Published in S. Miksch, J. Hunter, E. Keravnou (eds.): Artificial Intelligence in Medicine. Proceedings of the 10th Conference on Artificial Intelligence in Medicine, AIME 05, Aberdeen, Scotland, July 2005. Springer Verlag, Berlin, Heidelberg, New York, 2005. (c) Springer Verlag

Mining Clinical Data: Selecting Decision Support Algorithm for the MET-AP System

Jerzy Blaszczynski[1], Ken Farion[2], Wojtek Michalowski[3], Szymon Wilk[1]*, Steven Rubin[2], and Dawid Weiss[1]

[1] Poznan University of Technology, Institute of Computing Science, Piotrowo 3a, 60-965 Poznan, Poland
{jurek.blaszczynski, szymon.wilk, dawid.weiss}@cs.put.poznan.pl
[2] Children's Hospital of Eastern Ontario, 401 Smyth Road, Ottawa, ON, K1H 8L1, Canada
{farion, rubin}@cheo.on.ca
[3] University of Ottawa, School of Management, 136 Jean-Jacques Lussier Street, Ottawa, ON, K1N 6N5, Canada
wojtek@management.uottawa.ca

Abstract. We have developed an algorithm for triaging acute pediatric abdominal pain in the Emergency Department using the discovery-driven approach. This algorithm is embedded into the MET-AP (Mobile Emergency Triage - Abdominal Pain) system – a clinical decision support system that assists physicians in making emergency triage decisions. In this paper we describe experimental evaluation of several data mining methods (inductive learning, case-based reasoning and Bayesian reasoning) and results leading to the selection of the rule-based algorithm.

1 Introduction

Quest to provide better quality care to the patients results in supplementing regular procedures with intelligent clinical decision support systems (CDSS) that help to avoid potential mistakes and overcome existing limitations.

A CDSS is "a program designed to help healthcare professionals make clinical decisions" [1]. In our research we concentrate on systems providing patient-specific advice that use clinical decision support algorithms (CDSAs) discovered from data describing past experience. We evaluate various methods of data mining for extracting knowledge and representing it in form of a CDSA. The evaluation is based on the decision accuracy of the CDSA as the outcome measure.

CDSAs discussed in this paper aim at supporting triage of abdominal pain in the pediatric Emergency Department (ED). Triage is one of ED physician's functions that requires the following disposition decisions: discharge (discharge and possible follow-up by a family doctor), observation (further investigation in the ED or hospital), or consult (consult a specialist).

* Corresponding author

MET-AP [2] is a mobile CDSS that provides support at the point of care for triaging abdominal pain in the ED. It suggests the most appropriate dispositions and their relative strengths.

The paper is organized as follows. We start with a short overview of data mining methodologies used for developing CDSAs. Then, we describe the experiment and conclude with the discussion.

2 CDSA and Data Mining

One of the oldest data mining approaches applied in developing discovery-driven CDSA is Bayesian reasoning [3]. It is a well-understood methodology that is accepted by physicians [3], as it mimics clinical intuition [4]. Bayesian reasoning was adopted in the Leeds system [5] for diagnosing abdominal pain.

Other techniques of data mining, including case-based reasoning [3] and inductive learning [4], have not been as widely accepted in clinical practice [4]. However, some practical applications [3, 4] of CDSAs constructed with these techniques have attracted the attention of healthcare professionals.

Case-based reasoning [6] allows arriving at a recommendation by analogy. This approach was employed in the development of CDSA for the CAREPARTNER system [7] designed to support the long-term follow-up of patients with stem-cell transplant.

Inductive learning [4] represents discovered knowledge in the form of decision rules and decision trees that are later used in CDSAs. It is less often used in practice despite some successful clinical applications (e.g., for evaluating coronary disease [8]).

3 Selecting the Most Appropriate CDSA

In order to select the CDSA for triage of abdominal pain we designed and conducted a three-phase computational experiment.

The first phase of the experiment dealt with the development of CDSAs on the basis of learning data transcribed retrospectively from ED charts of patients with abdominal pain, who visited the ED in the Children's Hospital of Eastern Ontario between 1993 and 2002 (see Table 1). Each record was described by values of 13 clinical attributes. Charts were assigned to one of three decision classes by reviewing the full hospital documentation.

The following data mining methods were evaluated: induction of decision rules (LEM2 algorithm [9]); Bayesian reasoning (naive Bayesian reasoning [10]); case-based reasoning (IBL algorithm [11]); and induction of decision trees (C4.5 algorithm [12]).

During the second phase of the experiment we tested performance of the algorithms on new retrospectively collected independent testing data set (see Table 1). There is general consensus among the physicians that it is important to differentiate between the classification accuracies obtained for different decision classes (high accuracy for the consult class is more important than for

Table 1. Learning and testing data sets

	Number of charts	
Decision class	Learning set	Testing set
Discharge	352	52
Observation	89	15
Consult	165	33
Total	606	100

the discharge class). This means than the assessment and selection of the most appropriate CDSA should consider not only its overall classification accuracy, but also performance for critical classes of patients.

The results of the second phase are given in Table 2. Although all CDSAs had very similar overall accuracies, the most promising results considering the observation and consult classes were obtained for case-based CDSA and for the rule-based CDSA. The tree-based CDSA had the same accuracy for the consult class as the rule-based one, however, its performance for the observation class was much lower. Finally, naive Bayesian CDSA offered the highest accuracy for the discharge class, and the lowest accuracy for the consult class, which is unacceptable from the clinical point of view.

Table 2. Classification accuracy of the CDSAs

CDSA	Overall	Discharge	Observation	Consult
Rule-based	59.0%	55.8%	46.7%	69.7%
Naive Bayesian	56.0%	65.4%	20.0%	57.6%
Case-based	58.0%	57.7%	20.0%	75.8%
Tree-based	57.0%	59.6%	20.0%	69.7%

In the third phase of the experiment we tested and compared the CDSAs considering the misclassification costs (in terms of penalizing undesired misclassifications). The costs, given in Table 3, were defined by the physicians participating in the study and introduced into the classification scheme. We excluded rule-based CDSA from this phase, as its classification strategy did not allow cost-based evaluation.

The results of the third phase are given in Table 4. The most remarkable change comparing to the second phase is the increased classification accuracy for the observation class resembling the way a cautious physician would proceed. Such an approach also has drawbacks – the decreased classification accuracy for the consult class.

Table 3. Misclassification costs

	Recommendation		
Outcome	Discharge	Observation	Consult
Discharge	0	5	10
Observation	5	0	5
Consult	15	5	0

Table 4. Classification accuracy of the CDSAs (costs considered)

CDSA	Overall	Discharge	Observation	Consult
Naive Bayesian	56.0%	59.6%	46.7%	54.6%
Case-based	49.0%	42.3%	60.0%	54.6%
Tree-based	55.0%	59.6%	40.0%	54.6%

Comparison of the performance of the cost-based CDSAs with the performance of the rule-based CDSA from the second phase favors the latter one as the preferable solution for MET-AP. This CDSA offers high classification accuracies for crucial decision classes (observation and consult) without decreasing accuracy in the discharge class. It is important to note that a rule-based CDSA represents clinical knowledge similarly to practice guidelines, and thus it is easy for clinicians to understand. Taking all of the above into account, we have chosen to implement the rule-based CDSA in MET-AP.

4 Discussion

Our experiment showed better performance of the rule-based CDSA in terms of classification accuracy for crucial decision classes (observation and consult), even when misclassification costs were introduced. This observation proves the viability and robustness of the rule-based CDSA implemented the MET-AP system.

More detailed evaluation of the results suggested that the learning data did not contain obvious classification patterns. This was confirmed by a large number of rules composing the rule-based CDSA (165 rules) and their relatively weak support – for each class there were only a few stronger rules. This might explain why rule-based CDSA and case-based CDSA that consider cases in an explicit manner (a rule can be viewed as a generalized case) were, in overall, superior to other algorithms.

During further studies we plan to research for more possible dependencies between the characteristics of the analyzed data and data mining methods used to build CDSAs. This should enable us to select the most appropriate CDSA in advance without a need for computational experiments. Moreover, it will be worth checking how the integration of different CDSAs impacts the classification accuracy.

5 Acknowledgments

Research reported in this paper was supported by the NSERC-CHRP grant. It was conducted when J. Blaszczynski, Sz. Wilk, and D. Weiss were part of the MET Research at the University of Ottawa.

References

1. Musen, M., Shahar, Y., Shortliffe, E.: Clinical decision support systems. In Shortliffe, E., Perreault, L., Wiederhold, G., Fagan, L., eds.: Medical Informatics. Computer Applications in Health Care and Biomedicine. Springer-Verlag (2001) 573–609
2. Michalowski, W., Slowinski, R., Wilk, S., Farion, K., Pike, J., Rubin, S.: Design and development of a mobile system for supporting emergency triage. Methods of Information in Medicine **44** (2005) 14–24
3. Hanson, III, C.W., Marshall, B.E.: Artificial intelligence applications in the intensive care unit. Critical Care Medicine **29** (2001) 427–35
4. Kononenko, I.: Inductive and Bayesian learning in medical diagnosis. Applied Artificial Intelligence **7** (1993) 317–377
5. de Dombal, F.T., Leaper, D.J., Staniland, J.R., McCann, A.P., Horrocks, J.C.: Computer-aided diagnosis of acute abdominal pain. British Medical Journal **2** (1972) 9–13
6. Schmidt, R., Montani, S., Bellazzi, R., Portinale, L., Gierl, L.: Cased-based reasoning for medical knowledge-based systems. International Journal of Medical Informatics **64** (2001) 355–367
7. Bichindaritz, I., Kansu, E., Sullivan, K.: Case-based reasoning in CAREPARTNER: Gathering evidence for evidence-based medical practice. In Smyth, B., Cunningham, P., eds.: Advances in Case-Based Reasoning: 4th European Workshop, Proceedings EWCBR-98, Berlin, Springer-Verlag (1998) 334–345
8. Krstacic, G., Gamberger, D., Smuc, T.: Coronary heart disease patient models based on inductive machine learning. In Quaglini, S., Barahona, P., Andreassen, S., eds.: Proceedings of 8th Conference on Artificial Intelligence in Medicine in Europe (AIME 2001), Berlin, Springer-Verlag (2001) 113–116
9. Grzymala-Busse, J.: LERS - a system for learning from examples based on rough sets. In Sowiski, R., ed.: Intelligent Decision Support - Handbook of Applications and Advances of the Rough Set Theory. Kluwer Academic Publishers, Dordrecht/Boston (1992) 3–18
10. John, G.H., Langley, P.: Estimating continuous distributions in Bayesian classifiers. In: Proceedings of the 11th Conference on Uncertainty in Artificial Intelligence. Morgan Kaufmann, San Mateo (1995) 338–345
11. Aha, D., Kibler, D., Albert, M.: Instance-based learning algorithms. Machine Learning **6** (1991) 37–66
12. Quinlan, R.: C4.5: Programs for Machine Learning. Kaufmann Publishers, San Mateo, CA (1993)