

Machine learning techniques for decision support in anesthesia

Olivier Caelen¹, Gianluca Bontempi¹, and Luc Barvais²

¹ Machine Learning Group, Département d'Informatique,
Université Libre de Bruxelles, Bruxelles, Belgium

² Service d'Anesthésiologie-Réanimation, Faculté de Médecine,
Université Libre de Bruxelles, Bruxelles, Belgium

Abstract. The growing availability of measurement devices in the operating room enables the collection of a huge amount of data about the state of the patient and the doctors' practice during a surgical operation. This paper explores the possibilities of generating, from these data, decision support rules in order to support the daily anesthesia procedures. In particular, we focus on machine learning techniques to design a decision support tool. The preliminary tests in a simulation setting are promising and show the role of computational intelligence techniques in extracting useful information for anesthesiologists.

1 Introduction

Machine learning and data mining are key technologies in order to transform data into useful information for better diagnosis, event detection and decision aid. This paper deals with the anesthesia domain where several platforms have recently been made available to support the anesthesiologist in the operating room. An example is the TOOLBOX software [1] which has been used for several years by the group of anesthesiology of the ULB Erasme Hospital¹. This software

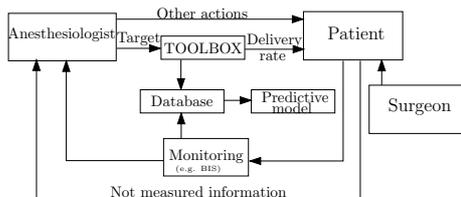


Fig. 1. The TOOLBOX software and the anesthesia procedure.

monitors the patient's state and acts as a servo-controller on the multiple intravenous drug infusions, whose setting is regularly adjusted by the anesthesiologist,

¹ L'Hôpital Erasme is the university hospital of Université Libre de Bruxelles (ULB), Brussels, Belgium.

by simultaneously using pharmacokinetic and pharmacodynamic principles [2] (Figure 1). Before and during the operation, TOOLBOX stores necessary statistics and monitoring information like: (i) basic details regarding the doctor, the patient and his general state, (ii) the type of surgery, (iii) the evolution of the hemodynamic and physiological parameters (e.g. the BIS) of the patient, (iv) the evolution of the drugs concentration levels chosen by the anesthesiologist. In this study, 910 surgical intervention sessions are used to build the database.

This paper discusses and assesses the role of machine learning techniques in extracting useful information from the database generated by TOOLBOX in order to test and develop a decision support tool to assist the anesthetist during his routine procedure. In particular, we will focus on the impact of the brain concentration of Propofol on the hypnosis upon monitoring with the bispectral index. The bispectral index (BIS) [3, 4] is a well-known measure adopted by anesthesiologists to rate the depth of the hypnosis. The BIS index represents the electro-encephalographic signal in a normalized range from 100 to 0, where 100 stands for the “awake” status and 0 stands for electrical silence. Propofol is a short-acting intravenous hypnotic agent used for the induction and maintenance of general anesthesia. According to some information about the patient and the target Propofol brain concentration, the decision support tool will give information on the future BIS value of the patient that will help the anesthesiologist to take the best decision regarding the drugs modification. A correlation technique is used to estimate the time between the Propofol target brain concentration modification and its impact on the BIS index.

In this paper, we will assess and compare a linear model and a local learning approach called the lazy learning [5]. The learning procedure is supported by a forward feature selection procedure to reduce the input dimensionality of the prediction problem. This step is very important since there is a large number of variables (e.g. the patient age, the surgery type, the phase of the operation, etc.) which could influence the value of the BIS signal.

The main contributions of this paper are (i) the application of a system identification procedure on a huge database concerning the relation between the drug modification and the impact on the BIS index, (ii) the comparison of the accuracy between a classical linear model and a local linear model (lazy learning) (iii) the execution of a forward variable selection to extract the most relevant input variables.

2 Learning the predictive model

The goal of our decision support architecture is to assist the anesthetist in adjusting the concentration of the Propofol drug in order to let the BIS of the patient attain the desired level. Suppose that the dynamics of the BIS index can be described by a single-input single-output (SISO) NARMAX (Nonlinear AutoRegressive Moving Average with eXternal input) discrete-time dynamic system [6]

$$B(t + \Delta t^*) = f(B(t), tpo(t), tpn(t), \Delta timeP, tr(t), a, w, h, s, lbm) + \epsilon(t) \quad (1)$$

where, at time t , $B(t)$ is the BIS value³, $tpo(t)$ (in $\mu g/ml$) is the old concentration of Propofol, $tpn(t)$ is the new concentration of Propofol (action of the anesthetist), $\Delta timeP$ is the time between t and the previous Propofol modification and $tr(t) \in [4, 6]$ (in ng/ml) is the concentration of Remifentanyl. Also, a , w , h , s , lbm are, respectively, the age, the weight, the height, the sex and the lean body mass of the patient. $\epsilon(t)$ is random noise and Δt^* is the time delay which maximizes the correlation between drug modification and BIS variation.

We apply a system identification procedure [6] to the samples collected by TOOLBOX to estimate the model

$$\widehat{B}(t + \Delta t^*) = \widehat{f}(B(t), tpo(t), tpn(t), \Delta timeP, tr(t), a, w, h, s, lbm, \alpha_N) \quad (2)$$

where α_N is a vector containing the parameters of the model.

Let us define as query point q the vector containing all the input variables. We identify the system by using a *training set* of $N = 1702$ measures $\{B_i(t + \Delta t^*), q_i(t)\}$, $i = 1, \dots, N$. The sample $(B_i(t + \Delta t^*), q_i(t))$ means that (i) we observed at time $t + \Delta t^*$ the BIS value $B_i(t + \Delta t^*)$, (ii) the target concentration of Propofol was set at time t and (iii) no other modification of the Propofol target occurred during the interval $[t, t + \Delta t^*]$.

The simplest learning approach boils down to a conventional linear identification [7]. However, when linear identification does not return a sufficiently accurate prediction, the designer may want to use alternative methods for learning non-linear relationships. This paper adopts a method of local modeling, called lazy learning, which proved to be successful in many problems of non-linear modeling [8] and in two international competitions on data analysis and time series prediction [9].

The learning procedure is preceded by a feature selection step in order to reduce the dimensionality of the problem. We use a *sequential forward selection* [10] where a leave-one-out cross-validation procedure is used to assess the accuracy of the input sets. This procedure is useful both for statistical reasons and to return to the anesthesiologist high-level information about which variables play a role on the evolution of the patient physiological parameters.

3 Results

This section summarizes the results of the different BIS predictors assessed during the forward selection procedure. Three leave-one-out criteria are used to assess the accuracy of the predictive models. Let $\widehat{E}_i^{loo} = \widehat{B}_{(-i)}(t + \Delta t^*) - B_i(t + \Delta t^*)$ be the leave-one-out error made on the sample i where $\widehat{B}_{(-i)}(t + \Delta t^*)$ is the prediction for the sample i returned by a model trained on all the samples except i . The first criterion is the *normalized mean squared error NMSE* = $\frac{\sum_{i=1}^N (\widehat{E}_i^{loo})^2}{\sum_{i=1}^N (\widehat{\mu}_b - B_i(t + \Delta t^*))^2}$ where $\widehat{\mu}_b = 1/N \sum_{i=1}^N B_i(t + \Delta t^*)$ is the average of the future

³ In order to smooth the fluctuations, $B(t)$ is the time average of the BIS over the interval $[t - 30, t]$ and $B(t + \Delta t)$ is the time average over the interval $(t + \Delta t - 30, t + \Delta t + 30)$.

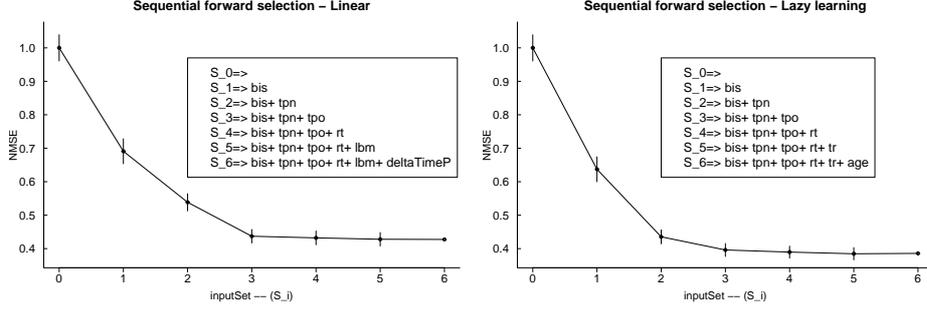


Fig. 2. Results of the sequential forward selection process for a linear model (left) and a lazy learning predictor (right).

BIS index. This quantity is greater than zero and normalizes the performance of the predictor with respect to the variance of the signal to be predicted. A value of NMSE= 1 means that we are simply predicting the average of the BIS series. The second criterion is the *mean of the absolute errors* $MAE = \frac{1}{N} \sum_{i=1}^N \left| \hat{E}_i^{loo} \right|$ which returns an indication of the average magnitude of the errors made by the decision support system. The last criterion returns the percentage (P) of times that the variation of the BIS is wrongly predicted.

$$P = 100 \left(1 - \frac{\sum_{i=1}^N I \left[\left(\hat{B}_{(-i)}(t + \Delta t^*) - B_i(t) \right) \cdot (B_i(t + \Delta t^*) - B_i(t)) \right]}{N} \right) \quad (3)$$

$$\text{where } I[A] = \begin{cases} 1 & \text{if } A \geq 0, \\ 0 & \text{if } A < 0. \end{cases}$$

In both cases (linear and lazy) the forward selection procedure confirms the importance of taking into account the value of the current BIS index as well as the new and the old targets of Propofol. Note that most of the other variables (age, sex, ...) are integrated in the pharmacokinetic model used by TOOLBOX and this could explain the fact that these variables are discarded by the selection procedure.

The last experiment compares a linear predictor $A_{S_3}^{lin}$ and a lazy predictor $A_{S_3}^{lazy}$ both taking as inputs the current BIS value, the previous and the current Propofol target. Table 1 reports the three criteria accuracy figures for the linear model $A_{S_3}^{lin}$ and the lazy predictor $A_{S_3}^{lazy}$. According to a paired-t test all the differences are significant. This means that the lazy predictor significantly outperforms the linear one and suggest the existence of a nonlinear relationship linking the target of Propofol and the BIS signal.

model	NMSE	MAE	P
$A_{S_3}^{lin}$	0.437	6.35	18.2
$A_{S_3}^{lazy}$	0.395	5.91	16.9

Table 1. Three measures of BIS prediction error for the linear ($A_{S_3}^{lin}$) and the lazy ($A_{S_3}^{lazy}$) model.

4 Conclusion and future work

This paper compares conventional linear and machine learning prediction techniques in a predictive modeling task. The encouraging results show that predictive models can extract useful information from historical data and provide support to the decisions of anesthetists during surgical operations. Future work will focus on the implementation of a prototype to be tested, in real conditions, during daily operations in the operating room.

Acknowledgement The work was funded by the FIRST Europe project number EP1A320501R059F/415717 of the Region Wallonne, Belgium. The authors thank their colleagues A. A. Miranda and G. Bejjani for their valuable comments.

References

1. F. Cantraine and E. Coussaert. The first object oriented monitor for intravenous anesthesia. *Journal of Clinical Monitoring and Computing*, 16(1):3–10, 2000.
2. J. M. Bailey, W. M. Haddad, and T. Hayakawa. Closed-loop control in clinical pharmacology: Paradigms, benefits and challenges. In *Proceedings of the 2004 American Control Conference*, pages 2268–2277, 2004.
3. J.C. Sigl and N.G. Chamoun. An introduction to bispectral analysis for the electroencephalogram. *Clin Monitor*, 10:392–404, 1994.
4. A. Gentilini, C. W. Frei, A. H. Glattfedler, M. Morari, T. J. Sieber, R. Wymann, T. W. Schnider, and A. M. Zbinden. Multitasked closed-loop control in anesthesia. *IEEE Engineering in Medicine and Biology*, pages 39–53, 2001.
5. G. Bontempi, M. Birattari, and H. Bersini. Lazy learning for modeling and control design. *International Journal of Control*, 72(7/8):643–658, 1999.
6. T. Söderström and P. Stoica. *System Identification*. Prentice Hall, 1989.
7. R. H. Myers. *Classical and Modern Regression with Applications*. PWS-KENT, Boston, MA, 1990.
8. G. Bontempi. *Local Learning Techniques for Modeling, Prediction and Control*. PhD thesis, IRIDIA- Université Libre de Bruxelles, 1999.
9. G. Bontempi, M. Birattari, and H. Bersini. Lazy learners at work: the lazy learning toolbox. In *Proceeding of the 7th European Congress on Intelligent Techniques and Soft Computing EUFIT '99*, 1999.
10. D. W. Aha and R. L. Bankert. A comparative evaluation of sequential feature selection algorithms. *Artificial Intelligence and Statistics*, 5, 1996.