

Automated Decision Support and Guideline Verification in Clinical Practice

R Mlynarski¹, G Ilczuk², E Pilat¹, A Wakulicz-Deja³, W Kargul¹

¹Electrocardiology Department of Silesian Medical Academy, Katowice, Poland

²HEITEC AG Systemhaus fuer Automatisierung und Informationstechnologie, Erlangen, Germany

³Institut of Informatics University of Silesia, Sosnowiec, Poland

Abstract

Applying international guidelines in medical, including cardiological, therapies is a guarantee of safe and modern treatment. Unfortunately, standards are often not obeyed.

In this paper we present an experimental software program based on rough sets methods. The main aim of this application is to improve patient care and help the decision process using guidelines verification. We concentrate on the practical aspects using these methods. Examples and clinical tests, which were based on real-life data of our patients, show that the accuracy of results reached on a large group of patients could be acceptable in clinical practice.

1. Introduction

EBM – Evidence Based Medicine has become very common over the past few years in the whole medical world. The basis of EBM determined the multidisciplinary and multicenter researches which were carried out on a large group of patients. These trials are a source of information for experts in many fields of medicine including cardiology [1]. For example, the latest guidelines update for the management of patients with chronic stable angina were published by ACC/AHA in 2002, but every year many standards in many subdivisions of cardiology are published [2]. Applying international guidelines and standards in medical therapies is a guarantee of safe treatment. Unfortunately, standards are often not obeyed. There are many possible explanations for this fact. The most important may be fact that doctors often do not have enough time to study these documents. To help doctors improve their decision making process and to make treatment safe for patients, we decided to create a computerised tool, which could be very helpful for eg. younger doctors who are less experienced. And because experience comes with time, novel mathematical technologies like rough sets implemented into user-friendly software could be very useful.

1.1. Aim of the study

The aim of this study was to prepare rough sets based software and to test this application in practice using real-life data from an Electrocardiology Clinic.

2. Methods

We divided the work of preparing a software program into following sections.

2.1. Creating the database

The first step was to create the clinical database. In hospitals there are many commercial and open-source databases. Unfortunately, most of them have been prepared keeping statistical analysis and patient's documentation in mind and they need some adaptations before they can be useful in research. Nevertheless, they contain a lot of valuable data which can be very useful for doctors if it was available.

For our research we imported data about 2039 patients hospitalized between 2003-2005 in the Electrocardiology Department of the Silesian Medical Academy (Katowice, Poland) from an internal hospital system. Then we extracted information about current state of the patient and patient's drug treatment saved in the form of free-text reports and coded this information into our database using binary values so that, for example, if a phrase "Morbus ischaemicus cordis" (eng. Ischaemic heart disease) was found in the description of the patient status, a value of 1 was assigned to a column named "I25.1". This recognition process will be extended and automated in the final version of the described system.

This data pre-processing phase produced a data set containing 2039 objects described by a value of 84 binary attributes (columns). A value of each attribute was set to 1 if a disease was diagnosed and to 0 otherwise. To demonstrate the method presented in this paper we joined some of the 84 attributes into 14 grouped attributes as shown in Table 1.

Table 1. Groups of attributes and their description.

Attribute	Description
AVBL	Atrioventricular block
DIAB	Diabetes
PTACH	Paroxysmal tachycardia
HYPERCHOL	Hypercholesterolaemia
CARDIOMYO	Cardiomyopathy
ATHEROSC	Atherosclerosis
AFF	Atrial fibrillation and flutter
HYPERTEN	Hypertension (High Blood Pressure)
CIHD	Chronic ischaemic heart disease
OBESITY	Obesity
SSS	Sick Sinus Syndrome
PACEMAKER	Pacemaker stimulation
TYROIDG	Disorders of thyroid gland
MIOLD	Myocardial infarction in past

As a decision attribute we chose two groups of drugs commonly prescribed in treating heart diseases: ACE Inhibitors (ACE – Angiotensin Converting Enzyme) and Beta-Blockers.

2.2. Rough sets

Rough sets theory proposed by Pawlak in the 1980s has already been applied in many machine learning, knowledge discovery and expert systems [3-6]. The ability of rough sets to handle imprecision and uncertainty in input data without any preliminary or additional information about it made this approach very common in medical domain.

In this paper we present the results of MLEM2 rule induction algorithm, which is based on the original LEM2 algorithm [7,8]. The main advantage of the MLEM2 in comparison to LEM2 is a smaller number of rules from the same input data, which are more understandable and verifiable by humans which in the end improves their testability.

2.3. Decision trees

The algorithm rules generated from MLEM2 were used to build a kind of decision tree to be presented to domain experts from the Electrocardiology Department of the Silesian Medical Academy for evaluation.

2.4. Creating software

The experimental software used for all presented calculations was written in Java version 1.4. We also used a relational PostgreSQL database for data storage.

3. Results

We tested the accuracy of the generated rules using 10-times repeated 10-fold cross-validation with instance randomization after each run. This methodology provides a very stable replicability as presented in [9]. Table 2 shows the results achieved in summary.

Table 2. Overall results of accuracy of the rough sets based methods using 10-times repeated 10-fold cross-validation with instance randomization after each run

Decision attribute	Number of rules	Accuracy [%]
ACE	13	76,60
Beta-Blocker	28	66,80

From table 2 it can be seen that the accuracy of classification after cross-validation is higher for the ACE decision attribute than for the Beta-Blockers. Also the number of generated rules, which is more than two times higher for the Beta-Blockers attribute which could be attributable to the increased complexity of correct prediction.

The overall high prediction accuracy for the ACE attribute shows that the selection of the grouped attributes was good enough to generate strong decision rules. An example of the generated rules for the ACE decision attribute is shown below:

Rule 1: (HYPERTEN=1)&(MIOLD=0)&(DIAB=0)=>
ACE=yes=715, ACE-no=183

Rule 2: (HYPERTEN=0)&(DIAB=0)&(MIOLD=0)=>
ACE=yes=120, ACE-no= 292

Rule 3: (HYPERTEN=0)&(DIAB=1)&(MIOLD=0)=>
ACE=yes=230, ACE-no=55

Rule 4: (DIAB=0)&(MIOLD=1)&(HYPERTEN=1)=>
ACE=yes=166, ACE-no=40

Rule 5: (MIOLD=1)&(HYPERTEN=0)&(DIAB=0)=>
ACE=yes=68, ACE-no=30

Each presented rule consists of two parts: the first one describes conditions which must be fulfilled to make a decision (in this example to classify an object to the class ACE=yes or ACE-no). Which decision will be made depends on the number of cases the rule correctly classify during the learning phase. These numbers are shown in a second line for each rule. From this short example it can be seen that for a larger number of generated rules a human ability to validate them decreases. To avoid this

situation we are working on a method to preset the rules in the form of a decision tree as shown in figure 1.

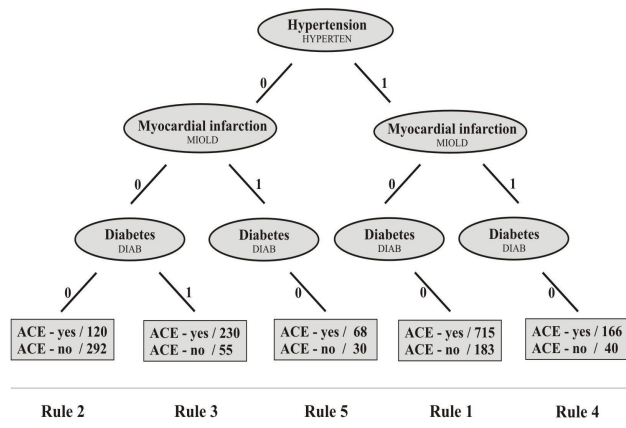


Figure 1. The example of a generated decision tree showing treatment by ACE in patients with or without hypertension, myocardial infarction and diabetes.

This graphical and hierarchical result representation allows knowledge management and validation at different levels of details in a way which is more suitable for human consumption.

4. Discussion

In this paper we have shown an example of a decision system built based on information collected from free-text narrative medical reports. We used MLEM2 algorithm to generate decision rules and then used two methods to test their accuracy: 10-times repeated 10-fold cross-validation and validation by domain experts. For the human validation we transformed the generated rules into decision trees, which speeded-up this process. The achieved results were then discussed by the Electrocardiology Department staff and the most experienced doctors.

It is almost impossible for humans to check hundreds of rules generated by the computer. We try to resolve this problem by random generating examples of rules. The next step was to check each of them step by step by the most experienced doctors from the Electrocardiology Department. We divided the analysis into 2 sections: correctness use of ACE and separately Beta-blockers. Using decision tree algorithms, we generated a graph of 9 common diseases and 27 combinations of them (e.g. atrial fibrillation, AV-blocks, ischemia etc.). In each case use of ACE and Beta-blockers was checked and verified by comparison with the latest international standards using these drugs.

The results were quite satisfactory. In 33 cases (91.7%), the results generated by the computer were correct and were according to the guidelines. In one case it was wrong (2.8%) and in 2 cases (5.6%) we have queries about the connection of “disease – drug” generated by the computer (e.g. B-blocker in hypercholesterolemia). Even in these 2 cases the decision was correct (B-blocker = 0). In ACE we have the following results: in 86,1% decision was correct and in 13,8% the answer was wrong. In Figure 2 we present a statistical graph which compares accuracy in % achievement results by domain experts with the results achievements by computer.

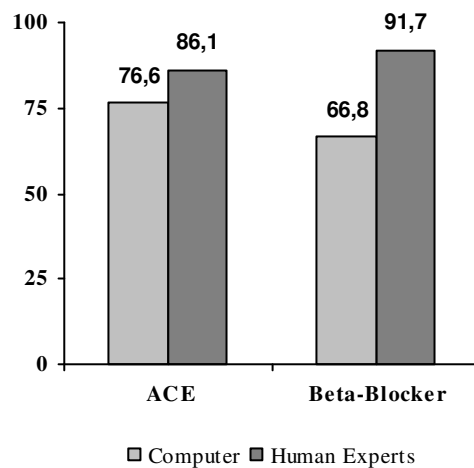


Figure 2. Statistical graph presenting in % accuracy of achievement results according to guidelines.

Connected results from all combination disease-drug were similar and an overall correct decision was taken in 89.7% of cases. These events are very promising. In our opinion the differences are caused by the fact that we analysed actual patients. In some of their cases using some drugs is recommended in primary prevention.

4.1. Conclusions

Based on results we may extract the following conclusions:

1. Rough sets systems can help in the process of diagnosis or treatment
2. This software can also be useful in automated decision support in clinical practice
3. Additional research is necessary to determine whether this kind of application can be used in clinical practice

Acknowledgements

We would like to thank Anna Drzewiecka MD from the Electrophysiology Department of the Silesian Medical Academy in Katowice for giving us feedback and Maciej Pacyno MS from the statistical department of Upper Silesia Medical Center for his assistance in exporting the databases.

References

- [1] Haynes RB, Devereaux PJ, Guyatt GH. Physicians' and patients' choices in evidence based practice. *BMJ* 2002; 324: 1350
- [2] Gibbons RJ, Abrams J, Chatterjee K, Daley J, Deedwania PC, Douglas JS, Ferguson TB Jr, Fihn SD, Fraker TD Jr, Gardin JM, O'Rourke RA, Pasternak RC, Williams SV. ACC/AHA 2002 guideline update for the management of patients with chronic stable angina - summary article: a report of the American College of Cardiology / American Heart Association Task Force on Practice Guidelines. *J Am Coll Cardiol* 2003; 41:159-68.
- [3] Pawlak Z. Rough Sets. *International Journal of Information and Computer Science* 1982; 11(5): 341-356
- [4] Pawlak Z. Rough Sets: Theoretical Aspects of Reasoning about Data. Kluwer Academic Publishers 1992: 229
- [5] Lin TY, Cercone N. Rough Sets and Data Mining: Analysis of Imprecise Data 1997. Kluwer Academic Publishers, 1997: 259-260
- [6] Lin TY, Yao YY, Zadeh L. Data Mining, Rough Sets and Granular Computing. Physica-Verlag 2002: 406-413
- [7] Grzymala-Busse JW. MLEM2 - Discretization During Rule Induction, Proceedings of the IIPWM'2003, Springer-Verlag 2003: 499-508,
- [8] Chan C, Grzymala-Busse JW. On the attribute redundancy and the learning programs Id3, PRISM, LEM2, Department of Computer Science, University of Kansas 1991: 91-14
- [9] Bouckaert R, Eibe F. Evaluating the Replicability of Significance Tests for Comparing Learning Algorithms, *PAKDD* 2004: 272-281

Address for correspondence

Cardiological section

Rafal Mlynarski
Klinika Elektrokardiologii SAM
ul. Ziolowa 45/47
Katowice 40-635, Poland
email: Rafal_Mlynarski@mp.pl

Computer section

Grzegorz Ilczuk
Am Europakanal 36
91056 Erlangen., Germany
email: Grzegorz.Ilczuk@Ilczuk.com