Applying Clustering based on rules on WHO-DAS II for Knowledge Discovery on functional disabilities.

K. Gibert, D. Statistics and Opertation Research Universitat Politècnica de Catalunya, C. Pau Gargallo 5. Barcelona, SPAIN (karina@eio.upc.es)

R. Annicchiarico, IRCCS Santa Lucia Foundation Via Ardeatina 306 Roma, Italy

J. Rodas, U. Cortés, Software Department Universitat Politècnica de Catalunya

U. Nocentini, C. Caltagirone IRCCS Santa Lucia Foundation U. Tor Vergata, Roma, Italy

Abstract

The senior citizens represent a fast growing proportion of the population in Europe and other developed areas. This increases the proportion of persons with disability and reducing quality of life. The concept of disability itself is not always precise and quantifiable. To improve agreement on the concept of *disability*, the World Health Organization (WHO) developed a clinical test *WHO Disability Assessment Schedule,* (WHO-DASII) that is understood to include physical, mental, and social well-being, as a generic measure of functioning.

From the medical point of view, the purpose of this work is to extract knowledge on the performance of the test WHO-DASII on the basis of a sample of neurological patients from an Italian hospital.

This Knowledge Discovery problem has been faced by using *clustering based on rules*, a technique stablished on 1994 by Gibert which combines some Inductive Learning (from AI) methods with Statistics to extract knowledge on ill-structured domains (that is complex domains where consensus is not achieved, like is the case). So, in this paper, the results of applying this technique to the WHO-DASII results is presented..

Key words:

Disability, scale (clinical test), assessment, neurological disease, Knowledge Discovery, Clustering based on rules, Knowledge-based applications in medicine.

Introduction

The senior citizens represent a fast growing proportion of the population in Europe and other developed areas (). Today (2002), [Madrid, 2002] the number of persons aged 60 years or older is estimated to be 629 million around the world. This quantity is expected to grow to almost 2 billion by 2050, when the population of older persons will be larger than the population of children (0-14 years) for the first time in human history. The largest share of this old people, a 54%, live in Asia; Europe has the next largest share (24%). This ageing population increases the proportion of individuals with physical and/or mental impairment that need some kind of help for the daily tasks.

According to Laselett [Laselett, 1996: ''*the human life span is now divided into four ages: the first is an age of dependency, childhood, and education, the second is an age of independence, maturity, and responsibility, and although the third age is considered a period of fulfillment for physically and mentally fit people in retirement, the fourth age is associated with chronic diseases, disability and dependence''.*

The increasing number of people affected by chronic diseases is a direct consequence of the ageing of the population. Chronic illnesses, such as heart disease, cancer and mental disorders, are fast becoming the world's leading causes of death and disability, including the developing world. In fact, according to the World Health Report 2001, 59%of whole-world deaths relate to non-communicable diseases. In both developed and developing countries, chronic diseases are significant and costly causes of disability and reduced quality of life. The size and pattern of the fourth age is of critical importance not only for the quality of life of elderly people but also because disability is closely associated with the use of health and social services [Fried97].

Measurement of functional ability is increasingly important to the patient's health care and to health care research. This is because functional ability is an important determinant of quality of life and because it highly correlates with both physical and mental health [Wilson95].

Activities of Daily Living (*ADL*) rating scales are widely used in community studies with older people as a measure of functional ability [Katz63]. However, other authors report that *ADL* scales, as well as other disability scales, which typically include bathing, toileting, eating, dressing, and transferring from bed to a chair, have been subject to criticism.

There is, however, lack of consensus on the way of performing this measurement [Ostir01] on the scientific community. Regarding this controversy, the World Health Organization, has recently proposed a new International Classification of Functioning, Disability and Health (ICF) that defines components of functioning and disability, as well as activities and participation [WHO01].

To complement the process, a set of assessment instrument was developed. WHO-DAS II (World Health Organization Disablement Assessment Schedule) [Rehm90] is a scale which enables the assessment of disability levels in consistency with the ICIDH-2 (International Classification of Functioning and Disability [ICDH-2,99]). This instrument incorporates mental health factors related to disability together with physical health factors in the same set of instruments.

In this work, the starting point is the study of a set of neurological patients from an Italian hospital, which were evaluated using WHO-DAS II in order to determine their functional disability degree.

Main goal of the first part of the research is to extract the knowledge contained in the collected data base to see how the WHO-DAS II performs as a scale for disabled patients. Afterwards, relationships between WHO-DAS II and other existing scales will be analyzed as well.So, as a first approach, the work is focused to identify which are the typical answers to the WHO-DAS II and which are the characteristics of the groups of patients who provide this *types* of answers. In fact, this raises a clustering problem, which has to be solved. However, it has been seen [] that classical clustering techniques cannot well recognize the structure of certain sets of data, so producing some non-sense classes, i.e. which cannot be interpreted by the experts.

In fact, this arises when dealing with ill-structured domains [], where either numerical and qualitative information have to be treated together, and there exists some semantic additional (but partial) knowledge to be regarded for clustering.

Clustering based on rules [Gibert98] is a hybrid technique established in [Gibert94] which combines some inductive learning elements with statistical methods to improve clustering results on ill-structured domains. In fact, one of its main advantages is that it guarantees the semantic meaning of the resulting classes. So, on the basis that medical applications usually conform ill-structured domains, this study is faced using this technique.

Actually, the contents of the paper is the following: Introduction to the WHO-DAS II scale, description of the considered sample and the characteristics of the study, description of the *clustering based on rules* technique, application of it to the target sample, results, conclusions and future work.

The WHO-DAS II scale

Regarding the controversy about measurement of disability, the World Health Organization, has recently proposed a new International Classification of Functioning, Disability and Health (ICF) that defines components of functioning and disability, as well as activities and participation [WHO01]. Also, it has designed a scale for measuring Disability (WHO-DAS II) [Rehm90].

The newest version of the disability classification is ICF, which in fact is a review of ICIDH-2, moving away from a "*consequences of disease*" classification (1980 version) to a "*components of health*" classification. This latest model is designed to provide a common framework and language for the description of health domains and health-related domains. As a classification, ICF systematically groups different domains for a person in a given health condition (e.g. what a person with a disease or disorder does or can do). Functioning is an umbrella term encompassing all body functions, activities and participation; similarly, disability serves as an umbrella term for impairments, activity limitations or participation restrictions. ICF also lists environmental factors that interact with all these constructs. In this way, it enables the user to record useful profiles of individuals' functioning, disability and health in various domains.

WHO-DAS II is a scale that enables the assessment of disability levels in consistency with the ICIDH-2 classification system (and, as a consequence, with ICF). Besides, this instrument incorporates mental health factors related to disablement together with physical health factors in the same set of instruments.

This study used the WHO-DAS II version 3.1a (June 1999), which is a fully-structured, 36-item interview measuring self-reported difficulty of functioning in: *Understanding & Communicating* (6 items), *Getting Around* (5 items), *Self Care* (4 items), *Getting Along with People* (5 items), *Life Activities* (8 items) and *Participation in Society* (8 items). These are six major domains that encompass activities that are considered important in most cultures.

The WHO-DAS II employs a 5 point rating scale for all items on which 1 represents no difficulty and 5 represent extreme difficulty or the inability to perform the activity. Six WHO-DASII domain scores may be obtained by summing the answers in each domain, normalizing them on a 0 to 100 scale, and expressing these values as percentages such that higher values represent greater disability. Information related to the extent of disruption in life caused by these difficulties and extends of difficulties and extends of difficulty experienced in life and extends of dependence on assistive devices or other persons is considered as well.

The validation of WHO-DAS II is in progress in international field trials comprised of 16 centers in 14 countries from which 1564 men and women were drawn from the general population, and from persons with physical, mental, drug, and alcohol problems.

Experimental procedure

The sample of this study was comprised of 96 neurological patients between 17 and 80 years, who were recovering at the IRCCS hospital between October 1999 and February 2000. A control group of 20 healthy people, comparable for gender and mean age have been enrolled.

In the sample, 58 patients were males (60.4%) and 38 females (39.6%). The average age was 56 years. There were 20 patients with spinal cord injury (age 47.20, s= 17.6), 20 with Parkinson (age 69.25, s=6.53), 20 with stroke (age 63.40, s=15.96), 16 with depression (46.56,s=11.15) and 20 control individuals (55.05,s=15.57).

All patients were assessed with WHO-DAS instruments at admission and their functional status and health-related quality of life was measured. Each patient was evaluated according to two other standardized clinical scales. Functional status was quantified using the Functional Independence Measure (FIM). The FIM is a well-established measure for which reliability and validity have been proved [Dodds93]. Patients' Quality of Life (QOL) was quantified using the Short Form (SF)-36. The reliability and validity of this tool has been documented extensively [Were93].

Analysis: Clustering based on rules

As said before, the study is firstly focused on identifying different kinds of responses to the WHO-DAS II as well as to characterize the groups of respondents of any kind.

The classical way to face this problem is using a clustering algorithm for detecting groups of patients on the basis of the WHO-DAS II responses. It has been justified above the interest of using *clustering based on rules* (cis?) instead of classical clustering techniques. In this section a brief description of the whole proposed analysis methodology is presented:

1. *Descriptive statistics of every variable*: Use classical descriptive techniques to know the nature of the data.

Very simple statistical techniques may be used to describe the data and to get some preliminary information about it; histograms may be used to make graphical representations of the variability of measures; time series plots may be used to observe evolution along time; plots and multiple box-plots may be used to observe the relationship between some pairs of variables [Tukey77].

2. *Data cleaning*: includes missing data treatment or outlier detection. It is a very important phase, since the quality of final results directly depends on it. Decisions are taken on the bases of descriptive statistics and background knowledge of the experts.

3. *Cluster-classify data* (depending on the situation): The proposed method for that is *clustering based on rules*. The original development of this method is [Gibert98] and it was presented with details in [Gibert94]. It consists of a mixture of an AI process and a clustering one and it has been successfully used in several real applications previously [Gibert99, Gibert01, Comas01]. Our experience from previous applications is that using this *clustering based on rules* use to be better than using any statistical clustering method by itself, since clustering *based on rules* is able to produce *meaningful* classes. Also, it uses to be better than pure inductive learning methods since it reduces the effects of missing some implicit knowledge in the Knowledge Base.

The main goal of this paper is not going into specific details, but to give an intuitive idea of the method performance. For a detailed presentation see [Gibert98].
The general idea of *clustering based on rules* is:

- To use a (partial) Knowledge Base provided by the expert to induce an initial partition on part of the data; put the data not included in this partition into the *residual class.*
- To generate prototypes of each *rules-induced class*.
- To build the *extended residual class* as the set of prototypes of *rule-induced classes* together with the residual class data.
- To Uuse a hierarchical clustering method to weighted cluster the extended residual class. The result of *clustering based on rules* is a dendrogramme which graphically represents the clustering process in form of a binary tree where the height of each node indicates the distance (in a wide sense) between their children; the leafs of each node are the set of elements included in the class represented for this node.

2. To determine the final classes. This requires no new developments and consists in analyzing the

dendograme to choose the best horizontal cut (where the largest gap exists). This identifies a partition of the data in a set of classes.

3. To built a conceptual description of the resulting partition. This step requires automatic classes interpretation and it is mainly based on the use of some statistical tools to extract qualitative information from the classes. The aim is to obtain a meaningful description to the user and which indicates particularities of every class with respect the others.

In the next section the application of this methodology to the WHO-DAS II responses is described.

Results

First of all, descriptive analysis and data cleaning was performed. Basic summaries were already presented in section Experimental procedure. In data cleaning, variables with *don´t care* values have been properly treated. Once data prepared for analysis, a hierarchical clustering was performed. Reciprocal neighbors algorithm was used. Since both numerical and categorical variables have to be considered, a mixed metrics proposed by Gibert in [1] was used. A α–cut with 4 classes was done. Several tools were used to assist the interpretation of the final classes. From none of them it is clear why individuals with depression are distributed in several classes together with other diseases. Looking for a more accurate classification, α–cuts into greater number of classes were explored, but they do not improve the quality of the results. The distribution of the variables along the classes, do not help to explain why classes were formed. Neither physician could understand why the partition was formed, nor the meaning of the final classes. Table 1 compares the composition of the classes according to the diagnostic of the patients and it can be seen that depressed patients are distributed along almost all the classes, including the class of control patient, which in fact have no disease.

Clustering based on rules In fact patients with disabilities can be considered as an *ill-structured domain* (ISD) [Gibert94] and, as said before, clustering uses to be unable to capture the structure of those kinds of data by itself. Thus, results of previous section are not so strange. *Clustering-based on rules* [Gibert98] is a combination of inductive learning and statistics that can overcome these limitations. It is implemented in the software KLASS+ [Giber94] that is used for this research. It has been successfully used in other medical domains [Gibert99, Comas01, Rodas01].

The additional knowledge the expert could provide is expressed by means of logical rules, which represent a partial description of the domain (as usual for ISD, it is very difficult to make explicit a complete Knowledge Base for the domain, and this is a great handicap for using pure Artificial Intelligence methods). Since patients with emotive problems are not well grouped in the previous classification, the rules used in this application concern the items of the WHO-DAS II, which ask for emotive *behavior*.

There are 4 questions in the test related to emotive behavior :

- B4= *How do you rate your mental or emotional health in the past 30 days?*
- B9= *How much worry or distress have you had about your health in the past 30 days?*
- S5= *How much have you been emotionally affected by your health condition?*
- R2= *How much have the difficulties been caused by mental health or emotional problems?*

It is considered that people who answer higher values to those questions should have any kind of emotive problem. So, the proposed Knowledge base for biasing cluster is:

KB={ r1: If B 4 is in [4,5] then emotive-problems,
r2: If B9 is in [4,5] then emotive-problems,
r3: If S5 is in [4,5] then emotive-problems,
r4: If R2 is in [4, 5] then emotive-problems}

Clustering based on rules was used with this Knowledge Base (KB) on the sample of 170 patients. Rules divided the sample and clustering was done in the rules-induced part (which contains 56 patients) as well as on the extended residual class, building a global hierarchy. Finally, a set of 7 classes was recommended by the system. Three of them contain isolated patients that have outlier behavior; they will be studied individually. The description of the other four classes (Cd52, Cd53, Cr93, Cr89) can be partially seen in figure 1. Indeed, in this case, a clearer conceptual interpretation of the classes is available. In fact, there is a group of variables whose distribution clearly allows distinction among classes; more remarkable differences are provided by answers to questions:

- B2, B9, B10, S12, P11, which allows distinction between classes with emotive problems (cd52, cd53) and classes without (cr93, cr89);
- S3, S4, S6 which allows distinctions between cr52 and cr53;
- S5, S7, S8, S9, P21, P31, P62, R1, R2, R6, which allows distinction between cr93 and cr89;
- S1, P41, P51 that allows simultaneously distinction between cd52 and cd53 on the one side and between cr93 and cr89 on the other side.

So, for example S1= "*how much difficulties on standing up 30 minutes had you during last 30 days" is* very low for classes cd52 and cr93 and really high for classes cd53 and cr89. Also, B2= *"How do you value your physical health during last 30 days?"* takes low values (around 2) for cr93 and cr89 and high values for cd52 (4) and cd53 (5).

Looking at different variables, a description of the classes in terms of those features, which makes differences among them, can be done:

- Cd52: (B2) Physic health bad, (B9) worry or distress moderate, (S1) can stand up 30 minutes, (S12) mild difficulties in the daily work, they manifest extreme physical problems (R1) and moderate mental ones (R2).
- Cd53: (B2) Very bad physic health, (B9) moderate worry or distress, (S1) extreme difficulty for standing up 30 minutes, (S12) high difficulties in the daily work, the only class with high difficulty on learning a new compute (S3), participating at community activities (S4) and concentrating for 10 minutes (S6), and difficulties on (P11) communications and comprehension highly affects their life, they also manifest extreme physical problems (R1) and mild mental ones (R2).
- Cr93: (B2) Physic health good, (B9) non worry or distress, (S1) can stand up 30 minutes, (S8) no difficulties on toileting the whole body or (S9) dressing, the only class with non emotional involving on life conditions (S5), or being able of walking for ten minutes (S7), difficulties on transfer (P21) and selfcare (P31) highly affects their life, in fact, the only class with no problems at all during last 30 days (P62). They manifest no physical (R1) nor mental problems (R2).
- Cr89: (B2) Physic health good or excellent, (B9) mild or moderate worry or distress, (S1) extreme difficulty for standing up 30 minutes, (S8) moderate difficulties of toileting the whole body, (S9) and dressing. They manifest high physical problems (R1) and no mental problems at all (R2).

From the medical point of view, the composition of the classes is well corresponding with different types of functional disability. Moreover, it can be seen that Depressed patients (which are supposed to have emotive problems), are captured all of them by the rules, but divided into two main subgroups by the clustering. Such a subdivision corresponds to some interesting criteria: Cd53 have greater physical and mental problems (learning new computes, participating in community, concentrating, working) compared with Cd52, and also patients in Cd53 fell that those difficulties (including those of toileting and dressing) affects much more to their life; on the contrary, Cd52 cannot stand up, while Cd53 can do.

Classes cd52 and cr93 have no difficulties on interacting with other people and daily activities, while classes cd53 have high difficulties on these topics and cr89 moderate.

Finally, class Cr89 is formed of people without cognitive problems but with some moderate degree of physical problems (cannot stand up, …).

On the basis of this information, it can be stated that 4 classes of functional disabilities emerge from this study:

- Cr93: is the class of healthy patients, including all the people without any diagnosis;
- Cd53: patients with the worst performance, in terms of physical, cognitive and emotional function;
- Cd52: patients with ligther physical problems respect to the latter group, but with the same dregree of cognitive and emotional function…..
- Cr89 patients with moderate physical problems.

Conclusions and Future Work

In this application, again [Gibert98b], the use of *clustering based on rules* produces meaningful classes and improves from a semantics point of view the results of classical clustering.

The rules proposed by the experts selected all the depressed patients (which are supposed to have emotional problems). The final hierarchy suggests a subdivision of those patients into two classes one for those with greater problems on the cognitive aspects and the other for those with problems mainly on the physics.

It has been identified 4 groups of functional disabilities which are, in fact, related to higher degree of disabilities and it has been also seen which are the more characteristic items of every class. Classes, in fact could be ordered according to the gravity of the functional disability as: Cr93 (no problems), Cr89 moderate disabilities in physics, Cd52, moderate mental and/or cognitive disability and high physical disability, Cd53 extreme physical and cognitive and/or mental disability.

There is no a group with high cognitive and no physical disability, probably depending on the apraxia (impossibility of performing corordinate and finalised tasks) strongly related to severe degrees of cognitive impairment.

There may be some kind of correlation between the gravity of the disability and the depression. This is to be studied.

Relationship with DIM and SF-36 is actually in progress.

Response 5 puts together people with great difficulties in doing something and people absolutely disabled for that. This introduces some kind of ambiguity in the data that may disturb some results. In the future it should be convenient to consider maintaining this ambiguity.

Outliers have to be studied.

Acknowledgements

Thanks to the EURO-MDSBL e il WHO Assessment Classification and Epidemiology Group who designed the WHO-DAS II. Also, to project TIC-2000. Authors also want to thank Fabio Campana for his help.

References

[Comas01]Comas J., Dzeroski S. Gibert, K., Rodríguez-Roda I. and Sànchez-Marrè M. 2001. Knowledge Discovery by means of inductive methods in WWTP data. *AI Communications* 14(1):45-62, January 2001.

[Fried97] Fried, L. P., Guralnik, J. M. 1997a. Disability in older adults: evidence regarding significance, etiology, and risk. *J Am Geriatr Soc* 45: 92-100.

[Wilsson95] Wilson, I. B., Cleary, P. D. 1995a. Linking clinical variables with health-related quality of life. A conceptual model of patients outcomes. *JAMA.*

[Gibert98] Gibert, K., Cortés, U., `Clustering based on rules and Knowledge Discovery in ill-structured domains', *Computación y sistemas, revista iberoamericana de computación* **1**(4), 213227 (1998).

[Gibert98b]Gibert K., T. Aluja, U. Cortés, 1998 Knowledge Discovery with Clustering Based on Rules. Interpreting Results, LNAI 510, p 83-92, Springer-Verlag, 1998.

[Gibert97] Gibert, K., Cortés, U., 1997. Wheighing quantitative and qualitative variables in clustering methods', *Journal Mathware and Soft Computing* **4** (7), 251--266.

[Gibert94] Gibert K., Cortés U. 1994. Combining a knowledge based system with a clustering method for an inductive construction of models. LNS no 89, pp 351--360, New York, Springer-Verlag, 1994.

[Gibert99] K. Gibert and Z. Sonicki , 'Classification based on rules and medical research' , *Journal of Applied Stochastich Models and Data Analysis* , **15** (3), 319--324 (1999).

[Gibert01] K. Gibert, Z. Sonicki, J. C. Martín, 'Impact of Data encoding and Thyroids dysfinctions' Health Data in the Information Society, Technology and Informatics v 90 494-503, 2001.

[ICDH-299b]ICIDH-2:International Classification of Functioning and Disability. 1999 b. Geneva, World Health Organization

[Katz 63] Katz, S., Ford, A. B., et al. 1963a. Studies of illness in the aged. The index of ADL: a standardized measure of biological and psychosocial function. *JAMA* 185: 914-919.

[Laslett96] Laslett, P. 1996 A fresh map of life: the emergence of the third age. London: Macmillan Press.

[Ostir01] Ostir, G. V., Volpato, S., et al. 2001a. Summarizing amount of difficlulty in ADLs: A refined characterization of disability. *Aging Clin Exp. Res* 13: 465-472.

[Rehm90a] Rehm, J., Ustun, T. B., et al. 1990a. On the development and psychometric testing of the WHO screening instrument to assess disablement in the general population. *Inter J Meth in Psychiatric Research* 8(2): 110-122.

[Rodas01] J. Rodas, K. Gibert, and J. E. Rojo 'KDSM Methodology for Knowledge Discovery from Ill-Structured Domains Presenting Very Short and Repeated Serial Measures with Blocking Factor' LNAI 2504, p. 228-238 ff, 2001.

[Tukey77] Tukey, J.W Exploratory Data Analysis, Addison-Wesley, 1977

[WHO02] Second World Assembly on Ageing Madrid. 2002b. Building a Society for all Ages. Spain 8 -12 April.

[WHO01b] World Health Organization. 2001b. The International Classification of Functioning, Disability and Health-ICF. Geneva, Switzerland: WHO.