БЪЛГАРСКА АКАДЕМИЯ НА НАУКИТЕ . BULGARIAN ACADEMY OF SCIENCES

ПРОБЛЕМИ НА ТЕХНИЧЕСКАТА КИБЕРНЕТИКА И РОБОТИКАТА, 50
PROBLEMS OF ENGINEERING CYBERNETICS AND ROBOTICS, 50

София . 2000 . Sofia

Towards a Multifunctional Platform for EMG Studies[1]

Gennady Agre, A. Neltchinov

Institute of Information Technologies, 1113 Sofia, email: agre@iinf.bas.bg

1. Background

European Research Network for Intelligent Support of Electromyographic Studies (EMG-Net) is a research INCO-COPERNICUS project aims at establishing a research network combining information technology (IT) and medical teams from nine Western and Central/Eastern European countries (http://www.iinf.bas.bg/EMG-Net). Electromyography (EMG) is a set of electrophysiological techniques allowing defining neuromuscular dysfunction disease diagnosis. The EMG domain covers more than hundred existing diagnoses and exploits data from about four thousand tests of nervous or muscular structures. EMG-Net is a successor of ESTEEM AIM project that resulted in introducing a set of standards for EMG study and in developing some knowledge based systems for EMG diagnoses. The main objective of the on-going EMG-Net project is to complete and enhance the available EMG knowledge collected during ESTEEM project. This new knowledge will consist mainly in useful dependencies between parameters of EMG cases and can be acquired by using such advanced IT techniques as machine learning (ML) and data mining (DM). Furthermore, the expertise of other EMG labs, external to the previous consortium will be added to the current knowledge base augmenting and covering in such a way a greater part of the variety of the EMG practice.

The technical goal of EMG-net is to develop a multifunctional platform for EMG study aiming at assisting EMG practitioners in developing EMG standard examination procedures as well as in analysing and evaluating of EMG methods. The paper describes architecture of such a platform with emphasis on one of its modules – the data mining system DaMiS that is under development in the Institute of Information Technologies – BAS, which is one of the EMG-Net IT partner. The structure of the papers is as follows: the next section introduces taxonomy of some open problems identified in the EMG domain that could be solved by applying ML and DM tech-

[1] The research was partially supported by EMG-Net INCO-COPERNICUS Project 977069 funded by the European Union and by the National Science Foundation, Grant I-606/96.

niques. The section 3 presents the architecture of the multifunctional platform for EMG studies and relates the developed taxonomy to some of its modules. Section 4 gives a closer look onto the data mining module of the platform, and the last section is an overview of the current state of the system development and directions for future work.

2. The taxonomy of EMG open problems

2.1. The current state of EMG interpretation Process

The Fig. 1 presents a model of the EMG diagnostic process developed during ESTEEM project [7]. The interpretation begins with selecting of one or more *tests* that are of relevance to the initial goal of the investigation.

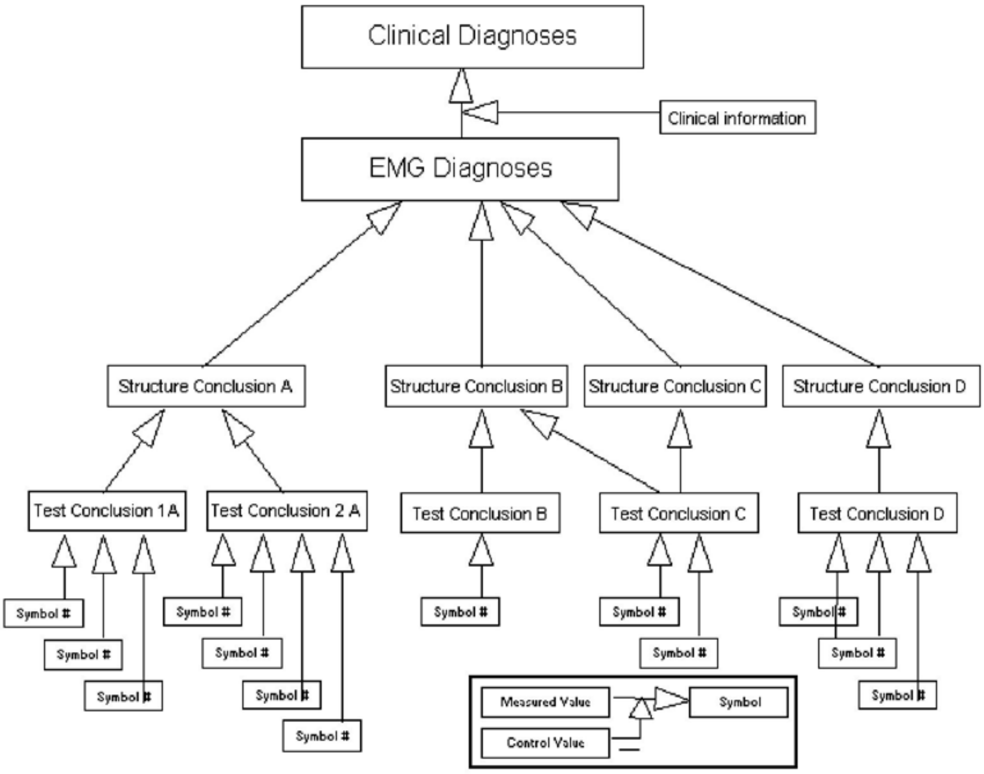

Fig. 1. Model of EMG diagnostic process

Following the performance of the tests, the individual numerical parameters derived from each test are compared with an expected normal range for the given parameter. This process results to a symbolic value associated with each parameter. The symbols derived from each parameter originating from a single test are then combined to give a *test conclusion* that may be considered as a multi-dimensional representation of the pathological process that could give rise to the observed findings. The next step is a reinterpretation of each test conclusion in the light of test conclusions from other tests to derive a *structure conclusion* represented as a *pathophysiological state* too. Also at this level, the test conclusions from a number of tests on the same structure may be combined to give a single structure conclusion. Then all of the structure conclusions are interpreted to derive an *EMG diagnosis*, which is represented in the form of one or more focal or diffuse pathological processes which may involve the peripheral nervous system. And finally, the *clinical diagnoses* is derived from the EMG diagnoses together with all clinical information (history, clinical examination, clinical chemistry, etc.

2.2. The open problems

As a result of intensive discussion between medical and IT partners of the EMG-Net project the following open problems in the EMG interpretation proccess have been identified:

At the numeric level of EMG reasoning:

1. *Inferring the test conclusions from numeric rather then symbolic data*. The main reasons for rising this problem are:

- There is no concensus on calculating the normal (reference) values played the basic role in the process of inferreing the test conclusions.
- Reference values for a specific parameter are often dependent on age, sex, height and specific examination conditions. The current process of calcuationg such values does not take into account these dependences.

2. *Verification of the consistency of test conclusions.* In the moment it is not clear when and if the obtained data is enough for inferring the reliable test conclusions. That is why it will be useful to try to find a minimal subset of data, which is necessary for inferring the consistent test conclusions for a given examination technique.

At the symbolic level of EMG reasoning:

3. *Inferring consensus diagnostic rules with a focus on Polyneuropathy.* It has been reported that in the current moment there is no consensus on the rules for inferring the Polyneuropathy. It has been proposed to try to infer such rules from examples from the consensus case base starting not from test or structure conclusions but from a more abstract level – so called "disassociation patterns". Inferring such patterns can be done by means of the existing EMG expert system DARE (Cruz, Barahona, 1997) developed by the Portugal EMG-Net partner – University of Lisboa.

4. *Refinement of existing diagnostic rules.* The existing diagnostic rules are heuristic and neither absolutely correct nor complete. It will be useful to try to refine such rules by going down to the low (data) level as the experts often do it. As a base for such experiments the existing diagnostic rules used in the EMG expert system KANDID [15] or DARE can be explored.

5. Check the consistency of diagnosis conclusions. In the moment it is not clear when and if the inferred diagnostic solutions are reliable given a specified set of test conclusions. Finding the solution of such a problem will allow defining the conditions when it is necessary to reinterpret the raw data or to take into account some additional information in order to improve the reliability of the current diagnostic conclusion.

At the planning level of EMG reasoning:

One of the "hot" topic discussed between EMG-Net medical partners is what is missing in the current database. It has been recognised that the information about *test utility* would be very useful for improving the quality of the EMG dianostic process. That is why the problem of *determining the test utility based on the existing information about the sequence of tests execution* has been defined as an open problem to be tried to solve by DM or ML methods.

3. The platform architecture

The main objective for developing the multifunctional platform for EMG studies is to provide the EMG physicians with radically new possibilities for analysing EMG cases and for discussing examination procedures.

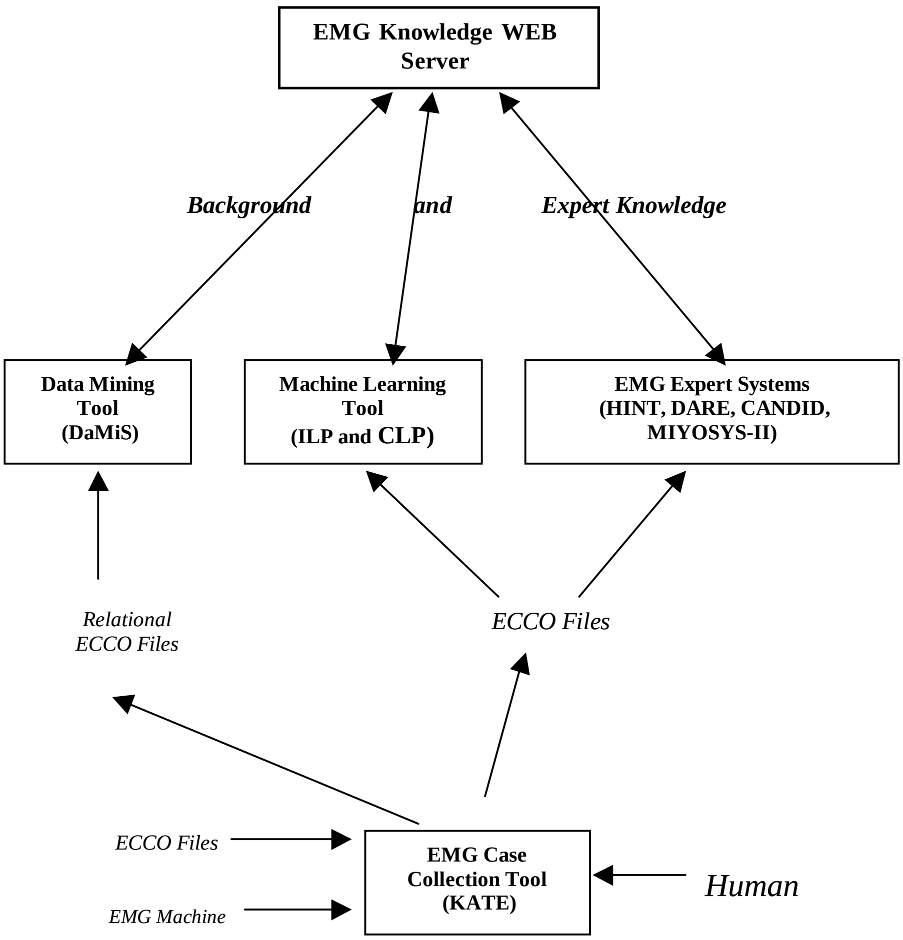

Fig. 2. The multifunctional platform architecture

Such a platform may be considered as a first basic step toward a development of a virtual clinical electromyographic laboratory fully integrating all sources of EMG knowledge and data with advanced IT tools for processing them. The proposed architecture for such a platform is shown on Fig. 2. It consists of four main components – data collection, problem solving, learning and educational components.

3.1. The data collection component

The role of the data collection component of the platform is played by the KATE system – an advanced tool for collecting and managing EMG cases, which is under development by the Slovak University of Technology, Bratislava – the Slovak IT Partner of EMG-Net project. It bases on the existing Case Collection Tool (CASE) [7] developed during ESTEEM project. KATE can collect EMG cases at several levels: local, national and European level. The local level is represented by a particular EMG workplace. The national level becomes important when normative data collection and evaluation is considered. The European level is intended mainly for consensus exercises aiming at the standardisation of the EMG studies. The tool is designed in such a way that a patient is considered as a central system entity. A physician can work with the patient data (general and clinical information) and create several EMG examinations for the patient as well as to work with several EMG cases (for the same patient or for different patients) at the same time. KATE supports three formats:

- ECCO binary format (ECCO Files), which has been designed during ESTEEM project as an efficient format from the storage point of view. It is retained in the new case tool in the form of export/import capabilities. Such a format is used by a set of existing EMG expert systems (see Fig. 2). Moreover, ECCO files will be used (after translating them to Prolog files) as input to ML tools based on Inductive and Constraint Logic Programming, which are under development by the Research Institute for Informatics, Bucharest – the Romanian IT partner of EMG-Net project.
- The relational format (relational ECCO Files) is used for storing EMG cases (i.e. for creating relational data bases with EMG cases) and for exploring by data mining tools.
- Finally, the XML format is suitable for exchanging EMG cases between medical EMG-Net partners. Moreover, the direct import of signals from EMG machine would be supported too.

3.2. The problem solving component

The problem-solving component of the proposed multifunctional platform is presented by a set of knowledge-based systems developed within ESTEEM project. These systems mainly belong to two categories:

- The systems from the first category (HINT [14] and DARE) start with a set of test results (usually provided by the CASE or KATE tool) and infer an EMG diagnosis off-line (they do not do test planning). Such systems perform reasonably well on the "clear-cut" cases, especially when clinical information is not crucial in reaching a final diagnosis.
- The systems from the second category (KANDID and MYOSYS-II [18]) do test planning as well. This sort of facility is extremely useful in practice, since the diagnosis task cannot really be separated from test planning.

3.3. The learning component

The learning component of the platform consists of data mining and machine learning tools. The objective of the DM tools is to "mine" new EMG knowledge from the data acquired by the data collection component. The role of a data mining tool is played by the DaMiS system, which is under development in the Institute of Information Technologies – BAS. This system is intended to be a main tool for solving most of the mentioned above open EMG problems (at least 1, 2 and 3) and will be described in more details in section 4.

The main role of ML tools is to refine the existing expert rules for EMG diagnosis (see Open problem 4). Starting from an existing diagnostic system (such as KANDID), an anatomical database, a set of lesion localization rules as well as a (partial) set of diagnosis rules, containing in EMG knowledge Web server, this ML component will try to refine these rules to a (more) complete and correct set by learning in cases in which these rules are incomplete or incorrect.

Such theory refinement problem has been addressed in (Richards, Mooney, 1995), but these results need to be adapted in the context of EMG expert systems. The learning system, which is under development by the Research Institute for Informatics, Bucurest, will be based on an extension of ILP by a flexible, complete and non-redundant refinement operator, which is absolutely necessary for solving this EMG problem.

3.4. The educational component

The educational component of the platform is represented by the EMG knowledge Web server, which will contain all available consensual knowledge on EMG domain. EMG knowledge concepts are split into three types: initial concepts (symptoms, sign, risk factors, etc.), concepts representing deduced information (diseases, test procedures, test results), and secondary concepts used to infer the previous ones (anatomical concepts and topography, heuristic lesion localization rules, etc.). The server, practically, reflects the current state of expertise in EMG domain and provides a possibility to access to this expertise via World Wide Web (Fig. 3).

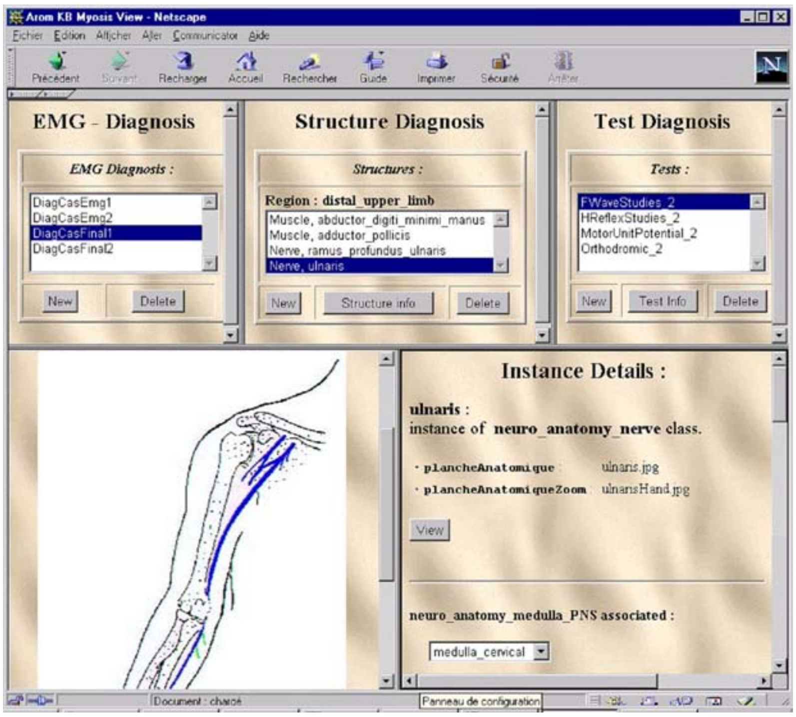

Fig. 3. An example of the EMG knowledge Web server screen

The server is implemented on AROM – an Object Based Knowledge Representation system written in Java and developed by the members of the ROMANS project at INRIA Rhone-Alpes – the French IT partner of EMG-Net project [11]. It provides both an original knowledge representation model based on classes and associations and on an Application Programming Interface, which allows the user to write his own applications using the proposed knowledge representation. The user interface allows domain experts to design knowledge bases without learning the AROM modelling language.

4. DaMiS functional architecture

DaMiS system is an important part of the multifunctional platform for EMG studies. It designed to be an assistant of the data analyst during the entire DM process (Fig. 4). After understanding the domain objectives the analyst can use the system for data understanding, preparation and modeling as well as for evaluation of obtained results. In the next subsections the functional architecture of DaMiS is presented from the DM process model point of view [Chapman et al 99], i.e. the description of the system functions is given with respect to the main DM tasks these functions are intended to solve.

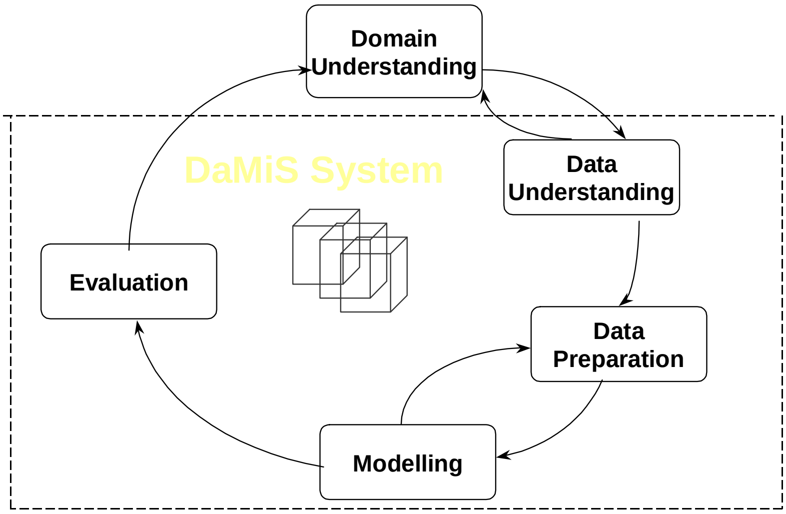

Fig. 4. The scope of DaMiS Functionality

4.1. Data understanding functions

The data understanding activities can be seen as a sequences of four main tasks – collection of initial set of data, describing this set, exploration of collected data and verification of data quality.

4.1.1. Functions for Initial Data Collection

Data collection functions aim at creating an initial database, which can be further manipulated and investigated by different data mining algorithms. The system is able to create such an initial data collection either via import from external sources (relational databases or text files) and from its own files previously processed and stored by the system itself. Relational databases are the main source of data for DaMiS system. To facilitate description and exploration of the data DaMiS provides a set of facilities for manual and query-based automatic selection of data tables, records and attributes.

DaMiS is able to collect data from text files stored in several well-known repositories for machine learning databases that are used mostly for comparing the behaviour of different data mining algorithms. Most of these files have a similar format (the differences are mainly in the separators used, comment marks, etc.). All such files may be used as an input to the DaMiS system with a few pre-processing activities.

4.1.2. Functions for Data Description

Data description functions provide the analyst with an initial understanding of collected data mainly from statistical point of view. The functions report about such characteristics of data set as number of rows (examples) and columns (attributes), type and domain of attributes, value distribution of nominal attributes or values of mean, median, mode and standard deviation of a continuous attribute, etc. DaMiS allows evaluating attributes importance in accordance with a specified class attribute by means of attribute correlation matrix or by applying the attribute weighting technique ReliefF [10]. Another variant of this algorithm may be used for attribute selection.

4.2. Data preparation functions

Data preparation functions allow selecting and cleaning the data as well as constructing new data attributes and preparing (formatting) data for specific data modelling methods.

4.2.1. Functions for Data Selection

Data selection functions aim at constructing *modelling data sets* (so called modelling data mining table - *MDMT)* intended for using by different data modelling algorithms. Each MDMT represents a subset of data selected according to specified criteria. For example, it is possible to select examples manually, or based on a specified value of a class attribute, or to construct MDMT as a stratified sample based on a specified percentage of examples from the whole size of initial data table, or based on a specified SQL query etc.

4.2.2. Functions for Data Cleaning

DaMiS provides functions for processing the missing values and outliers. A way of processing missing values depends on a type of the corresponding attribute – nominal or continuous. A missing nominal value is most often replaced by "?" ("unknown") symbol or simply discarded. The treatment of missing values of a continuous attribute is more difficult since the assignment of a numeric value to missing attribute values will change the distribution and statistic of this attribute. In DaMiS it is possible to segment data using distribution of a class attribute and assign the average or modal value of each segment to missing values from the corresponding segment of a specified continuous attribute. (Segmentation may be done using arbitrary attribute using "Set Class Attribute" function setting an attribute as the target one).

An outlier, or outlying value, is an example which value for a particular attribute is outside the attribute's normal range, as defined by (most often) ~99 % of the other possible values. If the outlying value is extreme, it could seriously alter the accuracy of a model built on the data. Sometimes outlying values are useful and should not be removed (e.g. in the task of deviation detection). In DaMiS you can mark or remove examples (outliers) that are outside of the range *defined by the user* for a specified or all attributes.

4.2.3. Functions for Data Transformation

Data transformation functions aim at changing a way of processing (interpretation) of an attribute as well as at constructing new attributes. These functions may be generally split to data normalization functions, data discretization functions and attribute construction functions.

DaMiS normalization functions are able to normalize continuous attributes into the range [0,1] or to apply Z-normalization method.

In DaMiS, the user can manually change the attribute type, specify an arbitrary attribute as a target (class) attribute or discretize an attribute applying his/her own intervals. More advances functions allow unsupervised discretization of continuous attributes on a specified number of equal length or equal frequency intervals [Weiss & Kulikowski 90]. It is also possible to discretize attributes by Recursive Minimal Entropy Partitioning method [6] based on the minimal entropy heuristic.

Construction of new attributes is done by means of SQL queries to MDMT tables.

4.3. Data visualization functions

Data visualization functions are intended to visualize the results of applying functions for data collection, selection and preparation. DaMiS system is going to be equipped with several modules (called visualizers) that are able to visualize such data from different points of view. In the moment the 2D parallel attribute axes (2D PAA) visualizer has already been implemented [1] (Fig. 5).

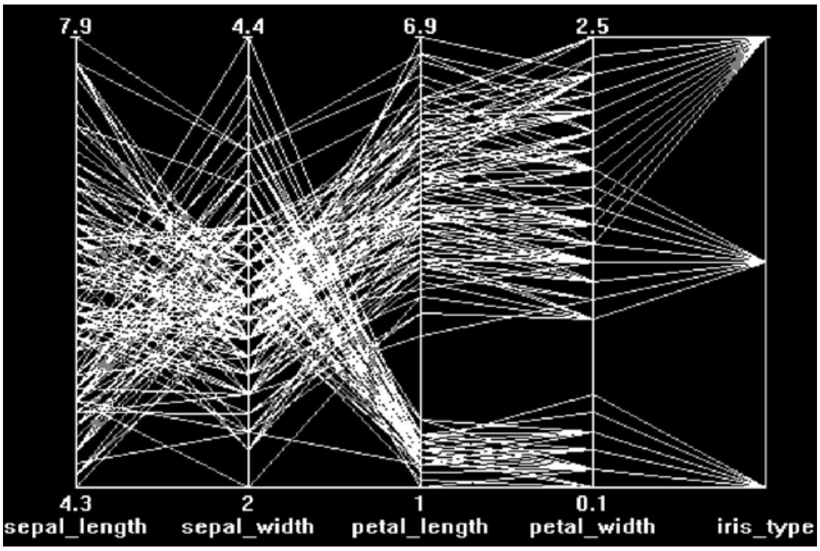

Fig. 5. 2D Parallel Attribute Axes Visualization of data

Each visualizer uses a specific visualization technique, however all of them are able to run a unified set of visualization functions, which does not depend on the technique used. All such functions operate either on *an active MDMT (A-MDMT)* or on its own graphical data structure corresponding to such A-MDMT (or on both). The developed set of the visualization functions is reach enough and assists the user in the sophisticated process of data analysis.

4.4. Data modelling functions

4.4.1. Functions for Knowledge Generation

Data modelling is the main phase of the DM process allowing the analyst to achieve the DM objectives by selecting and applying the corresponding modelling techniques to each particular type of DM problems. In principle, DaMiS should be able to solve all these problem types, however, the prototype of the system contains a restricted set of data modelling techniques designed as data modelling functions. Each data modelling function uses A-MDMT as learning data set and produces a new data structure called Knowledge Table (KT) that is specific for the model used. In the moment the following modelling functions are fully or partially implemented:

• *K-Nearest Neighbours Classification*. A-MDMT is used as the data model, which can be extended by weights of examples and attributes. As parameters of the model the number of nearest neighbours, the distance metrics and the way of weighting the neighbours are used.

• *Symbolic Nearest Mean Classification*. The data model is represented by a set of artificially constructed examples – prototypes, the number of which is determined automatically for each class. The prototypes are built by means of a variant of SNMC algorithm (Datta, Kibler, 1997). As a parameter of the model the distance metrics is used.

• *Typicality-Based Classification*. The data model is represented by a subset of existing examples selected from the learning set on the base of their typicality. The model is built by means of a variant of TIBL algorithm [17]. As a parameter of the model the similarity metrics is used.

• *Bayesian Classification*. The data model is represented by a set of numbers, which are estimations of conditional probability of all attribute values given a class and prior probabilities of the classes. The model is built by means of a variant of naïve Bayesian classifier [4].

• *Integrated Inductive and Instance Based Classification*. The data model is represented by a set of covering rules built by means of a version of RISE algorithm [5]. As a parameter of the model the similarity metrics is used.

• *Decision Tree Based Classification*. The data model is represented by a set of discriminating rules built by means of a variant of the C4.5 algorithm [12].

• *K-Nearest Neighbours Regression*. A-MDMT is used as the data model, which can be extended by weights of examples and attributes. As parameters of the model the number of nearest neighbours, the distance metrics and the way of weighting the neighbours are used.

• *K-Means Clustering*. The data model is represented by a set of all learning examples associated with a label (cluster name). As parameters for this partitioning clustering algorithm the number of clusters (K) to be built and the distance metrics are used.

4.4.2. Functions for Visualization of Data Models

As usual in DaMiS, all data models can be illustrated and analysed by different visualizers. The set of the models able to be built (or used) in DaMiS consists of three main types:

• *Exemplar Based Models* (K-NN classification and regression, SNMC, TIBL, K-MEANS) – the main model element is *an exemplar*. The exemplar has the same representation as a single datum – an example or a single row in a MDMT and may have some additional properties such as example or attribute weights.

• *Rule Based Models* (RISE, DT and RT based rules, Association Rules) – the main model element is *a rule*. The rule has the following general form:

RuleName: **IF** Cond(Att_{i1}) **AND** Cond(Att_{i2}) … **AND** (Cond(Att_{ik})
THEN *Cond*(Att_{ik+1}) ***AND***… *Cond*(Att_{kn})

Where:

$$Cond(Att_i) = \begin{cases} Att_i = True \\ Att_i = Val_{ik}, \quad if\ \ Att_i - symbolic \\ Att_i \in [Val_{i1}, Val_{i2}], \quad if\ \ Att_i - continuous \end{cases}$$

• *Probabilistic Models* (K-Means, Bayesian Classifiers, Hierarchical Clustering) – the main model element is *a probabilistic exemplar*. The probabilistic exemplar E_k may be described as:

$$E_k = \ll \vec{V}_1^k,...,\vec{V}_n^k >, < \vec{P}_1^k,...,\vec{P}_n^k >, c_k >,$$
$$where:$$
$$\vec{V}_i^k = \begin{cases} \mu_i^k, & if\ A_i - continuous \\ < v_{1i}^k,...,v_{ni}^k >, & if\ A_i - symbolic \end{cases}$$
$$\vec{P}_i^k = \begin{cases} \sigma_i^k, & if\ A_i - continuous \\ < p_{1i}^k,...,p_{ni}^k >, & if\ A_i - symbolic \end{cases}$$

In spite of different representations of main data model elements we assume that each model is represented as a *knowledge table*. Each row of a KT is a main model element (rule or exemplar) and each column in this table is a meaningful part of such an element (a single rule condition, a certain value (or set of values) of an attributes etc.). By applying different knowledge generation functions to the same A-MDMT it is possible to build a set of different KT. Each KT can be visualized by different visualizers using different visualization techniques (see example of 2D PAA visualization of K-NN algorithm and of a classification rule – Fig. 6). However all of them are able to run a unified set of visualization functions, which does not depend on the technique used. All such functions operate either on *an active KT (A-KT)* or on its own graphical data structure corresponding to such A-KT (or on both). The developed set of such visualization functions is reach enough and assists the user in the sophisticated process of data analysis.

4.4.3. Functions for Knowledge Manipulation

In DaMiS it should be possible to use data models produced *outside* of DaMiS – by other ML or DM programs as well as to export the models built by the system for use in other applications. The user of DaMiS will also able *to edit manually* the data models built by or imported to DaMiS.

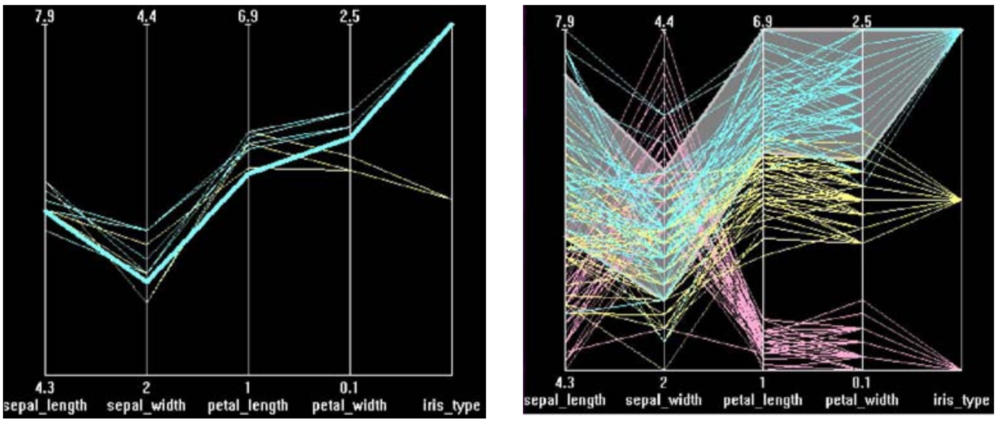

Fig. 6. Examples of 2D PAA visualization of K-NN algorithm and a rule

4.5. Knowledge evaluation functions

DaMiS has a set of functions for assessing the specified data model. Some models have specific procedures for evaluating the model quality, however, most types of data models built in DaMiS may be evaluated via their predictive accuracy. The process of evaluating the data model quality may be accomplished in two modes: interactive and non-interactive.

The interactive mode of model evaluation allows assessing the quality of *different parts* of the model or/and evaluating the detailed model behaviour on one or several test examples. Such examples may be constructed manually by the user, may be selected from existing DMTs or may be exported from text files or databases. Examples used for testing the data model are placed into a *test MDMT* (T-MDMT). In the interactive evaluation mode the active knowledge table (A-KT) is used as the data model to be assessed.

The non-interactive mode of model evaluation allows assessing the quality of the specified data model *as a whole*. Such an evaluation produces the summary of statistics describing predictive accuracy of the tested model. The specific feature of this evaluation mode is that one evaluates *the given modelling technique* applied to the given dataset (A-MDMT) rather than the concrete model built from this data. For these purposes it is possible to use such evaluation techniques as hold out, random sampling, cross-validation and stratification [16, 9]. In practice, the non-interactive evaluation consists of three main phases: splitting the A-MDMT on learning and testing parts according to the specified method, generating a set of specified models on the learning data, testing the models on training data and calculating the final statistics.

5. Conclusion and future work

Development of the multifunctional platform for EMG studies is aiming at assisting EMG practitioners in creating EMG standard examination procedures as well as at analysing and evaluating the EMG methods. The platform will integrate EMG signal analysis systems, the EMG knowledge base, clinical databases and the data mining system. In order to implement the proposed multifunctional platform the following problems have been solved:

- *Establishing a consensual view of the EMG domain.* An attempt to reuse and update already existing knowledge included into the EMG knowledge based systems developed during ESTEEM project have been done. The existing consensual knowledge have been be extracted from such systems and placed into the EMG knowledge Web server. This corpus of knowledge would be constantly extended on a consensual basis.
- *Developing taxonomy of the EMG problem domain* having in mind the possibility of applying such advanced IT methods as Data Mining, Machine Learning and Web-based and Database technologies. The EMG problems, which possible solutions will be really useful for EMG practitioners and which could be lead for further standardisation of EMG diagnostics process, have been identified.
- *Relating the developed taxonomy to the advanced IT systems that are under development by the IT partners.* An attempt to reuse, adapt or enhance the existing software systems or to develop new ones in such areas as Data Mining, Machine Learning, Database Management and Web-based Processing have been done.

The future work will emphasize on *finding a sufficient level of compatibility between such systems* providing their integration into a uniform multifunctional platform.

As a basis for such integration the database of consensus EMG cases will be used. Both ECCO file format and a relational format compatible with the SQL standard will be used for communicating between different modules of the proposed multifunctional platform. An accent will be put on enhancing the developed prototypes of the advanced case collection and management tool and the data mining system DaMiS.

References

1. Agre, G., P. Michailov. A system for two-dimensional visualization of multidimensional data. – In: Proceedings of the Intern. Conference "Computer Systems and Technology – CompSysTech'2000", Sofia, June 2000 (forthcoming).
2. Chapman, P., J. Clinton, T. Khabaza, T. Reinart z, R. Wirth. The CRISP-DM process model. – CRISP-DM Discussion Paper, March 1999, CRISP-DM Consortium. J. Cruz and P. Barahona, A Causal-Functional Model Applied to EMG Diagnosis, Lecture Notes in Artificial Intelligence, vol. 1211, Springer Verlag, March 1997, 249-260.
3. Datta, P., D. Kibler. Symbolic nearest mean classifiers. – In: Proc. of AAAI'97, AAAI Press, 1997.
4. Domingos, P. Beyond Independence: Conditions for the Optimality of the Simple Bayesian Classifier. – Proceedings of the 13th International Conference of Machine Learning (ICML'96), Bary, Italy, July 3-6, 1996, 105-113.
5. Domingos, P. Unifying Instance-Based and Rule-Based Induction. Machine Learning **24(2)**, 1996, 141-168.
6. Fayyad, U. M., K. B. Irani. Multi-interval discretization of continuous-valued attributes for classification learning. – In: Proceedings of the 13th International Joint Conference on Artificial Intelligence, Morgan Kaufmann, 1993, 1022-1027.
7. Johnsen, B., A. Fuglsang-Frederiksen. USER-4.1: Updated specification of a common EMG dataset. ESTEEM report Advanced Informatics in Medicine (A2010), **38**, 1993.
8. Johnsen, B., A. Fuglsang-Frederiksen. USER-5: EMG dataset specification version 3. ESTEEM report Advanced Informatics in Medicine (A2010), 1994, 1-44.
9. Kohavi, R. Wrappers for Performance Enhancement and Oblivious Decision Graphs, Ph.D. Thesis, Dept. of Computer Sciences, Stanford University, 1995.
10. Kononenko, I., M. Robnik-Sikonja, U. Pompe. Relief for estimation and discretization of attributes in classification, regression, and ILP problems. – In: A. Ramsay (Ed.). Artificial Intelligence: Methodology, Systems, Applications. IOS Press, 1996, 31-40.
11. Page, M., J. Gensel, C. Capponi, C. Bruley, P. Genoud, D. Ziebelin, Representation de connaissances au moyen de classes et d'associations: le systeme AROM. – Actes Langages et Modeles a Objets (LMO'00), Mont Saint-Hilaire (QC, CA), (25-28 janvier), 2000, 91-106.
12. Quinlan, J. R.. C4.5: Programs for Machine Learning, Morgan Kaufmann, 1996.
13. Richards, B. L., R.J. Mooney. Automated refinement of first-order Horn-clause domain theories, – Machine Learning, Boston , Kluwer Academic Publishers, 1995.
14. Schofield, I. S.. HINT – High level inference tool. J. Intelligent Systems. – Special Issue on Computational Intelligent Diagnostic Systems in Medicine, **8**, 1998, 1-2, 85-98.
15. Vingtoft, S., A. Fuglsang-Frederiksen, J. Ronager, J. Petrera, B. Stigsby, R. Willison, J. Jarratt, P. Fawcett, I. Schofield, G. Otte, G. Sieben, A. Vila. KANDID – An EMG decision support system – evaluated in a European multicenter trial. – Muscle and Nerve, **16**, 1993, 520-529.
16. Weiss, S., C. Kulikowski. Computer Systems That Learn. Morgan Kaufmann, Sant Mateo, 1990.
17. Zhang, J. Selecting Typical Instances in Instance-Based learning. – In: Proceeding of the ML'92 Workshop, Morgan Kaufmann, San Mateo, California, 1992, 470-470.
18. Ziebelin, D., A. Vila. Modeling planning knowledge in electromyographic diagnosis. – In: Proceedings of the PECO-INRIA-IIT Familiarization Workshop on Integrated Reasoning Architectures. IIT Working Papers, **IIT/WP-24**, ISSN 1310-652X, December 1996, 85-94.

Многофункциональная платформа для исследовании в области электромиографии

Геннадий Павлович Агрэ, А. Нелчинов

Институт информационных технологий, София

(Р е з ю м е)

Описана многофункциональная платформа для исследований в области электромиографии (ЭМГ), которая разрабатывается коллективом международного проекта EMG-NET 977069 в рамках европейской научной программы INCO-COPERNICUS. Архитектурно-программные средства платформы разделены на четыре уровня. Нижний уровень реализован в виде средства для сбора и обработки ЭМГ информации и ее сохранения в виде реляционной базы данных (система KATE). Оперативный уровень представлен в виде набора экспертных систем по ЭМГ диагностике. Уровень самообучения представлен ввиде системы извлечения закономерностей из данных (система DaMiS), а обучающий уровень реализован ввиде Web-сервера, содержащего различные виды знаний по ЭМГ. Основное внимание в работе уделено подробному описанию функций системы DaMiS.