Predicting evolution of critically ill patients

Jan Ramon
Dept. of Computer Science
K.U.Leuven
Celestijnenlaan 200A 3001 Leuven, Belgium
Jan.Ramon@cs.kuleuven.be

ABSTRACT

In this paper we describe an interesting application of temporal data mining, predicting the evolution of critically ill patients. We point out several issues which make this application particularly challenging. We outline our work in progress and discuss directions for further work.

Keywords

Intensive care, Temporal data

1. INTRODUCTION

Critically ill patients are patients in need of vital monitoring and/or support. People can need intensive care after major trauma (unplanned emergencies, e.g. severe burns, acute organ failure) or after major surgery (often planned, e.g. cardiac surgery).

Crucial in intensive care is to provide the correct treatment and to detect clinical problems early enough so preventive or curative measures can be taken in time. In practice, an intensivist analyses all the patient related data in order to detect a change in the condition of his patients. However, the data may contain much more information than one can currently extract from it. Human intellectual capacity is known to be not capable of handling more than 5 to 7 different parameters at the same time [5]. In the early eighties already, it was estimated that an average ICU patient is described by approximately 250 different parameters.

Therefore, recently there is a growing interest in the intensive care domain for using data mining methods. In particular the advent of automatic monitoring systems, producing a wider range of measurements at a much higher resolution, has emphasized the need for automatic tools for better analysing the data.

In this paper, we will describe the domain of intensive care from the point of view of temporal data mining. We will point out the major challenges we see and discuss a number of preliminary results. The remainder of this paper is structured as follows: Section 2 describes the problem domain. In Section 3 we briefly review some related work. In Section 4 we discuss the major temporal data mining challenges in this domain. In Section 5 we summarize our preliminary results. We conclude in Section 6 and outline our future work.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
TPTDM'06 August 20, 2006, Philadelphia, Pennsylvania, USA
Copyright 2006 ACM 1-59593-442-1 ...$5.00.

2. INTENSIVE CARE DATA

In a typical intensive care unit, a wide range of data is available. A major part of the data consists of information collected during the stay in the intensive care unit. As usually this data is recorded regularly, it has a temporal character. This data mainly consists of measurements of the state of the patient (measurements of patient parameters, laboratory data, bacteriological data, subjective observations, ...) and information about the treatments (medication administered, treatments, feeding, treatment policy, ...).

Online monitoring systems usually produce measurements for more than 20 parameters (temperature, heart rate, ...) every one to three minutes. In this way, a large amount of temporal data is generated. Still, apart from the temporal data, several other data sources may be available, and they too provide an important part of the knowledge needed for decision making. Such background information can include demographic data, historical data (earlier illnesses of the patient), medication information (detailing to use and the effects of medications, including ontologies on medications) and a wide variety of expert knowledge,

3. RELATED WORK

In the field of critical care, a wide range of approaches has been studied in the hope to be able to better predict the evolution of patients.

A first strategy, the construction of hand-crafted models, is quite popular in intensive care research [2]. These models are usually intended to describe a small part of the functioning of the human body. E.g. [12] reviews specific models for the inflammatory response. They get complex quite quickly, e.g. models using 20 differential equations to describe one phenomenon are not exceptional. For a more global or longer term picture, the detailed knowledge to build a mechanistic model is often lacking. Therefore, there is a growing interest in statistical approaches.

Although to the best of our knowledge no systematic study has been conducted on a large intensive care dataset, in recent years several efforts have shown the interestingness of applying data mining methods. One of them is the work [6, 7] which combines expert knowledge with statistical methods on a dataset of 148 patients which contains records of

118 attributes registered every minute (for a total of 679817 records). The result is a system that recommends therapeutic interventions. A number of approaches specialise on learning about a particular body function (e.g. [3] analyses the respiratory pressure-volume curves of mechanically ventilated patients), a particular problem ([8] analyses hospital infections and [10, 11] analyse threshold selection for models predicting length of stay in ICU) or considers a subpopulation of the intensive care patients (e.g. [1] builds a predictive model for the survival of patients that undergo damage control surgery).

4. CHALLENGES

In this section we discuss a number of factors that make (temporal) data mining in the field of intensive care particularly challenging.

- A first important issue is the *amount of data*. An intensive care unit with 250 beds measuring a moderate number of parameters of its patients every 3 minutes generates several tens of gigabytes of temporal (numeric) data a year.
- Even with accurate measuring systems, *noise* is unavoidable. It is well-known that e.g. changing the position of the patient can influence some measurements. A procedure which is sometimes used is to validate data, i.e. to regularly check the equipment and insert a tag in the data indicating whether the equipment was functioning correctly.
- Searching temporal patterns is made more difficult by the different time scales used by several processes: some processes take a few minutes while other processes may take several days. Therefore, when algorithms are looking for certain trends, looking for these trends at the correct time scale is important. There exists some literature on the subject of dealing with varying time scales, but it does not seem trivial to translate this to the intensive care domain.
- A final challenge is to take into account the (relational) background knowledge in the models built. E.g. it is well-known that patient specific characteristics (e.g. age, sex, average heart rate in normal conditions, ...) are important factors for prediction. Also, taking into account ontologies on treatments and medications turns out to be necessary in order to be able to use these classes of information.

5. PRELIMINARY RESULTS

In this section we summarize a number of preliminary results we obtained in this domain.

5.1 Daily data

In a first preliminary study we used a database with 1503 patients for which only daily measurements were stored. We tried to predict, given the measurements of the patient up to now, several state variables of the patients for one up to four days in the future. State variables include survival, length of stay and endangering states such as kidney failure, inflammation and shock. As an illustration, Table 1 shows a selection of the prediction tasks and results. [9] describes in more detail this data and experiments.

Task	DT	FORF	NB	TAN
Survival	79%	82%	88%	86%
Stay longer than 3 days	75%	79%	83%	83%
Inflam. in 2 days	63%	66%	69%	70%
Severe inflam. in 3 days	79%	83%	81%	82%
Inflam.-shock in 3 days	85%	90%	89%	92%
Kidney failure in 4 days	83%	88%	87%	87%

Table 1: Some prediction tasks and results (area under the ROC-curve). Except for the first two, all tasks try to predict that a patient will get into an endangering state assuming that he is not in this state today.

In general, results are promising (above 80% area under the ROC curve for the majority of the prediction tasks) and for some target variables comparable to predictions of doctors and nurses.

We used relational decision trees (DT), first-order random forests (FORF), naive Bayes (NB) and tree augmented naive bayes (TAN) algorithms for constructing models. In general, the random forests perform better than single decision trees. Naive Bayes and tree augmented naive Bayes perform better than decision tree based methods but have the disadvantage that they are more difficult to interprete by doctors.

5.2 Using Gaussian processes to predict short-range evolutions

In a second stage, we are considering a database with measurements every minute. We are using Gaussian processes to predict the evolution of patient parameters in the near future. Early experiments shows promising results, but things get more difficult when predicting further ahead. The main problem here is that complex processes cause interactions between parameters which are difficult to model at a high resolution time scale. Only more global predictions are possible. We are interested in applying techniques that allow us to work on different time scales, but it seems not trivial to combine this with a number of other strategies (e.g. for noise-handling and including background information) that are essential for our application.

As patient-specific background information (such as age, sex, average parameter values in healthy conditions, ...) include important predictive information accordig to domain experts we are recently investigating ways to enrich the input of the Gaussian processes with such patient-specific data [4]. We have observed that these global parameters have little effect on the predictive performance of the short-term evolution of parameters, and therefore expect that they will play a more prominent role when predicting further ahead in the future.

6. CONCLUSIONS

In the intensive care domain, large amounts of (mainly) temporal data are available. Finding interesting patterns can potentially yield large gains in terms of lifes and treatment costs. However, despite some early exploration yielding promising results on small-scale databases, several challenges remain. Many of the properties commonly listed as the currently major problems in temporal data mining such as database size and noise are present. In addition, the application domain and the processes one should model are

very complex and naive methods seem to hit a barrier when trying to perform more difficult tasks. We are especially interested in investigating ways to combine common solutions for the size and noise problems with ways to handle multi-relational data and ways to handle temporal data on different time scales.

Acknowledgements

Jan Ramon is a postdoctoral fellow of the Fund for Scientific Research (FWO) of Flanders. This work is also supported by the Interdisciplinary research fund of the K.U.Leuven (IDO 3M030483). The author wants to thank Fabian Guiza, Daan Fierens (computer science) and Geert Meyfroidt and Greet Van Den Berghe (Intensive care) for the interesting discussions.

7. REFERENCES

[1] N. Aoki, M. Wall, J. Demsar, B. Zupan, T. Granchi, M. Schreiber, J. Holcomb, M. Byrne, K. Liscum, G. Goodwin, J. Beck, and K. Mattox. Predictive model for survival at the conclusion of a damage control laparotomy. *American Journal of Surgery*, 180(6):540–544, Dec 2000.

[2] T. G. Buchman. Nonlinear dynamics, complex systems, and the pathology of critical illness. *Current Opinion in Critical Care*, 10:378–382, 2004.

[3] S. Ganzert, J. Guttmann, K. Kersting, R. Kuhlen, C. Putensen, M. Sydow, and S. Kramer. Analysis of respiratory pressure-volume curves in intensive care medicine using inductive machine learning. *Artificial Intelligence in Medicine*, 26(1-2):69–86, Sep-Oct 2002.

[4] F. Guiza, J. Ramon, and H. Blockeel. Gaussian processes for prediction in intensive care. In N. D. Lawrence, A. Schwaighofer, and J. Quinonero, editors, *Proceedings of the Gaussian Processes in Practice Workshop*, Bletchley Park, U.K., June 2006.

[5] G. A. Miller. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, 63:81–97, 1956.

[6] K. Morik, P. Brockhausen, and T. Joachims. Combining statistical learning with a knowledge-based approach – a case study in intensive care monitoring. In *Proceedings of the 16th International Conference on Machine Learning (ICML-99)*, Bled, Slovenia, 1999.

[7] K. Morik, M. Imhoff, P. Brockhausen, T. Joachims, and U. Gather. Knowledge discovery and knowledge validation in intensive care. *Artificial Intelligence in Medicine*, 19(3):225–249, 2000.

[8] S. Moser, W. Jones, and S. Brossette. Application of data mining to intensive care unit microbiologic data. *Emerging Infectious Diseases*, 5(3):454–457, May-Jun 1999.

[9] J. Ramon, F. Guiza, D. Fierens, G. Meyfroidt, H. Blockeel, M. Bruynooghe, and G. Van Den Berghe. Mining data from intensive care patients. *Advanced Engineering Informatics*, 2005. submitted.

[10] M. Verduijn, N. Peek, F. Voorbraak, E. de Jonge, and B. de Mol. Dichotomization of ICU length of stay after cardiac surgery based on model precision. In A. Nowé, T. Lenaerts, and K. Steehaut, editors, *Benelearn 2004: Proceedings of the Annual Machine Learning Conference of Belgium and the Netherlands*, pages 162–169, 2004.

[11] M. Verduijn, N. Peek, F. Voorbraak, E. de Jonge, and B. de Mol. Dichotomization of ICU length of stay based on model calibration. In S. Miksch, J. Hunter, and E. Keravnou, editors, *Proceedings of the 10th Conference on Artificial Intelligence in Medicine (AIME 2005)*. Springer-Verlag, 2005.

[12] Y. Vodovotz, G. Clermont, C. Chow, and G. An. Mathematical models of the acute inflammatory response. *Current Opinion in Critical Care*, 10:383–390, 2004.