sex, race, ethnicity, respirator setting, etc. For practical reasons, it would be also very important to use historical data where BPD diagnosis is not a dichotomic variable, but is a grade of severity. Having such data, we could better predict not only the probability of the disease, but also its intensity. Providing a user with the possibility of searching for the best model not only by accuracy, but also by the sensitivity or specificity – depending on requirements – would be also very useful. The expert system presented is only the first step in such a type of doctors' decision support, but we think that this is the direction worth developing.

Conflict of interest statement

None declared.

Appendix A. Overview of the classifiers used

A.1. Logistic regression

Probability of the dependent variable equaling a BPD positive diagnosis ($y_k = 1$), on the condition that the vector of explanatory variables (features of specific case k) equals $X_k = (x_{1,k}, x_{2,k}, \dots, x_{n,k})$, is defined as

$$p_k = P(y_k = 1 | X_k) = \frac{e^{a_0 + \sum_{i=1}^n a_i x_{i,k}}}{1 + e^{a_0 + \sum_{i=1}^n a_i x_{i,k}}}, \quad (\text{A.1})$$

where $x_{i,k}$ is an explanatory variable value for feature i of case k , a_i is a regression coefficients, and n is the number of features.

In contrast to linear regression, where normal distribution of the independent variables is assumed, and because variances of explanatory variables are not equal, the method of least squares cannot be used to calculate regression coefficients. Thus, using learning data they

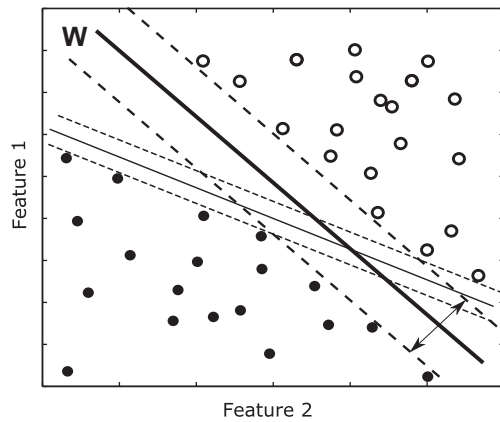


Fig. A1. The idea of maximizing margin between classes with hyperplane W in SVM method. Border points which limit the largest possible margin are called the Support Vectors.

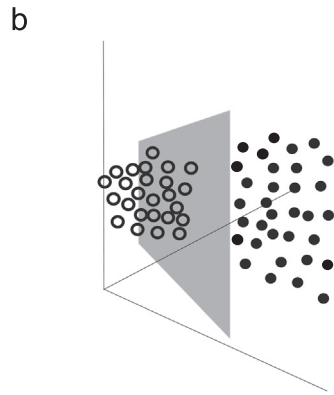
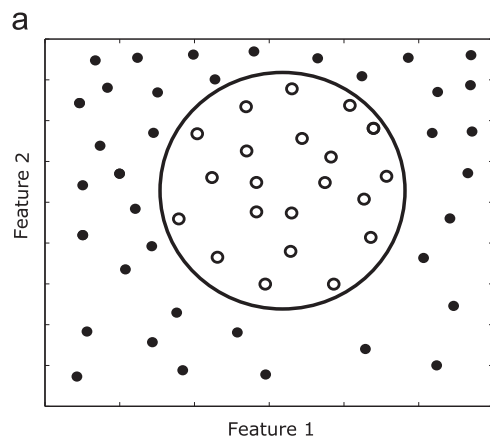


Fig. A2. Making data linearly separable with dimensional transformation with kernel function: (a) data in R^2 space and (b) data transformed to a new R^3 space.

are usually obtained with maximum likelihood estimation by maximizing likelihood function (L), which is the same problem as minimizing its negative logarithm:

$$L = \prod_{y_k=1} p_k \prod_{y_k=0} (1 - p_k), \quad (\text{A.2})$$

$$\ln(L) = \sum_{k=1}^m [y_k \cdot \ln(p_k) + (1 - y_k) \cdot \ln(1 - p_k)], \quad (\text{A.3})$$

where k is an observation number (learning case), y_k is a diagnosis for case k , and m is the number of observations.

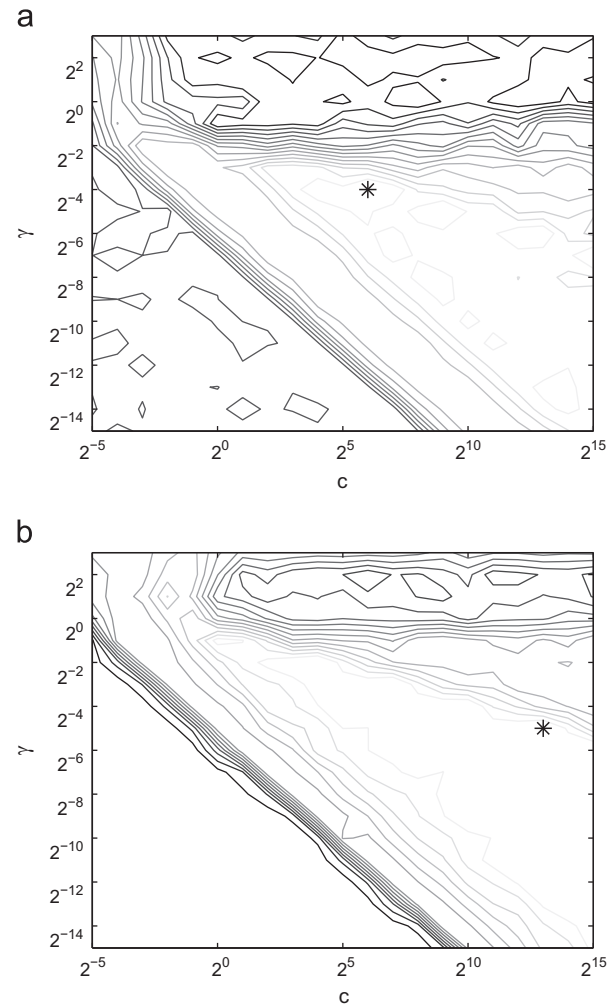


Fig. C1. Optimization landscapes for two example models, where γ is the sigmoid parameter, Eq. (1), and c is the penalty parameter, Eq. (A.6). The contour lines correspond to the mean ACC values, as described in Algorithm 1: (a) eight-parameter model and (b) five-parameter model.

Having regression coefficients a_i , the probability of positive BPD diagnosis case X can be easily predicted using Eq. (A.1).

A.2. Support vector machine

The learning data D divided into two classes y is defined as

$$D = \{(X_k, y_k) | X_k \in R^n, y_k \in \{1, -1\}\}_{k=1}^m. \quad (\text{A.4})$$

The solution which is being searched for is a hyperplane W

$$W \cdot X + b = 0 \quad (\text{A.5})$$

that separates the two classes, and provides maximum margin, as in Fig. A1, which is an equivalent of L -minimizing problem

$$L(W) = \frac{\|W\|^2}{2} + c \cdot \sum_{k=1}^m \varepsilon_k, \quad (\text{A.6})$$

with conditions

$$y_k (W \cdot \phi(X_k) + b) \geq 1 - \varepsilon_k \quad (\text{A.7})$$

where $\varepsilon \geq 0$ is a slack variable, $c > 0$ is a penalty parameter for each incorrectly classified point, and ϕ is a kernel function:

$$K(X_i, X_j) \equiv \phi(X_i)^T \phi(X_j), \quad (\text{A.8})$$

Linearly inseparable problems, thanks to the use of a kernel function, could be transformed into some higher dimensional space, in which there is much higher likelihood that classes would be separable. It should be noted that we use the function K Eq. (A.8) which defines only a dot product of X_i and X_j in some higher dimensional space \mathcal{H} . K can be interpreted as measure of similarity between border points (Support Vectors) and the ones being classified in the space mentioned. Owing to what is known as “kernel trick”, hyperplane W is never explicitly calculated, and it is not even required to know function ϕ which maps data to some (possibly infinite dimensional) space \mathcal{H} [29]. Since in our research we used sigmoid kernel function Eq. (1) for which ϕ exist only in an infinite dimensional space, we were not able to produce any figure presenting that transformation even for two features. Therefore, what Figs. A1 and A2 show is just the ideas of maximizing the margin and data remapping, but they do not contain any real data.

Appendix B. AA and FiO_2 relation

$$AA = \frac{pO_2}{p_{ATM} \cdot FiO_2 - pCO_2}, \quad (B.1)$$

where pO_2 is the oxygen partial pressure, p_{ATM} is the atmospheric pressure, pCO_2 is the carbon dioxide partial pressure, and FiO_2 is a fraction of inspired oxygen.

Appendix C. Examples of $\gamma \times c$ surfaces optimized in SVM computations

See Fig. C1.

References

- [1] B.J. Stoll, N.I. Hansen, E.F. Bell, S. Shankaran, A.R. Laptook, M.C. Walsh, E.C. Hale, N.S. Newman, K. Schibler, W.A. Carlo, et al., Neonatal outcomes of extremely preterm infants from the NICHD Neonatal Research Network, *Pediatrics* 126 (3) (2010) 443–456, <http://dx.doi.org/10.1542/peds.2009-2959>.
- [2] M.C. Walsh, S. Szefer, J. Davis, M. Allen, L. Van Marter, S. Abman, L. Blackmon, A. Jobe, Summary proceedings from the bronchopulmonary dysplasia group, *Pediatrics* 117 (Supplement 1) (2006) S52–S56, <http://dx.doi.org/10.1542/peds.2005-0620I>.
- [3] T. Farstad, D. Bratlid, S. Medbø, T. Markestad, Bronchopulmonary dysplasia-prevalence, severity and predictive factors in a national cohort of extremely premature infants, *Acta Paediatr.* 100 (1) (2011) 53–58, <http://dx.doi.org/10.1111/j.1651-2227.2010.01959.x>.
- [4] J.L. Tapia, D. Agost, A. Alegria, J. Standen, M. Escobar, C. Grandi, G. Musante, J. Zagarra, A. Estay, R. Ramirez, Bronchopulmonary dysplasia: incidence, risk factors and resource utilization in a population of south-American very low birth weight infants, *J. Pediatr.* 82 (1) (2006) 15–20, <http://dx.doi.org/10.1590/S0021-75572006000100005>.
- [5] S. Ryan, J. Nycyk, B. Shaw, Prediction of chronic neonatal lung disease on day 4 of life, *Eur. J. Pediatr.* 155 (8) (1996) 668–671, <http://dx.doi.org/10.1007/BF01957150>.
- [6] N. Subhedar, A. Hamdan, S. Ryan, N. Shaw, Pulmonary artery pressure: early predictor of chronic lung disease in preterm infants, *Arch. Dis. Child.—Fetal Neonatal Ed.* 78 (1) (1998) F20–F24, <http://dx.doi.org/10.1136/fn.78.1.F20>.
- [7] C. Romagnoli, E. Zecca, L. Tortorolo, G. Vento, G. Tortorolo, A scoring system to predict the evolution of respiratory distress syndrome into chronic lung disease in preterm infants, *Intensive Care Med.* 24 (5) (1998) 476–480, <http://dx.doi.org/10.1007/s001340050599>.
- [8] S.S. Toce, P.M. Farrell, L.A. Leavitt, D.P. Samuels, D.K. Edwards, Clinical and roentgenographic scoring systems for assessing bronchopulmonary dysplasia, *Am. J. Dis. Child.* 138 (6) (1984) 581–585, <http://dx.doi.org/10.1001/archpedi.1984.02140440065017>.
- [9] J. Corcoran, C. Patterson, P. Thomas, H. Halliday, Reduction in the risk of bronchopulmonary dysplasia from 1980–1990: results of a multivariate logistic regression analysis, *Eur. J. Pediatr.* 152 (8) (1993) 677–681, <http://dx.doi.org/10.1007/BF01955247>.
- [10] G. Noack, W. Mortenson, B. Robertson, R. Nilsson, Correlations between radiological and cytological findings in early development of bronchopulmonary dysplasia, *Eur. J. Pediatr.* 152 (12) (1993) 1024–1029, <http://dx.doi.org/10.1007/BF01957230>.
- [11] B. Yuksel, A. Greenough, J. Karani, Prediction of chronic lung disease from the chest radiograph appearance at seven days of age, *Acta Paediatr.* 82 (12) (1993) 944–947, <http://dx.doi.org/10.1111/j.1651-2227.1993.tb12605.x>.
- [12] V.K. Bhutani, S. Abbasi, Relative likelihood of bronchopulmonary dysplasia based on pulmonary mechanics measured in preterm neonates during the first week of life, *J. Pediatr.* 120 (4) (1992) 605–613, [http://dx.doi.org/10.1016/S0022-3476\(05\)82491-6](http://dx.doi.org/10.1016/S0022-3476(05)82491-6).
- [13] Y.D. Kim, E.A.-R. Kim, K.-S. Kim, S.-Y. Pi, W. Kang, Scoring method for early prediction of neonatal chronic lung disease using modified respiratory parameters, *J. Korean Med. Sci.* 20 (3) (2005) 397–401, <http://dx.doi.org/10.3346/jkms.2005.20.3.397>.
- [14] C.A. Bhering, C.C. Mochdece, M.E. Moreira, J.R. Rocco, G.M. Sant’Anna, Bronchopulmonary dysplasia prediction model for 7-day-old infants, *J. Pediatr.* 83 (2) (2007) 163–170, <http://dx.doi.org/10.1590/S0021-75572007000200011>.
- [15] R.A. Sinkin, C. Cox, D.L. Phelps, Predicting risk for bronchopulmonary dysplasia: selection criteria for clinical trials, *Pediatrics* 86 (5) (1990) 728–736.
- [16] M.A. Rojas, A. Gonzalez, E. Bancalari, N. Claure, C. Poole, G. Silva-Neto, Changing trends in the epidemiology and pathogenesis of neonatal chronic lung disease, *J. Pediatr.* 4 (126) (1995) 605–610, [http://dx.doi.org/10.1016/S0022-3476\(95\)70362-4](http://dx.doi.org/10.1016/S0022-3476(95)70362-4).
- [17] D.D. Marshall, M. Kotelchuck, T.E. Young, C.L. Bose, L. Kruyer, T.M. O’Shea, Risk factors for chronic lung disease in the surfactant era: a North Carolina population-based study of very low birth weight infants, *Pediatrics* 104 (6) (1999) 1345–1350, <http://dx.doi.org/10.1542/peds.104.6.1345>.
- [18] W. Oh, B. Poindexter, R. Perritt, J. Lemons, C. Bauer, R. Ehrenkranz, B. Stoll, K. Poole, L. Wright, Neonatal research network. association between fluid intake and weight loss during the first ten days of life and risk of bronchopulmonary dysplasia in extremely low birth weight infants, *J. Pediatr.* 147 (6) (2005) 786–790, <http://dx.doi.org/10.1016/j.jpeds.2005.06.039>.
- [19] N. Ambalavanan, K.P. van Meurs, R. Perritt, W.A. Carlo, R.A. Ehrenkranz, D. K. Stevenson, J.A. Lemons, W.K. Poole, R.D. Higgins, Predictors of death or bronchopulmonary dysplasia in preterm infants with respiratory failure, *J. Perinatol.* 28 (6) (2008) 420–426, <http://dx.doi.org/10.1038/jp.2008.18>.
- [20] G.S. Cunha, F. Mezzacappa-Filho, J.D. Ribeiro, Risk factors for bronchopulmonary dysplasia in very low birth weight newborns treated with mechanical ventilation in the first week of life, *J. Trop. Pediatr.* 51 (6) (2005) 334–340, <http://dx.doi.org/10.1093/tropej/fmi051>.
- [21] M.M. Laughon, J.C. Langer, C.L. Bose, P.B. Smith, N. Ambalavanan, K.A. Kennedy, B.J. Stoll, S. Buchter, A.R. Laptook, R.A. Ehrenkranz, et al., Prediction of bronchopulmonary dysplasia by postnatal age in extremely premature infants, *Am. J. Respir. Crit. Care Med.* 183 (12) (2011) 1715–1722, <http://dx.doi.org/10.1164/rccm.201101-0055OC>.
- [22] Y.D. Kim, K.S. Kim, E.A.R. Kim, J.J. Lee, S.J. Park, S.Y. Pi, Perinatal risk factors for the development of bronchopulmonary dysplasia in premature infants less than 32 weeks’ gestation, *J. Korean Soc. Neonatol.* 8 (1) (2001) 78–93.
- [23] J.R. Groothuis, D. Makari, Definition and outpatient management of the very low-birth-weight infant with bronchopulmonary dysplasia, *Adv. Ther.* 29 (4) (2012) 297–311, <http://dx.doi.org/10.1007/s12325-012-0015-y>.
- [24] I. Sosenko, E. Bancalari, New developments in the pathogenesis and prevention of bronchopulmonary dysplasia, In: *The Newborn Lung: Neonatology Questions and Controversies: Expert Consult-Online and Print*, 2012, pp. 217–233.
- [25] R. Gilbert, J. Keighley, The arterial-alveolar oxygen tension ratio, an index of gas exchange applicable to varying inspired oxygen concentrations, *Am. Rev. Respir. Dis.* 109 (1) (1974) 142.
- [26] P. Stoch, Prediction of bronchopulmonary dysplasia in preterm neonates using statistical and artificial neural network tools (Ph.D. dissertation) (in Polish), AGH University of Science and Technology, Kraków, 2007, pp. 60–72.
- [27] A.H. Jobe, The new bronchopulmonary dysplasia, *Curr. Opin. Pediatr.* 23 (2) (2011) 167, <http://dx.doi.org/10.1097/MOP.0b013e3283423e6b>.
- [28] L. Kuenzel, Predicting and understanding bronchopulmonary dysplasia in premature infants, *Stanf. Undergrad. Res. J.* 10 (2011) 36–43.
- [29] C.J. Burges, A tutorial on support vector machines for pattern recognition, *Data Min. Knowl. Discov.* 2 (2) (1998) 121–167, <http://dx.doi.org/10.1023/A:1009715923555>.
- [30] M. Ochab, W. Wajs, Bronchopulmonary dysplasia prediction using support vector machine and logit regression, *Inf. Technol. Biomed.* 4 (2014) 365–374, http://dx.doi.org/10.1007/978-3-319-06596-0_34.
- [31] M. Ochab, W. Wajs, Bronchopulmonary dysplasia prediction using support vector machine and libsvm, In: *Proceedings of the 2014 Federated Conference on Computer Science and Information Systems, Annals of Computer Science and Information Systems*, vol. 2, IEEE, Warsaw, 2014, pp. 201–208, <http://dx.doi.org/10.15439/2014F111>.
- [32] C.-C. Chang, C.-J. Lin, Libsvm: a library for support vector machines, *ACM Trans. Intell. Syst. Technol.* 2 (2011) 27:1–27:27, software available at (<http://www.csie.ntu.edu.tw/~cjlin/libsvm>). <http://dx.doi.org/10.1145/1961189.1961199>.
- [33] H.L. Jones, Jackknife estimation of functions of stratum means, *Biometrika* 61 (2) (1974) 343–348, <http://dx.doi.org/10.1093/biomet/61.2.343>.
- [34] A.K. Akobeng, Understanding diagnostic tests 1: sensitivity, specificity and predictive values, *Acta Paediatr.* 96 (3) (2007) 338–341, <http://dx.doi.org/10.1111/j.1651-2227.2006.00180.x>.
- [35] Z. Ali, P. Schmidt, J. Dodd, D.L. Jeppesen, Bronchopulmonary dysplasia: a review, *Arch. Gynecol. Obstet.* 288 (2) (2013) 325–333, <http://dx.doi.org/10.1007/s00404-013-2753-8>.