IOSR Journal of Engineering (IOSRJEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
Vol. 08, Issue 12 (December. 2018), ||V (II) || PP 54-56

Modeling Crohn's disease and colon disease diagnosis using rules actionability

Agnieszka Dardzinska[1,a], Anna Kasperczuk[1,b] , Osman Gürdal1[2,c]
[1]Bialystok University of Technology, Department of Biocybernetics and Biomedical Engeenering,ul. Wiejska 45a, 15-351 Bialystok, Poland
[2]SuleymanDemirel University School of Medicine, 32069 Isparta, Turkey
Corresponding Author: Agnieszka Dardzinska

Abstract: This paper presents the method binding statistical and data mining techniques, which aims to support the decision-making process in selected diseases in digestive systems. Currently, there is no precise diagnosis for these two diseases: ulcerative colitis (UC) and Crohn's disease (CD). Specialist physicians must exclude many other diseases occurring in the colon. The first goal of this study is a retrospective analysis of medical data of patients hospitalized in the Department of Gastroenterology and Internal Diseases and finding the symptoms differentiating the two analyzed diseases. The second goal is to build a system that clearly points to one of the diseases UC or CD, which shortens the time of diagnosis and facilitates the future treatment of patients. The work focuses on building a model that can be the basis for the construction of classifiers, which are one of the basic elements in the medical recommendation system.

Keywords: colon disease, decision system, information system, action rule, statistical methods

---

Date of Submission: 29-11-2018 Date of acceptance: 13-12-2018

I. INTRODUCTION

Let us assume that $S = (X, A, V)$ is an information system, where X is a nonempty, finite set of objects, A is a nonempty, finite set of attributes, V is a set of all attributes values[3,5,9].Then, $a : X \to V_a$ is a function for any $a \in A$, that returns the value of the attribute of a given object. The set of attributes can be divided into three subsets $A = A_1 \cup A_2 \cup D$. A_1 is a set of stable attributes, A_2 is a set of flexible attributes and a set of decision attributes is described by$D$[1,2,4,13]. By stable attributes we mean attributes with unchangeable values(e.g. age, gender), while for the flexible attributes the values can be changed (e.g. blood pressure, weight, haemoglobin level).Example of the information system $S = (X, A, V)$ is presented in Tab.1. The set of objects consists of five elements $X = \{x_1, x_2, x_3, x_4, x_5\}$. The set of attributes consists of two subsets [1,6,13] A_1 which includes stable attributes $\{a\}$, and A_2 which includes only flexible attributes $\{b, c\}$.

Table 1. Information System

Object	Stable attributes A_1	Flexible attributes A_2	
	Attribute a	**Attribute b**	**Attribute c**
x_1	a1	b1	H
x_2	a1	b2	H
x_3	a1	b3	I
x_4	a2	b3	I
x_5	a2	b2	H

Information systems can be also seen as decision tables. In Tab.2 we have decision System $S = (X, A, V \cup \{d\})$, with one stable attribute a, two flexible attributes b and c, and the decision attribute d.

Table 2. Decision System

Object	Stable attribute A_1	Flexible attributes A_2		Decision d
	Attribute a	**Attribute b**	**Attribute c**	
x_1	a1	b1	H	-
x_2	a1	b2	H	-
x_3	a1	b3	I	+
x_4	a2	b3	I	-
x_5	a2	b2	H	+

In this case we have objects classified initially into two separate classes $(d, +)$ and $(d, -)$. On the basis of knowledge extracted from decision systems we can model and simulate decision processes.This knowledge is represented as sets of connectionsbetween classification attributes and decision ones. Classification attributes are easily accessible and can be represented as measurements, parameters, personal data, etc.The decision is a value connected with the extracted knowledge, e.g. given straight by expert or coming from the observation (diagnose, medical treatment method). The decision allocated the objects from decision system forming in such way easy pattern model. Our goal is using the pattern model to classify new objects using their classification attributes.

II. CHRONIC ENTERITIS

Chronic intestinal inflammations have probably afflicted mankind since the dawn of time. As soon as science, discovered the world of microorganisms and proved their relationship with many diseases, people began to look for the causes of these ailments in infections. In the medical literature, the term chronic inflammatory bowel disease, also referred to as IBD (Inflammatory Bowel Disease), has been functioning since 1932, when New York doctor Burrill B. Crohn and his colleagues Leon Ginzburg and Gordon D. Oppenheimer publish a work in which they present a new disease unit called "regional ileitis" - from this point medicine knows the disease as Crohn's disease (Crohn's disease, CD) [2,11].

2.1. Crohn's disease (CD)

Crohn's disease is a chronic inflammatory disease that can affect every part of the digestive tract, from the mouth to the anus. Manifested symptoms depend on where the inflammatory changes are located. Generally,it can be distinguished: abdominal pain (usually in the right iliac fossa), bloating, weight loss, weakness and fever. It is also observed the formation of perianal fistulas and abscesses. Inflammatory changes in the gastrointestinal tract are discontinuous, which means that next to the inflamed mucous membrane, it is unchanged inflammation. In addition, inflammatory changes in addition to the mucous membrane also include all the other parts of the intestinal wall[2,11].

2.2. Ulcerative colitis (UC)

Like CD, it is a chronic inflammatory disease, affects only the colon. Symptoms include frequent diarrhea with an admixture of blood and / or mucus and painful pressure on the stool. Fistulas and abscesses are rare. Permanent inflammatory changes are localized along the entire length of the large intestine in the form of superficial ulcerations. The histological image shows granulocytes and lymphocytic infiltrates of the mucous membrane. At an advanced stage, i.e. after about 10 years of the disease, epithelial dysplasia changes leading to tumor degeneration may appear [6,11].

III. METHODOLOGY IN ACTIONABILITY PROCESS

Action rules are logical terms defining knowledge for desirable actions related to the hidden objects in a database. The intent here is to concentrate on objective measures for actionability, which is defined as the extent to which a user can gain benefits from the discovered patterns, such as in the medical domain [1,2,5]. Suppose an actionable goal of $\mathbf{r} = [\boldsymbol{\omega} * (\boldsymbol{\alpha} \rightarrow \boldsymbol{\beta}) \rightarrow (\boldsymbol{\theta} \rightarrow \boldsymbol{\psi})]$, where $\boldsymbol{\omega}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\theta}$, and ψ are descriptions of objects, e.g. in the case of patients, where $\mathbf{p}$ is described as the satisfactions of a designed condition and the changeable measure of $(\boldsymbol{\alpha} \rightarrow \boldsymbol{\beta})$ for patients who registered in a database with the expected result, $(\boldsymbol{\theta} \rightarrow \boldsymbol{\psi})$.There are two conceivable perspectives in terms of the strategies of actionability. One is the constituent of post-analysis at the back-end of the KD system [5,7]. This approach does not utilize the prior knowledge of the expert systems to lead the rule-generation process, which is purely subjective. The other approach is solely objective. It implements the input knowledge of the domain to control the rule generation process, which leads to determination of instrumental knowledge and comparing it with some standard beliefs. In this paper, we concentrate on the object-driven approaches of actionability. Object-driven patterns can be generated straightforward from the dataset, and then implemented for the final outcome [9, 10]. By object-driven action rule rin an information system S, the expression can be expressed as: $\mathbf{r} = [[(\mathbf{a1} = \boldsymbol{\omega}\mathbf{1}) * (\mathbf{a2} = \boldsymbol{\omega}\mathbf{2}) * \ldots * (\mathbf{aq} = \boldsymbol{\omega}\mathbf{q})] * (\mathbf{b1}, \boldsymbol{\alpha}\mathbf{1} \to \boldsymbol{\beta}\mathbf{1}) * (\mathbf{b2}, \boldsymbol{\alpha}\mathbf{2} \to \boldsymbol{\beta}\mathbf{2}) * \ldots * (\mathbf{bp}, \boldsymbol{\alpha}\mathbf{p} \to \boldsymbol{\beta}\mathbf{p})] \to [(\mathbf{d}, \mathbf{k1} \to \mathbf{k2})]$, where $\{\mathbf{b1}, \mathbf{b2}, \ldots, \mathbf{bp}\}$ are flexible and $\{\mathbf{a1}, \mathbf{a2}, \ldots, \mathbf{aq}\}$ are stable attributes in S. Further, it is assumed that $\boldsymbol{\omega}\mathbf{i} \in \mathbf{Dom}(\mathbf{ai}), \mathbf{i} = \mathbf{1}, \mathbf{2}, \ldots, \mathbf{q}$ and $\boldsymbol{\alpha}\mathbf{i}, \boldsymbol{\beta}\mathbf{i} \in \mathbf{Dom}(\mathbf{bi}), \mathbf{i} = \mathbf{1}, \mathbf{2}, \ldots, \mathbf{p}$ when $(\mathbf{ai} = \boldsymbol{\omega}\mathbf{i})$ the value of the attribute becomes $\mathbf{ai}$andis equal to $\boldsymbol{\omega}\mathbf{i}$, and $(\mathbf{bj}, \boldsymbol{\alpha}\mathbf{j} \to \boldsymbol{\beta}\mathbf{j})$, and it shows that value of the attribute $\mathbf{bj}$ has been changed from $\boldsymbol{\alpha}\mathbf{j}$ to $\boldsymbol{\beta}\mathbf{j}$. That is to say, object x∈Ssupports an action rule rin S, if there is an object $\mathbf{y} \in \mathbf{S}$ such that: $(\forall \mathbf{i} \leq \mathbf{p})[\,[\mathbf{bi}(\mathbf{x}) = \boldsymbol{\alpha}\mathbf{i}] * [\mathbf{bi}(\mathbf{y}) = \boldsymbol{\beta}\mathbf{i}]], (\forall \mathbf{i} \leq \mathbf{q})\ [\mathbf{ai}(\mathbf{x}) = \mathbf{ai}(\mathbf{y}) = \boldsymbol{\omega}\mathbf{i}]$, $\mathbf{d}(\mathbf{x}) = \mathbf{k1}$ and $\mathbf{d}(\mathbf{y}) = \mathbf{k2}$. Object-driven perspective aforementioned induces a set of structures that are implemented mathematically to evaluate a data set. By implementing objective approach with action rules, some of the chosen objects may be reclassified from one stage to another one by modifying some of the relevant flexible attributes. In [12] we suggested and formulated a simple rule extraction algorithm to build action rules of a single classification rule, which is named Action Rules Discovery Based on Agglomerative Strategy (ARAS). It is a bottom-up approach in a breath-first manner to form all frequent item sets with a qualified part of length $\mathbf{k}$, before forming those qualified part of length $\mathbf{k} + \mathbf{1}$. More information on the application domain of an experiment of ARAS is available in [3,4,11].

IV. RESULTS AND CONCLUSIONS

In the experiment, we useddata of patients suffering from ulcerative colitis (86 cases) and Crohn's disease (66 cases).The analysis was based on the construction of action rules and their application to real medical data. We discretized the data of the selected patients and extracted the highest related attributes in test values. Then we validated the results along with the patients' history and physical examination. The results are very promising. More than 95% of the patients were correctly reclassified.In the future, the full optimization of the object-driven algorithm can be tested rigorously for much lower computational complexity for clinical conformity, especially accuracy of diagnoses, in other areas of the medical domain. In addition, rigorous error analyses need to be studied. These include missing a decision scenario or incorrect logic, along with coding errors, such as bugs in the developed software.

Ethical approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.Informed consent was obtained from all individual participants included in the study.

REFERENCES

[1]. Agrawal R., Srikant R. (1993),Fast algorithm for mining assocation rules, International Conference on Very Large Databases, 487-499.
[2]. Crohn B.B., Ginzburg L., Oppenheimer G.D. 1932, Regional ileitis. A pathologic and clinical entity, J Am Med Ass. 99:1323-1329.
[3]. Dardzinska A. (2013), Action Rules Mining. Springer, pp.90.
[4]. Dardzinska A., Romaniuk A. (2016), Mining of Frequent Action Rules, Machine Intelligence and Big Data in Industry: 6th International Conference on Pattern Recognition and Machine Intelligence, 87-95 .
[5]. Dardzinska A., Kasperczuk A. (2018), Decision-making Process in Colon Disease and Crohn's Disease Treatment, Acta Mechanica et Automatica, Vol. 12 no. 3, pp. 227-231.
[6]. Gürdal O., Dardzinska A. (2017), A New Approach to Clinical Medicine by Action Rules, International Journal of Development Research, 7(1), 11032–11039.
[7]. Han J., Kamber M. (2006), Data Mining: Concepts and Techniques, Morgan Kaufmann Publishers, Second Edition, 21-27.
[8]. Hosmer D., Lemeshow S., Sturdivant R. (2013),Applied Logistic Regression, Wiley Series in Probability and Statistics.
[9]. Jong de P., Heller G. Z., (2008), Generalized Linear Models for Insurance Data, Cambridge University Press, Cambridge.
[10]. Kasperczuk A. and Dardzinska A., (2016), Comparative Evaluation of the Different Data Mining Techniques Used for the Medical Database, Acta Mechanica et Automatica, Vol. 10 no. 3, pp. 233-238.
[11]. Kirsner J. B. 1988, Historical aspects of inflammatory bowel disease, J Clin Gastroenterol ,10:286-297.
[12]. Powers D., (2011), Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness& Correlation, Journal of Machine Learning Technologies. 2 (1): 37–63.
[13]. Ras Z., Dardzinska A. (2011), From Data to Classification Rules and Action,. International Journal of Intelligent Systems, Wiley, 26(6), 572-590

Agnieszka Dardzinska"Modeling Crohn's disease and colon disease diagnosis using rules actionability"IOSR Journal of Engineering (IOSRJEN), vol. 08, no. 12, 2018, pp. 54-56