Journal Pre-proof


How to best predict short bowel syndrome outcome with machine learning approaches?

Michaela Servi , Elisa Mussi , Riccardo Coletta , Antonino Morabito , Adrian Bianchi , Rocco Furferi , Yary Volpe

PII: S2666-9900(21)00015-X
DOI: https://doi.org/10.1016/j.cmpbup.2021.100016
Reference: CMPBUP 100016

To appear in: *Computer Methods and Programs in Biomedicine Update*

Received date: 29 March 2021
Revised date: 26 April 2021
Accepted date: 4 June 2021

**Highlights**

- Predict enteral autonomy in patients affected by short bowel syndrome
- Dataset with demographic data composed of 86 subjects
- Spot-checking of machine learning approaches
- Evaluation of accuracy, precision, recall and F1-score
- Best performing method decision tree with accuracy 0.85

**How to best predict short bowel syndrome outcome with machine learning approaches?**

**Michaela Servi [1], Elisa Mussi [1], Riccardo Coletta [2], Antonino Morabito [2], Adrian Bianchi [3], Rocco Furferi [1] and Yary Volpe [1]**

[1] Department of Industrial Engineering, University of Florence, via di Santa Marta, 3, Florence, 50139, Italy

[2] Department of Pediatric Surgery, Meyer Children's Hospital, University of Florence, Viale Gaetano Pieraccini, 24, Florence, FI 50139, Italy

[3] Royal Manchester Children's Hospital, Manchester University NHS Foundation Trust, Manchester, United Kingdom

Corresponding author:

Elisa Mussi

University of Florence

Department of Industrial Engineering

Via di Santa Marta 3, 50139, Florence, Italy

E-mail: elisa.mussi@unifi.it



Manuscript word count: 4952

Abstract word count: 173

Number of figures: 2

Number of tables: 3

**Abstract:** In recent years, there has been an extensive use of machine learning techniques in the medical field for diagnostic or therapeutic prediction purposes. In the field of short bowel syndrome, numerous statistical studies have been proposed, but to date no machine learning techniques have been exploited to predict the outcomes of the surgery commonly performed in paediatric patients suffering from this pathology. One reason for this lack can be identified in the fact that this is a rare condition and therefore it is difficult to have a large dataset. This paper investigates the possibility of processing demographic data of paediatric short bowel syndrome patients by spot-checking machine learning algorithms on a dataset of 86 patients. The experimental setup was developed to ensure the best performance of each algorithm and to take into account the moderate unbalance of the output classes. The Decision Tree algorithms proved to be the best solution in terms of accuracy, precision, recall and F1-score (obtaining values of 0.85 for each metric considered), capable of better understanding the data model.

**Keywords:** Short Bowel Syndrome, Machine Learning, Enteral Autonomy, Decision Tree, Paediatric



## 1. Introduction

Paediatric short bowel syndrome (SBS) is a ravaging clinical condition causing malabsorption and intestinal dysmotility and has an incidence of 25 per 100,000 live births, resulting from rare congenital malformations or acquired diseases leading to extensive surgical resections of the small intestine (e.g. necrotising enterocolitis, intestinal volvulus, gastroschisis and intestinal atresias) [1]. To live an acceptable life, patients with SBS need intravenous or parenteral (PN) support that provides adequate fluids, electrolytes and/or nutrients to ensure proper digestive function, prevent permanent organ damage, avoid malnutrition and/or dehydration-related diseases, and maintain life. PN can have adverse effects on liver function and various complications and it is therefore ideal to wean the patient off PN as soon as possible [2]. In order to achieve enteral autonomy, the course of short bowel patients may involve various surgical strategies such as autologous gastrointestinal reconstructive surgery (AGIR) [3,4] and, in more advanced stages of the disease, intestinal transplantation. Surgical intervention is not always a guarantee of survival and/or enteral autonomy, therefore many studies have focused on identifying factors that can determine the degree of success of the intervention in terms of survival and ability to wean off PN [5–14]. Specifically, through statistical analyses such as the chi-square test, factors like preoperative bowel length, the presence of the

Ileocecal valve, the conjugated bilirubin and enteral nutrition were investigated in order to identify possible correlations with the surgical outcome.

On the basis of these studies, the idea is to create a tool that, starting from the patient's demographic data, can predict weaning off PN after AGIR surgery.

This predictive data can be used by clinicians with the aim to plan in advance medical treatment to reduce the need for PN. Teduglutide, glucagon-like peptide 2 (GLP-2) analogue has recently been approved for use in paediatric SBS patients older than 1 year of age as a novel agent to augment intestinal adaptation [15]. The naturally occurring gut hormone, GLP-2, is a pleiotropic intestinotrophic hormone that enhances digestive and absorptive capacity. Knowing in advance which patient can be the best candidate could open new prospective in SBS treatment [16].

Specifically, in the present study the authors investigate the possibility of using machine learning approaches which have hitherto not been explored as a predictive approach applied to this medical context.

The study was conducted on data collected in two centres specialised in the field of SBS: Royal Manchester Children's Hospital (UK) and Meyer Children's Hospital (Italy). Since SBS is a rare disease, the dataset involved a limited number of patients compared to other diseases commonly analysed with this kind of instruments. Moreover, considering weaning off PN as a condition to be predicted, the data are unbalanced, i.e. a high percentage of patients out of the total is enterally autonomous after the surgery.

Nevertheless, as already highlighted in the literature, the problem of predicting enteral autonomy is a hot topic in the scientific community, therefore the authors considered it significant to investigate and pave the way for new predictive approaches. The novelty of the study and the peculiarity of the dataset required an exhaustive analysis to define the most suitable approach to address this problem.

Specifically, in order to determine which type or class of algorithms is appropriate for identifying the structure of the problem, a spot-checking experiment was carried out testing approaches, parameters and metrics to best describe the SBS problem.

The algorithms and metrics investigated are described in section 2. Section 3 presents the results obtained in the different tests. Finally, these results are analysed in section 4 and conclusions are drafted in Section 5.

## 2. Materials and Methods

### *2.1 Dataset*

A total of 86 patients with SBS were selected between 2002 and 2021 at the Royal Manchester Children's Hospital (UK) and Meyer Children's Hospital (IT). The dataset includes 47 females and 39 males who all underwent intestinal lengthening surgery (their characteristics are collected in Table 1). The demographic information in the dataset are: sex, weight on admission, gestational age, aetiology (short bowel, gastroschisis, malrotation, necrotizing enterocolitis (NEC), other), patient's age at the time of surgery, bowel dilatation, presence or absence of ileocecal valve (ICV), bowel length before surgery, bowel length after surgery, resulting percent elongation, whether bowel transplantation took place and weaning off PN after surgery.

The decision to use these parameters as predictive values is in line with previous studies [5] in which an attempt was made to retrospectively analyse values that could be decisive in understanding how severe a case of SBS can be considered. In [5], in addition to the parameters mentioned above it appears conjugated bilirubin, which was not included in this study as it was not available for a considerable number of patients.

**Table 1.** Summary of demographic data.

| | | |
|---|---|---|
| **Sex** | | |
| Male | 47 | 54.7% |
| Female | 39 | 45.3% |
| **Weight (gr)** | 2477±713.5 | |
| **Gestational Age** | 35.3±3.4 | |
| **Etiology** | | |
| Short bowel | 17 | 19.8% |
| Gastroschisis | 38 | 44.2% |
| Malrotation | 16 | 18.6% |
| NEC | 11 | 12.8% |
| Other | 4 | 4.7% |
| **Age at surgery (months)** | 26.5±39 | |
| **Dilatation (cm)** | 5.8±1.3 | |
| **ICV** | | |
| Absent | 53 | 61.6% |
| Present | 33 | 38.4% |
| **Length pre surgery (cm)** | 44.2±32.8 | |
| **Length post surgery (cm)** | 71.4±52.9 | |
| **Length increase (%)** | 72%±25 | |
| **Transplant** | | |
| No | 82 | 95.3% |
| Yes | 4 | 4.7% |
| ***Off PN*** | | |

| | | |
|---|---|---|
| No | 28 | 32.6% |
| Yes | 58 | 67.4% |

*2.2 Dataset analysis*

SBS in infants is an uncommon but highly morbid condition. The rarity of the syndrome hampers data collection and, therefore, the creation of a large database to allow in-depth analysis of the relationships between demographic characteristics, surgical outcome and follow-up. In recent years, the advancement of surgical techniques has significantly increased the survival rate (95% [6]) and life expectancy of paediatric patients with SBS. With these premises, the scientific community approaches this problem by focusing on the probability of weaning off PN, i.e. obtaining enteral autonomy, after the lengthening surgery, rather than the prediction of survival. Accordingly, in this work the weaning off PN (*off PN Yes* or *No*) represents the variable to be predicted. The dataset results to be slightly unbalanced on the variable to be predicted (as can be observed in Table 1) and, at first sight, numerically limited (86 patients); nonetheless, the dataset's uniqueness opens up new ways for research in the field of the rare SBS pathology.

*2.3 Experimental setup*

To date, to the best of the authors' knowledge, outcome prediction in the context of SBS has never been addressed with machine learning approaches. The idea of the present work is therefore to investigate different machine learning approaches with a spot-checking process to determine the best strategy, taking into account in the definition of the algorithms and their parameters the peculiarity of the database. Specifically, the experimental setup was created using a strategy that guarantees a high level of diversity among the tested algorithms, as well as the ability for each algorithm to achieve its best performance by manipulating the input parameters and feature coding. The following paragraphs describe the tested algorithms and the metrics used to analyse the results obtained from each of them.

Neural Network

A neural network [17] is a set of algorithms that attempts to recognise the underlying relationships in a set of data through a process that mimics the way the human brain operates.  A neural network is an interconnected system of perceptrons. The perceptron is a type of binary classifier that maps its inputs into an output value calculated with an activation function ($\chi$) that evaluates the scalar product between the input vector ($x$) and a vector of weights ($w$) added to a constant bias value ($b$). The activation formula is therefore:

$$f(x) = \chi(\langle w, x \rangle + b) \quad (1)$$

By modifying the vector of the weights $w$ through a specific learning algorithm [18], it is possible to modulate the output of a perceptron, with the aim of obtaining learning properties. Several layers of perceptrons form a neural network.

When designing a neural network for each new application area, it is important to define parameters such as the number of layers and the number of neurons per layer. The network implemented in this study is a fully connected network for which the number of hidden layers was varied from 1 to 2, considering the limited size of the number of features and examples. For each hidden layer the number of neurons was varied between 5 and 15; considering that the number of input features is 11, it was chosen to test both a lower number of neurons and a higher number of neurons. Moreover, a test was performed by assigning different weights to the two output classes, which is considered a good approach in case of unbalanced dataset [19].

For this experiment the preprocessing of the input features followed two strategies: i) in the first strategy the categorical feature (Etiology) was encoded with One-hot encoding, the dichotomous features (Sex, ICV and Transplant) were mapped to -1 and 1 and numerical features were normalised in the range [-1,1]; ii) in the second strategy, all of the features were converted to number and normalized in the range [-1,1].

Support Vector Machine

The second class of tested algorithms are Support Vector Machine (SVM) [20] classifiers. SVMs map input vectors into a high-dimensional nonlinear feature space and construct a hyperplane that linearly separates the data into different classes. The hyperplane equation for classifying can be written as:

$$H: w^T\,(x) + b = 0 \quad (2)$$

This hyperplane is called maximum margin hyperplane because it maximises the distance between the nearest points of the two classes.

In some cases, to achieve a more accurate separation of the data, the dot product can be replaced by a non-linear kernel function. In this work the linear and polynomial (with degree equal to 3) feature division hyperplanes were tested. Again, to address the problem of unbalanced dataset, a test was carried out by assigning each class a weight. To test this algorithm all the features were made numerical and normalized in the range [-1,1].

Naive Bayes

Naive Bayes methods [21] are a set of supervised learning algorithms based on the application of Bayes' theorem with the "naive" assumption of conditional independence between each pair of features given the value of the class variable. The algorithm was tested assuming a Gaussian distribution; the probability density of $v$ given a class $C_k$ , $p(x = v\ |C_k)$ can be computed as:

$$p(x = v\ |C_k) = \frac{1}{\sqrt{2\pi\sigma_k^2}} e^{-\frac{(v-\mu_k)^2}{2\sigma_k^2}} \tag{3}$$

where $\sigma_k^2$ and $\mu_k$ are the mean and variance of the assumed distribution. To test this algorithm, all the features were converted to number and normalized in the range [-1,1].

K-Nearest Neighbor

The idea behind nearest neighbour methods [22] is to identify a predefined number of training samples that are closest in distance to the new point and use them to predict the label. In k-nearest neighbour learning (K-NN), the number of samples can be a user-defined constant $k$, according or it can vary to the local density of the points, radius-based neighbour learning. The distance can, in general, be any metric measure, standard Euclidean distance is the most common choice:

$$d(x,y) = \sqrt{\sum_{1}^{k}(x_i - y_i)^2} \tag{4}$$

Despite its simplicity, nearest neighbors has been successful in a large number of classification problems; being a non-parametric method, it is often successful in classification situations where the decision boundary is very irregular.

The optimal choice of the value is highly data-dependent, in general a larger $k$ suppresses the effects of noise, but makes the classification boundaries less distinct. Values of $k$ equal to 4 and 5 were tested, considering the limited size of the number of features and examples.

To test this algorithm, all the features were converted to number and normalized in the range [-1,1].

Decision Tree

The objective of decision trees [23] is to create a model that predicts the value of a target variable by learning simple decision rules deduced from the characteristics of the data. In this work, three different decision tree algorithms were tested: ID3 [24], C4.5 [25] and CART [26].

ID3 produces a multi-way tree in which the attribute that provides the maximum information gain for the goals is identified for each node. The trees grow to their highest height, and then a pruning process is normally added to boost the ability of the tree to generalize invisible details. Input data must be categorical.

C4.5 is the successor to ID3 and has eliminated the constraint that features must be categorical by dynamically specifying a discrete attribute that divides the continuous value of the attribute into a discrete series of intervals. C4.5 transforms the trained trees into if-then rule sets. The consistency of each rule is then tested to determine the order in which it should be performed. The pruning phase is performed by deleting a rule precondition if the consistency of the rule increases without it.

CART (Classification and Regression Trees) is similar to C4.5, but differs in that it accepts numeric target variables and does not calculate rule sets. CART builds binary trees using a feature and threshold that generates the most information on each node.

Moreover, the Random Forest variant [27], an aggregation by bagging of decision trees, was tested. To test the decision tree algorithms the features were not modified.

In order to deal with the problem of unbalanced dataset all the algorithms were tested after performing an oversampling step. In detail, the Synthetic Minority Oversampling Technique (SMOTE) algorithm was used to balance the number of examples belonging to the two classes [28]. In brief SMOTE creates synthetic examples for the minority class by selecting examples that are close in feature space, drawing a line between them and drawing a new sample at a point along that line.

2.4 Evaluation metrics

In order to analyse the results obtained from the tests, different metrics were taken into consideration which, if integrated, allow an exhaustive analysis of the algorithm performances. The definition of the metrics (shown in Table 2) uses the acronym:

- TP to indicate the number of True Positive,
- TN to indicate the number of True Negative,
- FP to indicate the number of False Positive
- FN to indicate the number of False Negative.

The first metric is *accuracy* which represents the ability of an algorithm to correctly predict the class of an example. *Precision* indicates the proportion of positive identifications that are actually correct. *Recall* indicates what proportion of TPs are correctly identified. The *F1-score* can be interpreted as a weighted average of *precision* and *recall*, where an *F1-score* reaches its best value at 1 and its worst at 0**.**

**Table 2.** Definition of performance assessment metrics.

| **Metric** | **Equation** |
|---|---|
| Accuracy | $(TP + TN)/(TP + TN + FP + FN)$ |
| Precision | $TP/(TP + FP)$ |
| Recall | $TP/(TP + FN)$ |
| F1-score | $(2 * precision * recall)/(precision + recall)$ |

Finally, for an immediate reading of the performance of the classifiers, the confusion matrices were computed. The predicted values are represented in each column of the matrix, while the actual values are represented in each row and the number of correctly classified examples is indicated on the diagonal.

**3. Results**

The experimental setup is schematised in Figure 1. All algorithms were implemented in Python programming language using dedicated libraries. All the computations were carried out on a computer running the Windows 10 64-bit operating system, with Intel Core TM i7-7700 CPU at 3.60 GHz and 16.0 GB of installed random access memory.

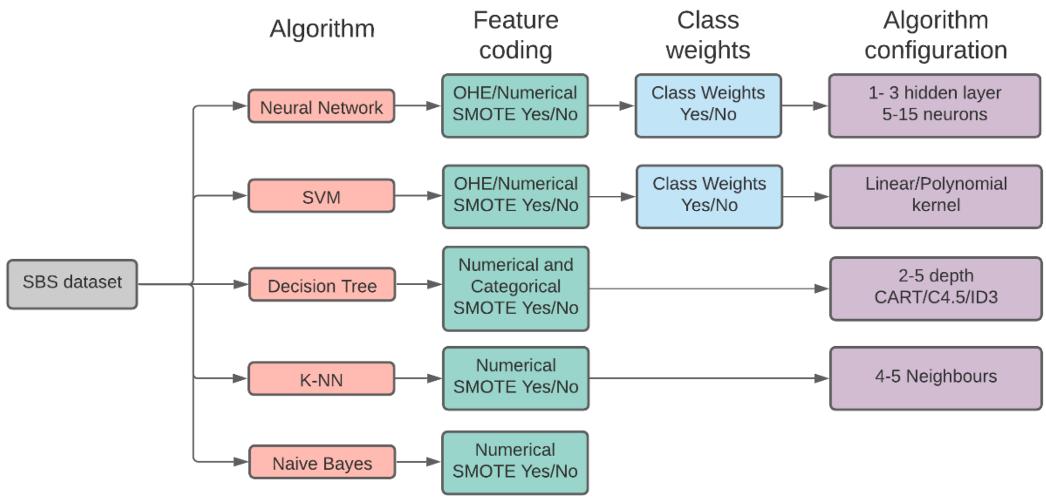


**Figure 1.** Draft of experimental setup.

A total of 43 tests were performed, including 16 tests for neural networks, 12 tests for SVM, 12 tests for decision tree, 2 tests for K-NN and one test for Naive Bayes. The dataset was split in 80% for the training set and 20% for the test set and the two output classes to predict are *off PN Yes* (1) or *No* (0). For reasons of space the results corresponding to the best configuration for each algorithm was chosen to be reported in Table 3. Specifically the values of *accuracy*, *precision*, *recall* and *F1-score* are indicated in this Table and Figure 2 shows the obtained confusion matrices.

**Table 3.** Summary of demographic data.

| **Algorithm** | **Configuration** | **Metrics** |
|---|---|---|
| Neural Network | Numerical features, class weights, 1 layer, 5 neurons | Accuracy: 0.8 Precision: 0.8 Recall: 0.8 F1-score: 0.79 |
| SVM | Numerical features, Polynomial kernel | Accuracy: 0.75 Precision: 0.82 Recall: 0.75 F1-score: 0.71 |
| Decision Tree | Numerical and categorical features, depth 5, C4.5 | Accuracy: 0.85 Precision: 0.85 Recall: 0.85 F1-score: 0.85 |
| K-NN | Numerical features, 5 neighbours, SMOTE | Accuracy: 0.77 Precision: 0.78 Recall: 0.78 F1-score: 0.78 |
| Naïve Bayes | Numerical features, SMOTE | Accuracy: 0.7 Precision: 0.7 Recall: 0.7 F1-score: 0.68 |

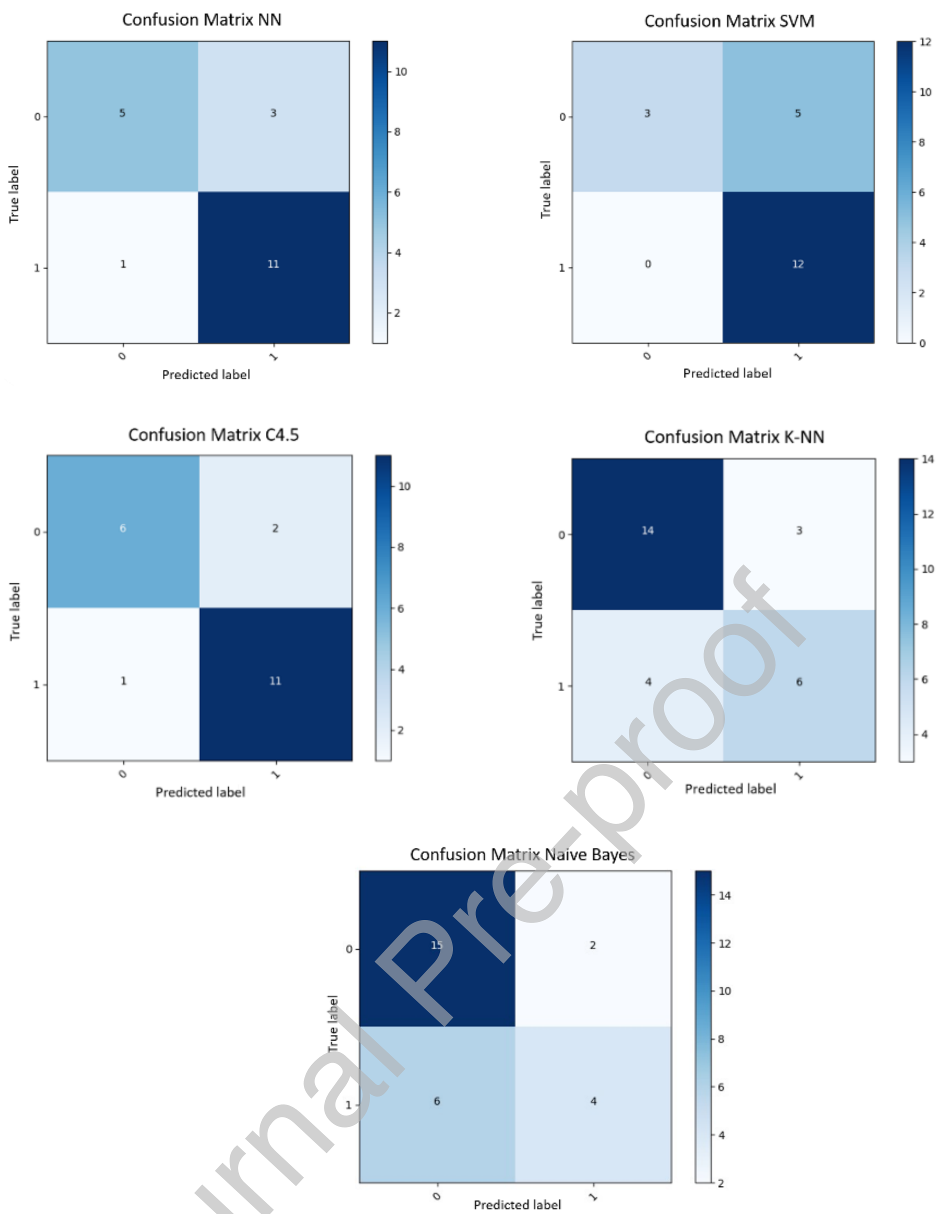


**Figure 2.** Best configuration confusion matrices.



## 4. Discussion

In this work, an experimental setup was developed to comprehensively analyse the dataset of SBS patients. The aim was to predict the possibility of achieving enteral autonomy following bowel lengthening surgery.

Analysing the results obtained with the best configurations for each tested algorithm, shown in Table 3, it can be observed that the algorithms achieve different weighted average value of *precision*, *recall* and *F1-score* and of overall prediction *accuracy*. From the confusion matrices (Figure 2) it is possible to derive the values of the three metrics for the single classes (0 and 1) which, as explained in paragraph 2.4, depend on TP, TN, FP, FN.

In detail, it can be observed that the neural network obtains high *precision* values both for the prediction of class 0 (*prec_0* = 0.83) and for class 1 (*prec_1* = 0.79). With reference to *recall*, this is lower for class 0 (*rec_0* = 0.62) and higher for class 1 (*rec_1* = 0.92). Consequently, the *F1-score* values are higher for class 1 (*F1_0* = 0.71, *F1_1* = 0.85).

With reference to the performance of the SVM algorithm, a higher *precision* is observed for class 0 than for class 1 (*prec_0* = 1, *prec_1* = 0.71). The *recall* values have an opposite trend (*rec_0* = 0.38, *rec_1* = 1). Consequently *F1-score* is higher for class 1 (*F1_0* = 0.55, *F1_1* = 0.83).

The C4.5 algorithm provides the best results for decision trees; in detail good *precision* values are obtained for both classes (*prec_0* = 0.86, *prec_1* = 0.85) and *recall* is higher in the case of class 1 (*rec_0* = 0.75, *rec_1* = 0.92). The *F1-score* values are respectively *F1_0* = 0.8, *F1_1* = 0.88.

The K-NN algorithm achieves higher performance when the SMOTE algorithm is applied to the data. Specifically, the values for the individual classes are *prec_0* = 0.82, *prec_1* = 0.7, *rec_0* = 0.82, *rec_1* = 0.7. Consequently, similar values are obtained for *F1-score* (*F1_0* = 0.82, *F1_1* = 0.7).

The Naïve Bayes algorithm also performs better when the SMOTE algorithm is applied to the data. In this case the *precision* values of class 1 are lower than class 0 (*prec_0* = 0.71, *prec_1* = 0.67) as well as the *recall* values (*rec_0* = 0.88, *rec_1* = 0.4). The resulting *F1-score* values are *F1_0* = 0.79, *F1_1* = 0.5.

Having defined the performances of all the algorithms it is important to establish the relevance of the metrics in relation to the SBS problem and therefore in relation to the prediction purpose. In a conservative approach, where it is important to monitor more carefully patients who do not reach enteral autonomy, it is crucial to identify all or as many patients as possible belonging to class 0 (corresponding to high *rec_0*). At the same time it is important to provide the physician with a reliable tool for positive prediction (class 1) to ensure that the weaning off PN prediction is not incorrect. In fact, an incorrect assignment of a patient to class 0 is clinically less risky (and therefore more conservative) than an incorrect assignment to class 1, because an *off PN Yes* prediction means predicting that surgery will be successful. This means that the classifier should be precise in predicting class 1 (corresponding to high *prec_1*). In summary the aim is not to risk a class 0 patient being assigned to class 1 or having the class 1 classification unreliable.

On the basis of these considerations, the most suitable algorithm for the application is the Decision Tree with the C4.5 algorithm and, to a lesser extent but still acceptable, the K-NN algorithm, with *accuracy* of 0.85 and 0.77 respectively. Among the other algorithms the neural network could represent a valid solution if the available dataset had a greater frequency of the class 0, this would allow a better learning and consequently a potential increase of *rec_0* at the moment not satisfying.

Previous studies of this clinical scenario have focused on statistical analyses to determine whether a single factor can be significant in predicting a patient's exit from parenteral nutrition [6,7,9]. Recently, Belza et al proposed a study [5] which analyses the relationships between the most commonly used factors with the aim of defining a severity prediction tool. This tool is based on a severity score which, by weighing certain parameters, is able to classify patients into three classes connected to the patient's probability of being weaned from parenteral nutrition.

Unlike the approaches previously investigated, in this work an instrument is proposed which, by independently evaluating all the collected parameters, defines the underlying model of the data. According to this model, for each new patient, the tool is able to predict weaning from parenteral nutrition, with a certain accuracy. The reliability of the proposed tool, which to date reaches a maximum of 85%, has the potential to improve performance by disposing of a larger database, that would allow the description of a best-fitting model.

## 5. Conclusions

SBS is a rare pathological syndrome that leads to intestinal malabsorption resulting in malnutrition and dehydration. The condition requires parenteral support, to provide adequate nutrients, which can be weaned after intestinal lengthening surgery. In this work, an attempt was made to predict the phenomenon of weaning off PN by analyzing the demographic characteristics of the patient using machine learning techniques. Since this medical context has never been studied using machine learning techniques before, and since the dataset has a moderate imbalance on the two output levels (*off PN Yes* and *No*), it was important to explore various methods and algorithms.

The test of the algorithms and their configurations highlighted the possibility of using the Decision Tree algorithm, with learning algorithm C4.5, to predict the output of PN with good performance. In particular, values of *accuracy* of 0.85, *precision* of 0.85, *recall* of 0.85 and *F1-score* of 0.85 were achieved.

In the clinical field, the developed tool has the potential to be used for multiple purposes, in fact, a natural use could be the study and discrimination of the target population based on the outcome of the intervention. In addition, after strengthening the robustness of the prediction, the tool can be used both to optimise the intervention, being able to rely on the descriptive capacity of the various predictive factors, and to reliably predict the patient's chances of weaning from PN.

The authors intend to perform a multicentre study in future works in order to build a larger database to be able to describe more accurately the characteristics of patients for whom the intervention is not resolving in terms of weaning off PN and in order to identify the demographic characteristics with greater impact on the outcome.



*Disclosures*

The authors have no relevant financial interests in this article and no potential conflicts of interest to disclose.

**References**

[1] P.W. Wales, N. De Silva, J. Kim, L. Lecce, T. To, A. Moore, Neonatal Short Bowel Syndrome: Population-Based Estimates of Incidence and Mortality Rates, in: J. Pediatr. Surg., W.B. Saunders, 2004: pp. 690–695. https://doi.org/10.1016/j.jpedsurg.2004.01.036.

[2] S.J. Wood, B. Khalil, F. Fusaro, S.E. Folaranmi, S.A. Sparks, A. Morabito, Early structured surgical management plan for neonates with short bowel syndrome may improve outcomes, World J. Surg. 37 (2013) 1714–1717. https://doi.org/10.1007/s00268-013-2011-z.

[3] R. Coletta, E. Mussi, F. Uccheddu, Y. Volpe, A. Morabito, Preoperative Planning of Spiral Intestinal Lengthening and Tailoring: A Geometrical Approach, Bioengineering. 8 (2021) 20. https://doi.org/10.3390/bioengineering8020020.

[4] E. Negri, R. Coletta, A. Morabito, Congenital short bowel syndrome: systematic review of a rare condition, J. Pediatr. Surg. 55 (2020) 1809–1814. https://doi.org/10.1016/j.jpedsurg.2020.03.009.

[5] C. Belza, K. Fitzgerald, N. de Silva, Y. Avitzur, P.W. Wales, Early Predictors of Enteral Autonomy in Pediatric Intestinal Failure Resulting From Short Bowel Syndrome: Development of a Disease Severity Scoring Tool, J. Parenter. Enter. Nutr. 43 (2019) 961–969. https://doi.org/10.1002/jpen.1691.

[6] F. Pederiva, A. Sgrò, R. Coletta, B. Khalil, A. Morabito, Outcomes in patients with short bowel syndrome after autologous intestinal reconstruction: Does etiology matter?, J. Pediatr. Surg. 53 (2018) 1345–1350. https://doi.org/10.1016/j.jpedsurg.2018.03.009.

[7] F.R. Demehri, L. Stephens, E. Herrman, B. West, A. Mehringer, M.A. Arnold, P.I. Brown, D.H. Teitelbaum, Enteral autonomy in pediatric short bowel syndrome: Predictive factors one year after diagnosis, J. Pediatr. Surg. 50 (2015) 131–135. https://doi.org/10.1016/j.jpedsurg.2014.10.011.

[8] C. Belza, K. Fitzgerald, N. De Silva, Y. Avitzur, K. Steinberg, G. Courtney-Martin, P.W. Wales, Predicting Intestinal Adaptation in Pediatric Intestinal Failure: A Retrospective Cohort Study, Ann. Surg. 269 (2019) 988–993. https://doi.org/10.1097/SLA.0000000000002602.

[9] G.C. Ives, F.R. Demehri, R. Sanchez, M. Barrett, S. Gadepalli, D.H. Teitelbaum, Small Bowel Diameter in Short Bowel Syndrome as a Predictive Factor for Achieving Enteral Autonomy, J. Pediatr. 178 (2016) 275-277.e1. https://doi.org/10.1016/j.jpeds.2016.08.007.

[10] M. Hukkinen, R. Kivisaari, L. Merras-Salmio, A. Koivusalo, M.P. Pakarinen, Small Bowel Dilatation Predicts Prolonged Parenteral Nutrition and Decreased Survival in Pediatric Short Bowel Syndrome, Ann. Surg. 266 (2017) 369–375. https://doi.org/10.1097/SLA.0000000000001893.

[11] D.W. Wilmore, Factors correlating with a successful outcome following extensive intestinal resection in newborn infants, J. Pediatr. 80 (1972) 88–95. https://doi.org/10.1016/S0022-3476(72)80459-1.

[12] R.E. Quirós-Tejeira, M.E. Ament, L. Reyen, F. Herzog, M. Merjanian, N. Olivares-Serrano, J.H.

Vargas, Long-term parenteral nutritional support and intestinal adaptation in children with short bowel syndrome: A 25-year experience, J. Pediatr. 145 (2004) 157–163. https://doi.org/10.1016/j.jpeds.2004.02.030.

[13] A.U. Spencer, A. Neaga, B. West, J. Safran, P. Brown, I. Btaiche, B. Kuzma-O'Reilly, D.H. Teitelbaum, K.D. Anderson, D.H. Teitelbaum, R.J. Touloukian, M.Z. Schwartz, P.R. Schloerb, J.C. Emond, J.A. Haller, Pediatric short bowel syndrome: Redefining predictors of success, in: Ann. Surg., Ann Surg, 2005: pp. 403–412. https://doi.org/10.1097/01.sla.0000179647.24046.03.

[14] M.M. Ziegler, Short bowel syndrome: Remedial features that outcome and the duration of parenteral nutrition, J. Pediatr. 131 (1997) 335–336. https://doi.org/10.1016/S0022-3476(97)80050-9.

[15] S. Hill, B.A. Carter, V. Cohran, S. Horslen, S.S. Kaufman, S.A. Kocoshis, D.F. Mercer, R.J. Merritt, M.P. Pakarinen, S. Protheroe, J.F. Thompson, C.P.B. Vanderpool, R.S. Venick, P.W. Wales, S.E. Smith, M. Yoon, A.A. Grimm, Safety Findings in Pediatric Patients During Long- Term Treatment With Teduglutide for Short- Bowel Syndrome–Associated Intestinal Failure: Pooled Analysis of 4 Clinical Studies, J. Parenter. Enter. Nutr. (2021) jpen.2061. https://doi.org/10.1002/jpen.2061.

[16] C.T. Tee, K. Wallis, S.M. Gabe, Emerging treatment options for short bowel syndrome: Potential role of teduglutide, Clin. Exp. Gastroenterol. 4 (2011) 189–196. https://doi.org/10.2147/CEG.S13906.

[17] R. Hecht-Nielsen, Theory of the Backpropagation Neural Network, Proceedings of the International Joint Conference on Neural Networks 1, 593–611, June 1989. © 1989 IEEE., in: Neural Networks Percept., Elsevier, 1992: pp. 65–93. https://doi.org/10.1016/b978-0-12-741252-8.50010-8.

[18] R. Rojas, R. Rojas, The Backpropagation Algorithm, in: Neural Networks, Springer Berlin Heidelberg, 1996: pp. 149–182. https://doi.org/10.1007/978-3-642-61068-4_7.

[19] M. Zhu, J. Xia, X. Jin, M. Yan, G. Cai, J. Yan, G. Ning, Class weights random forest algorithm for processing class imbalanced medical data, IEEE Access. 6 (2018) 4641–4652. https://doi.org/10.1109/ACCESS.2018.2789428.

[20] W.S. Noble, What is a support vector machine?, Nat. Biotechnol. 24 (2006) 1565–1567. https://doi.org/10.1038/nbt1206-1565.

[21] M.M. Saritas, A. Yasar, International Journal of Intelligent Systems and Applications in Engineering Performance Analysis of ANN and Naive Bayes Classification Algorithm for Data Classification, Orig. Res. Pap. Int. J. Intell. Syst. Appl. Eng. IJISAE. 7 (2019) 88–91. https://doi.org/10.1039/b000000x.

[22] V. García, R.A. Mollineda, J.S. Sánchez, On the k-NN performance in a challenging scenario of imbalance and overlapping, Pattern Anal. Appl. 11 (2008) 269–280. https://doi.org/10.1007/s10044-007-0087-5.

[23] D. Lavanya, K.U. Rani, Performance Evaluation of Decision Tree Classifiers on Medical Datasets, Int. J. Comput. Appl. 26 (2011) 1–4. https://doi.org/10.5120/3095-4247.

[24] J.R. Quinlan, Induction of Decision Trees, Mach. Learn. 1 (1986) 81–106. https://doi.org/10.1023/A:1022643204877.

[25] J.R. Quinlan, Learning Logical Definitions from Relations, Mach. Learn. 5 (1990) 239–266. https://doi.org/10.1023/A:1022699322624.

[26] W. Loh, Classification and regression trees, WIREs Data Min. Knowl. Discov. 1 (2011) 14–23. https://doi.org/10.1002/widm.8.

[27] L. Breiman, Random forests, Mach. Learn. 45 (2001) 5–32. https://doi.org/10.1023/A:1010933404324.

[28] N. V. Chawla, K.W. Bowyer, L.O. Hall, W.P. Kegelmeyer, SMOTE: Synthetic Minority Over-sampling Technique, J. Artif. Intell. Res. 16 (2011) 321–357. https://doi.org/10.1613/jair.953.