

How Decision System Trained on a Large Database Recognizes New Cases – Prelude before Clinical Implementation

R Mlynarski¹, A Wlodyka¹, G Ilczuk², E Pilat¹, W Kargul¹

¹Upper Silesian Cardiology Centre, Katowice, Poland

²University of Silesia, Sosnowiec, Poland

Abstract

In the research presented here, a joint team (cardiologists and software specialists) focused on the evaluation of previously created, multi-stage decision systems trained on real clinical data. The question is how a decision system trained on a large database recognizes new cases remains unanswered? Obtained results seem to confirm that decision systems trained on a large dataset works well on new data – which may confirm its usefulness in clinical practice. An acceptable level of accuracy for doctors' decisions was achieved. For more specialist areas (like pacemakers implantation) where strict guidelines are available, the accuracy was higher.

1. Introduction

Since the implantation of the first cardiac pacemaker in 1958 by Ake Sening and Rune Elmqvist in Stockholm, there has been gigantic progress in this field of medicine including both the range the equipment and the techniques of implantation. Today electrotherapy is a rapidly developing field of invasive cardiology [1,2,3]. There is a wide range of devices which can be implanted. The most important are:

- Cardiac pacemakers
- ICD - Implanted Cardioverter Defibrillators
- CRT - devices to Cardiac Resynchronization Therapy
- CRT - devices to biventricular stimulation with ICD functions.

There is wide group of indications to the implantation of earlier mentioned devices such as:

- * sins node dysfunction (SND),
- * A-V blocks (AVB),
- * atrial fibrillation and flutter (AF),
- * ventricular tachycardia (VT),
- * sustained ventricular tachycardia (sVT),
- * ventricular fibrillation (VF),
- * heart failure (HF) and many more.

Huge number of indications can cause problems with proper qualification of patients and next with follow-up.

The most common drugs in general cardiology as well as in the Electrocardiology are ACE (Angiotensin-converting Enzyme) inhibitors and B-blockers [4,5]. Both this groups have different indications including:

- * anti arterial hypertension,
- * slowing heart rhythm,
- * prevention of cardiovascular disorders,
- * congestive heart failure (CHF),
- * left ventricular dysfunction,
- * prevention of nephropathy in diabetes mellitus

In the research presented here, a joint team (cardiologists and software specialists) focused on the evaluation of previously created, multi-stage decision systems trained on real clinical data.

However, the question of how a decision system trained on a large database recognizes new cases remains unanswered?

1.1. Aim

Evaluate how a decision system trained on a large database recognizes new cases, using data about implantation of cardiac pacemakers and pharmacotherapy in general cardiology.

2. Methods

The decision system presented was trained using the medical records of 5425 patients hospitalized in Electrocardiology Department in 2003-2006. The data were imported from a clinical information system and transformed to binary attributes using our, based on regular expressions, hierarchical dictionary algorithm. Afterwards based on domain experts knowledge these binary attributes were joined into 14 grouped attributes. Using this data as input, decision rules were generated (Rough Sets Grzymala-Busse' MLEM 2; our own implementation).

Elements of the proposed system are presented below:

1. Import subsystem – responsible for importing data from medical information systems into our storage subsystem.

2. Preprocessing subsystem – transforms raw data into a form suited for further data processing. Additionally, noise and redundant data are removed based on a statistical analysis.

3. Feature selection module – responsible for selecting an optimal set of attributes for a generation of decision rules.

4. Rule induction subsystem – uses algorithms based on Rough Sets MLEM2 algorithm for generating decision rules. Rough Set is a mathematical approach to handle imprecision and uncertainty was developed by Pawlak and presented in 1982 [6,7].

5. Visualization module – transforms the collected knowledge to a form which is easily understandable and verifiable by humans. In this paper we used our transformation results achieved with an extended version of AQDT-2 method based on idea presented by Michalski [8] and/or with using rule-diagrams.

A set of methods to cover all stages necessary for a complete decision system were created and implanted in a user-friendly software based on JAVA 6.0.

System was presented in previous Computers in Cardiology Congresses in Lyon, Valencia and Durham [9-12]. In this year accuracy of prediction was improve and additional method of visualization (rule-diagrams) was added.

Accuracy of generated decision rules was validated by the system using 10-times repeated 10-fold cross-validation with a set of a new data – 198 patients hospitalized in 2007. In addition, the medical relevance of the generated rules was checked by cardiology experts.

3. Results

Selected results are presented in the table 1 and table 2 as accuracy of prediction in % (train dataset / test dataset) in the selected decisions connected with cardiac pacemaker implantation. Additionally number of generated decision rules was also presented.

DECISION	Pacemaker implantation		
ACCURACY	<i>TRAIN</i> 87,92%	<i>TEST</i> 85,32%	15 rules

DECISION	Pacemaker DDD type implantation		
ACCURACY	<i>TRAIN</i> 89,43%	<i>TEST</i> 75,95%	21 rules

DECISION	Pacemaker VVI type implantation		
ACCURACY	<i>TRAIN</i> 83,56%	<i>TEST</i> 70,95%	43 rules

Table 1. Accuracy in % for train and test subset of data for chosen examples of implantation cardiac pacemakers.

Pharmacotherapy in general cardiology was also tested using the 2 main groups of cardiological drugs: ACE inhibitors and B-blockers. Results for this group are presented in the table 2.

DECISION	ACE inhibitor		
ACCURACY	<i>TRAIN</i> 79,67%	<i>TEST</i> 68,05%	11 rules

DECISION	B-blockers		
ACCURACY	<i>TRAIN</i> 78,45%	<i>TEST</i> 73,54%	23 rules

Table 2. Accuracy in % for train and test subset of data for using pharmacotherapy in general cardiology.

Additionally in the table 3 average differences between test and train datasets. Minus (-) value means that accuracy of the system is lower in test dataset if compare to the train dataset.

	Difference in % <i>Test - Train</i>	p
Pacemaker implantation	-2,6%	$p=0,12$
PM* DDD type implantation	-13,48%	$P=0,00$
PM* VVI type implantation	-12,61%	$p=0,00$
ACE inhibitor	-11,62%	$p=0,06$
B-blockers	-4,91%	$p=0,01$

Table 3 Difference in % between test and train dataset. $P < 0,05$ = value statistically significant
*PM = pacemaker

4. Discussion and conclusions

Success of machine learning algorithms is usually dependent on a quality of a dataset they operate on. For datasets containing noisy, inadequate or irrelevant information these algorithms may produce less accurate results [13]. Therefore a common pre-processing step in data mining domain is a selection of highly predictive attributes. In this case study we select subsets of attributes using two different feature selection algorithms and compare achieved results for their medical appliance with original, non-reduced datasets [14].

This stage of the software development includes, but is not limited to, the importing data from a clinical database infrastructure, preprocessing data (removing noisy and irrelevant data, converting free text reports to binary attributes, joining binary attributes into group attributes, generating decision rules using our own implementation of the MLEM2 algorithm and finally visualizing the results.

In our experiments we selected subsets of attributes both from original training sets and from training sets after applying some noise reduction algorithms (over 1800 combinations pro decision attribute). The selected subsets were than used to generate decision rules using MLEM2 algorithm [6,7,15]. Afterwards two test sets were classified using these decision rules and results of these classifications were compared with results achieved with J48 and Decision Table from Weka package. In all cases our methods performed just as well as these two commonly applied algorithms. We saw some problems with classification for SSI and VDD decision attribute but we think, that they can be overcome with a extended classification algorithm. The most important result was for us a positive verification of selected attributes done by domain experts. In all

generated subsets there were attributes which are most important from medical perspective and this result was particularly visible in case of χ^2 . Nevertheless CFS is an interesting algorithm because of its ability to evaluate a complete set of attributes. If this feature is combined with a BestFirst search strategy then CFS can be used to estimate a number of needed attributes.

From clinical point of view, key seems to be possibility to recognize new cases based rules generated for the huge test dataset. In this experiment test data set was 5425 patients hospitalized in Electrophysiology Department in 2003 – 2006. Decision rules generated based on this dataset was tested on the completely new dataset of 198 patients hospitalized in 2007.

According to our opinion accuracy was lower then on trained dataset, however difference seems to be acceptable. Only in 2 cases – decision about DDD and VVI pacemaker implantation – differences was statistically significant ($p < 0,05$), but still very high.

We asked domain experts from the Electrophysiology Department of Upper Silesian Cardiology Center in Katowice, Poland for support. They agree that accuracy is still not ideal, but in some cases can be useful for decision support.

Additional research are necessary to increase accuracy and fully confirm usefulness in clinical practice.

4.1. Conclusions

The results presented seem to confirm that decision systems trained on a large dataset works well on new data – which may confirm its usefulness in clinical practice. An acceptable level of accuracy for doctors' decisions was achieved. For more specialist areas (like pacemakers implantation) where strict guidelines are available, the accuracy was higher.

Acknowledgements

We would like to thank Iwona Grzegorzczak MS from the statistical department of the Upper Silesia Medical Center for her assistance in exporting the databases.

References

- [1] Toff WD, Camm AJ, Skehan JD; United Kingdom Pacing and Cardiovascular Events Trial Investigators. Single-chamber versus dual-chamber pacing for high-grade atrioventricular block. *N Engl J Med.* 2005 Jul 14;353(2):145-55.
- [2] Vardas PE, Auricchio A, Blanc JJ, Daubert JC, Drexler H, Ector H, Gasparini M, Linde C, Morgado FB, Oto A, Sutton R, Trusz-Gluza M. Guidelines for cardiac pacing and CRT: the task force for cardiac pacing and CRT of the ESC. *Eur Heart J* 2007; 28:2256-95.

- [3] Bellott P.H., Byrt C.L. Recent developments in pacemaker implantation and lead retrieval. W: Barold S.S., Mugica J. red. New perspectives in cardiac pacing 2. Futura Publishing Company, Inc. Mount Kisco, Nowy Jork 1991, 105-131.
- [4] Molinaro G, Cugno M, Perez M, et al. Angiotensin-converting enzyme inhibitor-associated angioedema is characterized by a slower degradation of des-arginine - bradykinin. *J Pharmacol Exp Ther* 2002;303:232-7.
- [5] Benedict CR. Centrally acting antihypertensive drugs: re-emergence of sympathetic inhibition in the treatment of hypertension. *Curr Hypertens Rep.* 1999;1:305-12.
- [6] Pawlak Z, Skowron A. A Rough Set approach to decision rules generation. *Inst. Comput. Sci. Research Report* 1993; 23.
- [7] Pawlak Z. *Rough Sets: Theoretical aspects of reasoning about data.* Boston, USA: Kluwer Academic Publishers 1991.
- [8] Michalski RS, Ibrahim FI. Learning Problem-Oriented Decision Structures from Decision Rule: The AQDT-2 System. *Lecture Notes in Computer Science* 1994;869: 416-426.
- [9] Ilczuk G, Mlynarski R, Kargul W, Wakulicz-Deja A. New feature selection methods for qualification of the patients for cardiac pacemaker implantation. *Computers in Cardiology* 2007;34:423-426
- [10] Mlynarski R, Ilczuk G, Wakulicz-Deja A, Kargul W. A New Method of Data Preparation for Cardiological Decision Support. *Computers in Cardiology* 2006;33:273-276.
- [11] Mlynarski R, Ilczuk G, Pilat E, Wakulicz-Deja A, Kargul W. Automated Decision Support and Guideline Verification in Clinical Practice. *Computers in Cardiology* 2005;32:375-378
- [12] Ilczuk G, Mlynarski R, Wakulicz-Deja A, Drzewiecka A, Kargul W. Rough Set Techniques for Medical Diagnosis Systems. *Computers in Cardiology* 2005;32:837-840
- [13] Jensen R, Shen Q. Fuzzy-rough data reduction with ant colony optimization. *Fuzzy Sets and Systems* 2005;149:5-20.
- [14] Ilczuk G, Wakulicz-Deja A. Rough sets approach to medical diagnosis System. *Lecture Notes in Computer Science* 2005;3528:204–210.
- [15] Ilczuk G, Wakulicz-Deja A. Attribute selection and rule generation techniques for medical diagnosis systems. *lecture notes in computer science* 2005;3642:352-361.

Address for correspondence

Computer section

Grzegorz Ilczuk
Heuweg 12A
91334 Hemhofen, Germany
email:
Grzegorz.Ilczuk@Ilczuk.com

Cardiological section

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