The Knowledge Engineering Review, Vol. 00:0, 1–24. © 2004, Cambridge University Press
DOI: 10.1017/S000000000000000 Printed in the United Kingdom

A Generic and Extensible Automatic Classification Framework Applied to Brain Tumour Diagnosis in HealthAgents

Carlos Sáez[1], Juan Miguel García-Gómez[1], Javier Vicente[1], Salvador Tortajada[1], Jan Luts[2], David Dupplaw[3], Sabine Van Huffel[2] and Montserrat Robles[1]

[1] *IBIME, Instituto de Aplicaciones de las Tecnologías de la Información y de las Comunicaciones Avanzadas (ITACA), Universidad Politécnica de Valencia, Camino de Vera s/n, 46022 Valencia, España*
E-mail: {carsaesi, juanmig, javirob, vesaltor, mrobles}@upv.es
[2] *Department of Electrical Engineering (ESAT), Research Division SCD, Katholieke Universiteit Leuven, Kasteelpark Arenberg 10, B-3001 Leuven, Belgium E-mail: {jan.luts, sabine.vanhuffel}@esat.kuleuven.be*
[3] *Intelligence, Agents, Multimedia Group, School of Electronics and Computer Science, University of Southampton, Highfield, Southampton, SO17 1BJ, UK E-mail: dpd@ecs.soton.ac.uk*

Abstract

New biomedical technologies enable the diagnosis of brain tumours using non-invasive methods. HEALTHAGENTS is a European Union funded research project which aims to build an agent-based distributed Decision Support System (dDSS) for the diagnosis of brain tumours. This is achieved using the latest biomedical knowledge, Information and Communication Technologies (ICT) and Pattern Recognition (PR) techniques. As part of the PR development of HEALTHAGENTS, an independent and automatic Classification Framework (CF) has been developed. This framework has been integrated with the HEALTHAGENTS dDSS using the HEALTHAGENTS agent platform. The system offers: (1) the functionality to search for distributed classifiers to solve specific questions; (2) automatic classification of new cases; (3) instant deployment of new validated classifiers; and (4) the ability to rank a set of classifiers according to their performance and suitability for the case in hand. The CF enables both the deployment of new classifiers using the provided Extensible Markup Language (XML)[1] classifier specification; and the inclusion of new PR techniques that make the system extensible. These features may enable the rapid integration of PR laboratory results into industrial or research applications, such as the HEALTHAGENTS dDSS. Two classification nodes have been deployed and they currently offer classification services by means of dedicated servers connected to the HEALTHAGENTS agent platform; one node being located at the Katholieke Universiteit Leuven, Belgium, and the other at the Universidad Politécnica de Valencia, Spain. These classification nodes share the current set of brain tumour classifiers that have been trained from in-vivo magnetic resonance spectroscopy (MRS) data. The combination of the CF with a distributed agent system constitutes the basis of the Brain Tumour dDSS developed in HEALTHAGENTS.

1 Introduction

The purpose of the HEALTHAGENTS project (2005-2008) (13)[2] is to develop a dynamic tool for physicians that provides decision support for the diagnosis of brain tumours. It is also intended to enable the generation and sharing of knowledge that will help in the development of innovative, state-of-the-art improvements as well as analytical tools for brain tumour research. To achieve these goals, two lines of research are being investigated:

[1] Extensible Markup Language, `http://www.w3.org/XML/`
[2] `http://www.healthagents.net/`

(1) the applied research of PR techniques to develop automatic classifiers for brain tumour diagnosis by means of in-vivo MRS, as well as ex-vivo high resolution magic angle spinning spectroscopy (HR-MAS) and gene-expression microarray data;
(2) the design and implementation of a CF in order to make the automatic brain tumour classification functionality available to the HEALTHAGENTS system.

In (13) the main attributes of the HEALTHAGENTS dDSS were introduced. From the research on PR carried out in HEALTHAGENTS, a generic and extensible CF has been produced, that can be adapted to different problems and fields. This work describes in depth the full design and capabilities of such a CF and its integration in the HEALTHAGENTS dDSS. The CF presented here, in addition to integrating particular brain tumour classification problems into the system (using their corresponding classifiers), constitutes the core of the HEALTHAGENTS dDSS for brain tumour diagnosis.

The first Clinical Decision Support Systems (CDSSs) used in clinical practice were designed in the 1970s. Leaper et al. (19) developed a Clinical Decision Support System (CDSS) for the support of diagnosis and surgery of acute abdominal pain based on a naive Bayesian approach. In the early 80s, Shortliffe et al. (26) designed a CDSS for assisting physicians with the treatment of cancer patients receiving chemotherapy. Since then, a continuous series of specific CDSSs have been published in the clinical and technical journals. Most studies have focused on the solution of questions related to specific medical problems, such as breast (20), gastrointestinal (33; 22), hematologic (8; 25; 31), oral (24), lung (4), bladder (27), or prostate (18) cancer detection or diagnosis. The systems implemented for providing clinical solutions have mainly been non-distributed systems and few systems have invested effort in providing generic distributed solutions to standardize the incorporation of predictive models in CDSSs (12). The HEALTHAGENTS project has been one of the first projects where agent technology has been incorporated in a CDSS to provide a multidisciplinary network for clinical decision support (13). There are many software tools based on PR methodologies that are mainly focused on the design and building of predictive models. By contrast, the CF presented here has been conceived to offer decision support to industrial or research platforms/applications in clinical and/or biomedical environments. However, it can even be used in other environments.

An earlier European project, INTERPRET (2002) (29), provided the conceptual basis for the Decision Support System (DSS) for brain tumour diagnosis developed in HEALTHAGENTS. The current version of INTERPRET DSS (2.0) offers brain tumour classification by means of two linear discriminant analysis (LDA) classifiers for single-voxel short and long time of echo (TE) spectra. INTERPRET demonstrated that MRS-based brain tumour diagnosis could improve classical magnetic resonance imaging (MRI) diagnosis. Newly designed HEALTHAGENTS dDSS use this concept to offer a distributed system where multiple classifiers for different clinical questions, developed by state-of-the-art PR techniques and MRS advances, could be accessed from clinical centres in a transparent way. This knowledge-sharing environment enables a continuous improvement in the DSS capabilities. Other projects, such as eTumour (2008)[3], are offering a non-distributed DSS for brain tumour diagnosis. The HEALTHAGENTS approach takes advantage of the benefits both the CF and the distributed design provide:

(1) the system ensures — due to the CF capability to be easily extended with new methods — that new improvements in PR and brain tumour classification can be rapidly integrated from validated research results into clinical environments.

(2) new classifiers can be added to the framework by including their specification files. These specification files are created following the provided XML specification, as shown in (12). The use of the XML standard enables new classifier specifications

[3] `http://www.etumour.net/`

to be extended when the framework is improved with new methods. It also enables the inclusion of additional information about the field to which a classifier is related. This information can be displayed to the user to enhance the decision support task. In addition, XML-based documents can be interpreted by many computer tools and applications. This may enable the management of classifiers as well as its sharing with other systems or tools.

(3) the CF provides a classification ranking tool to sort the available classifiers in the network according to their performance, and the similarity to the test case of the contextual information in the training data. This is useful to the user since it gives a weighted measure of the performance of the classifiers, and the similarity of the case being evaluated with respect to the training cases. When comparing, for instance, two brain tumour classifiers with similar performances but built from two different datasets with populations centered on specific age ranges, the ranking will give a higher score to those candidates nearest to the age of the patient. This may occur, for instance, when classifying children cases.

(4) the HEALTHAGENTS dDSS allows physicians to classify new cases using distributed classifiers developed by different research groups. The system enables new, or updated, and validated classifiers to become automatically available on the network when deployed by their developers. By means of the distributed system created by HEALTHAGENTS, data can be shared with the scientific community on the network, enabling improved diagnostic tools to be offered by the dDSS.

Following an agent-based approach, the HEALTHAGENTS dDSS architecture is distributed over various nodes of the network (Figure 1). The network consists of clinical nodes, nodes that provide user interactions, and classification nodes. Clinical nodes have a double role: to provide data and become the final users of the dDSS. Each classification node offers a set of classifiers for different types of data, and different clinical questions, while using different feature extraction and/or classification methods. This architecture enables a clinical centre to act simultaneously as a clinical and classification node. This means a clinical centre can classify its data locally with local classifiers.

The multi-agent systems paradigm on which the HEALTHAGENTS dDSS is based, enables a classification using a set of distributed classifier agents. An agent in the network is able to request the classification of new cases from a set of classifier agents and obtain a set of results that can be combined, ranked, or displayed via a visualization on the CDSS. Although in current development and research, clinical nodes can assign permissions to their anonymised data to grant access to agents for the creation of new classifiers or the re-training of existing ones. The security implications of the network are thoroughly described in (32). The multi-agent approach, together with the CF, enables interaction within the multidisciplinary research network created by HEALTHAGENTS.

At the time of writing, two classification nodes and three clinical nodes are running within the HEALTHAGENTS system. Clinical nodes are located at Birmingham Children's Hospital, United Kingdom; the Universidad de Valencia, Spain; and the Universitat Autònoma de Barcelona, Spain. One of the classification nodes is located at the Katholieke Universiteit Leuven[4], Belgium, and another at the ITACA Institute in the Universidad Politécnica de Valencia[5], Spain. The classification nodes share 25 new classifiers for multiple binary and multiclass clinical questions regarding brain tumour diagnosis. The available classifiers are running under the CF and use the following PR techniques: Fisher LDA (7); k-nearest neighbour (KNN) (5); least square support

[4]`http://www.esat.kuleuven.be/`
[5]`http://ibime.upv.es/bmg/`

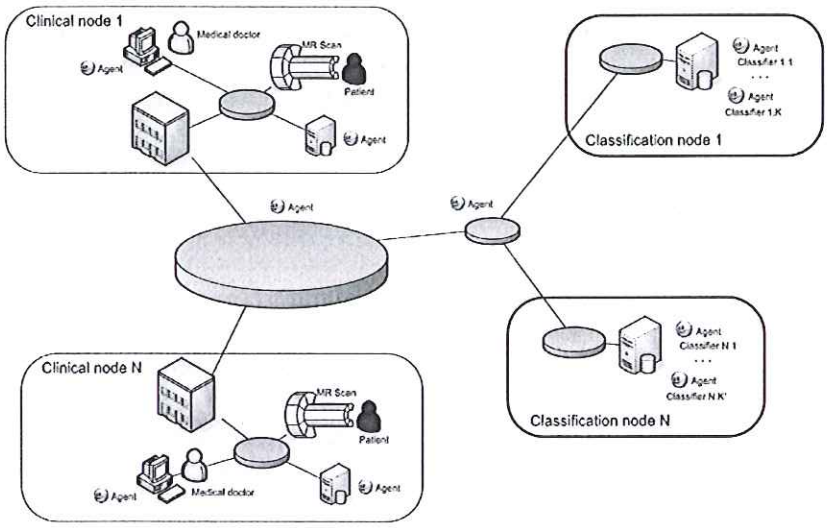

Figure 1 HEALTHAGENTS network displaying clinical and classification nodes.

vector machines (LS-SVM) (28); principal component analysis (PCA) (14); and independent component analysis (ICA) (3; 16; 17). Classifiers have been trained with data acquired during the INTERPRET and HEALTHAGENTS projects. The classification ranking tool included in the framework establishes a ranking of the results of a classification performed by a set of classifiers. Any clinical user in the HEALTHAGENTS network can use these CF functionalities for brain tumour diagnosis or research.

2 Materials and methods

2.1 Biomedical data for brain tumour diagnosis

Diagnosis and treatment of brain tumours is typically based on clinical symptoms, radiological appearance and histopathological diagnosis of a biopsy of the tissue. In the research framework of HEALTHAGENTS, information extracted from different biological levels has been acquired. Metabolomic profiles of brain tumours are acquired both in-vivo, with the non-invasive ^{1}H-MRS technique, and ex-vivo, by means of HR-MAS on a tissue sample extracted from the biopsy. Additionally, the gene-expression profile, determined using DNA microarrays, may enhance the classification of the tumour types not trivially distinguished by morphological appearance. Data availability and expertise on ^{1}H-MRS classification has enabled the development and integration of a complete set of ^{1}H-MRS brain tumour classifiers in the HEALTHAGENTS dDSS using the CF. In addition, recent studies (2) developed within the context of HEALTHAGENTS have produced the first gene-expression profiling classifiers — which can now be integrated into the CF.

The capability of ^{1}H-MRS to provide chemical information about various metabolites characterizing brain tumours makes this biomedical signal suitable for in-vivo diagnosis of brain tumour (15). Two typical acquisition protocols, short TE and long TE, are defined for brain tumour explorations. Short TE (20-35 ms) enables the observation of several metabolites and compounds useful for tumour classification — but whose signals have a stronger sensitivity to artifacts. Long TE (132-135 ms) is less informative, but easier to analyze than short TE. Additionally, clinical data such as age, gender, geographic origin of the patient, and brain tumour location, are being used together with ^{1}H-MRS for the radiological diagnosis of brain tumours. In addition, a combination of signals from long and short TE spectra might be useful for obtaining

better results in the classification of brain tumours (10). Prior to analysis, spectra are pre-processed according to INTERPRET protocol[6].

Histopathological classification of brain tumours is described in the World Health Organization (WHO) classification of tumours of the central nervous systems (21). This taxonomy is the base of the classes defined in the HEALTHAGENTS classifiers. In addition, groups based on tumour character, or grades, are also predicted by these classifiers.

2.2 Pattern recognition techniques for brain tumour diagnosis

Pattern recognition aims to automatically infer regularities in data by means of computational and mathematical tools. PR enables the extraction of useful knowledge from information, such as from a set of acquisitions of biomedical data. PR involves several stages before producing a classifier. In the first stage, after gathering the data to be used for the creation of the classifier, a pre-processing procedure is applied in order to normalize the data samples so that they are comparable and ready to use. Once the data has been properly pre-processed, the training process tunes the parameters of the mathematical model used by the PR method. This parameter tuning is carried out by means of repeated estimations of the performance. This estimation is generally obtained by using resampling techniques, typically with cross validation, although others, such as k-random sampling, have also been applied in HEALTHAGENTS (10). Finally, when the parameters are obtained, the evaluation of the classifier is performed by an independent test set in order to obtain the performance of the classifier. In studies where there is a lack of data, the classifier performance is provided by the performance obtained during the tuning stage.

In the context of brain tumour diagnosis, procedures such as data pre-processing, feature selection, and extraction are crucial — and can drastically affect the general accuracy of the system (23; 10). In this context, one PR technique may work better for a specific problem than another. Our experiments follow evaluation protocols to ensure the reliability of the obtained classifiers and their accuracies (11). A brief description is given of the PR techniques currently included in the CF used by the current set of brain tumour classifiers:

LDA: Given a set of samples, the objective of the LDA is to find a projection that minimizes the variance by class and maximises the separation between the projected means of the classes. Fisher's proposed solution is a ratio of the differences between the projected means and a measure of the dispersion of each class.

KNN: A widely-used algorithm in pattern classification. It is a non-parametric, instance-based learning method. Based on a metric and given any point in the feature space, it assigns the most frequent class label among the k training examples closest to this point. The special case where the class is predicted to be the class of the closest training sample (i.e. when $k = 1$) is called the nearest neighbour algorithm.

LS-SVM: The support vector machines (SVM) is a (non) linear technique for the classification and regression of functions defined in the context of statistical learning theory and structural risk minimization. SVM are based on the concept of hyperplanes defining decision boundaries that separate samples belonging to different classes. This kernel-based technique maps the original samples into a high-dimensional feature space where the linear-separating hyperplane with maximal margin is built. The explicit construction of a mapping to a higher dimensional feature space is avoided by using the kernel trick. LS-SVM introduces a least squares term in the cost function of the SVM. This transforms the problem into that of solving a set of linear equations. This characteristic makes LS-SVM attractive for solving high dimensionality problems, and it has been extensively applied to ^{1}H-MRS-based data.

[6]Available from `http://azizu.uab.es/interpret/mrsdata/mrsdata.html`

PCA: This is a feature extraction technique for reducing multidimensional data sets to lower dimensions for analysis. PCA is an orthogonal linear transformation that maps the data into a new coordinate system so that the greatest variance by any projection of the data comes to lie on the first coordinate (called the first principal component), and the second greatest variance on the second coordinate, etcetera. PCA can be used to reduce dimensionality in a data set — while retaining those characteristics of the data set that contribute most to its variance — by keeping lower-order principal components and ignoring higher-order ones.

ICA: A technique that extracts statistically independent components from a set of measured signals. The method is used for blind source separation and can extract a pure signal from a set of mixtures. The JADE (17) implementation of ICA has been included in the CF. An MRS signal from a specific voxel can contain components of different tissue types, known as the partial volume effect. This is the main motivation for applying ICA to a set of MRS or MRS imaging.

2.3 ICT technologies

One of the requirements of the dDSS is the continuous incorporation of research advances into the clinical environment in the shortest period of time. For that reason, the software to be developed must be scalable, robust and maintainable. Several software engineering practices are suitable for these objectives, and most of these practices can be applied to the design of the system architecture. Independent components of the system can be identified and modularized, improving the scalability and reusability of the software. It is essential for this type of software to provide a method for adding new features such as new research improvements, without modifying the underlying architecture. Object-oriented programming enables such flexibility in software development and offers most of the aforementioned attributes. Design patterns (1) and, more specifically, object-oriented design patterns provide a set of solutions to problems that commonly occur when developing systems with such requirements. These solutions have been reached through applied expertise; they have already been tested; and offer a guarantee for the developers (9). One example of a design pattern is the 'strategy' pattern, which allows a client to interchangeably use different algorithms or procedures using the same interface. This is the strategy followed in the development of the classification methods as each classifier can be seen as a set of puzzle pieces that can be assembled in the order specified by the classifier developers. The multi-agent software methodology is a good choice when the final system needs to be distributed on a network, or when entities that solve different problems work together to solve a new problem, and when entities communicate asynchronously when sharing data. This is the case of the HEALTHAGENTS dDSS. A specific agent framework has been developed to provide the required technological platform.

The HEALTHAGENTS agent framework is developed on top of the Java agent development framework JADE[7] — which follows the agent communication standard FIPA[8]. This specific agent framework offers an application program interface (API) for the development of specific agent systems, such as the HEALTHAGENTS agent framework. In addition, the Resource Description Framework (RDF)[9] based HEALTHAGENTS-agent language facilitates the creation of the messages that allow agents to communicate. In order to enable agents in the HEALTHAGENTS network to find each other, a directory service called Yellow Pages is used. The Yellow Pages agent enables agents to advertise the services they can offer so that agents in the system can decide the best services to communicate with; e.g. a dDSS client agent, by consulting the Yellow Pages, can easily find classifier agents able to classify specific data. Local and general Yellow Pages run throughout

[7]JADE, Telecom Italia, `http://jade.tilab.com/`
[8]Foundation of Intelligent Physical Agents, IEEE Computer Society, `http://www.fipa.org/`
[9]Resource Description Framework, `http://www.w3.org/RDF/`

the network in order to improve the general efficiency of the HEALTHAGENTS system — and make the directory service robust when faced with network dropouts.

XML is used as specification for customized data or languages. By means of XML, developers can describe standard documents specific to the application. Classifiers have been represented as XML documents that specify their descriptions and the steps necessary in order to obtain a classification. The use of XML ensures that the specification of classifiers can be extended when new capabilities are added to the system. It also enables extra information to be included — such as projections of the training data, or information related to the classifier's field. For assisting decission support in brain tumour diagnosis, this supplementary information can be the mean spectra of the training set, or the metabolite information used by the feature extraction method. Moreover, XML-based documents can be interpreted by most computer technologies, facilitating their handling by many tools and applications.

3 Design of the framework

The HEALTHAGENTS agent framework is provided as a Java API. Java technology is a current standard for building distributed applications because it provides platform independence. This means that the same Java software module can be used in different applications (web, local, mobile) and even on different hardware platforms and operating systems. These advantages, together with the direct incorporation into the HEALTHAGENTS system, meant that Java was the obvious choice as programming language for the present CF.

3.1 The Classification Framework

The CF is a software module, or library, which provides classification functionality. While the CF cannot be considered a stand-alone application; the framework provides an API that enables its integration into other systems. In parallel, the CF API is used to extend the system with new methods. The entry point for client modules requiring classification functionalities is the `Classifier` class, which is the main class in the framework.

The CF separates the specification of classifiers from the implementation of the PR techniques used. As described below, a classifier added to the system can use the specification files for using any of the techniques included in the CF. This approach enables technique implementations that are already part of the system to be reused. The CF mainly focuses on the classification process, leaving the training process to the laboratory tools that individual PR research groups are accustomed to using. Therefore, techniques for inclusion consist of the corresponding algorithms related to the classification process.

PR classifiers, as defined in this framework, include different actions for making predictions, including various methods for feature selection and extraction, data transformation, and classification. Conceptually, these methods can be considered as classifier actions, where each method takes an input and provides an output. A classifier can also pipeline a set of these actions before obtaining its final result. Classifier actions are represented in the CF by the `ClassifierAction` abstract class.

Incorporation of new classifier actions into the framework is an easy task for developers. Following the 'strategy' design pattern, the implementation of a classifier action is encapsulated as an interchangeable algorithm by writing to a specific API (Figure 2). The abstract `ClassifierAction` class can be extended for creating a new action. Therefore, the developer only needs to be concerned with the implementation of the method algorithms when implementing new classifier actions.

In statistical PR, a data sample consists of an array of features, and most types of data can be expressed in this way for data-mining, e.g. medical and biological data. Even if classifiers use combinations of different data, this array could be formed as a concatenation of the two original arrays. The CF can receive these arrays, and select from them the set of features a classifier needs

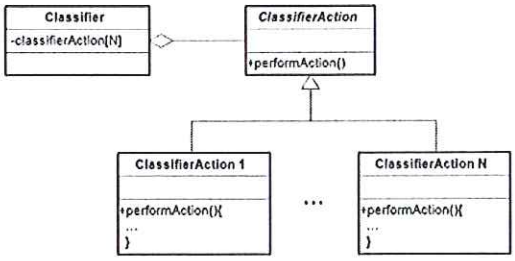

Figure 2 UML class diagram of the design of classifier actions in the CF. Using the 'strategy' design pattern, classifier action algorithms are encapsulated by the `performAction()` method.

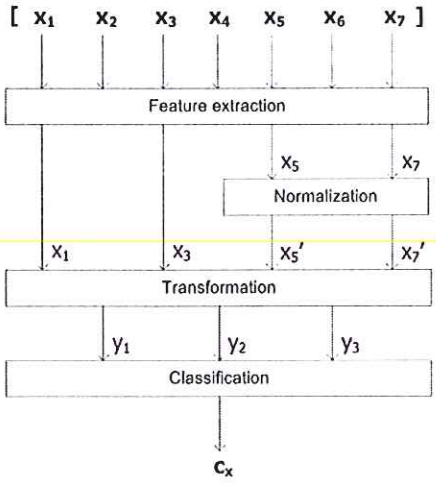

Figure 3 Example of a set of classifier actions pipelined to obtain a classification of an array of features. From the whole set of features (from x1 to x7) four are selected by the feature extraction filter (x1, x3, x5 and x7). Normalization is applied to x5 and x7. A transformation is then applied to x1, x3, and the normalized features x5' and x7'. Finally, the output of the transformation is sent to the classification module in order to obtain the posterior probabilities of each class.

to make a classification. In this way, the input of the framework becomes standard. An example of such functionality can be seen in Figure 3. Because the CF is offered as a generic tool: the classifiers are able to deal with most kinds of data and solve the questions that can be expressed by this data.

By allowing an open data input and the possibility of using any classification technique, or even a combination of techniques, the CF provides a flexible method of including classifiers in the system. An XML specification template has been defined in the CF. Therefore, to add a classifier to the system, it is only necessary to write the corresponding XML specification and add it to the framework.

The XML classifier specification is divided into three separate XML files (Figure 4):

`description.xml`: contains information related to the description of the classifier, such as what kind of input data is used, the question it solves, where it was developed, and what cases have been used for its training. Based on this file, the classifier gives information to the users so that they know whether it is suitable to their problem or not.

`profile.xml`: contains information related to the classifier's profile: classifier accuracy figures, statistical information on the training data, and dynamic values about the use of the classifier — such as a measure of the correctness of the classifier answers.

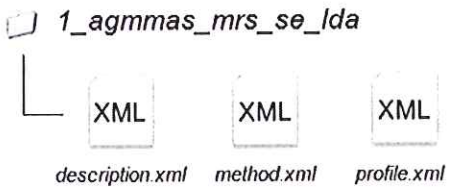

Figure 4 A classifier is represented by three XML specification files.

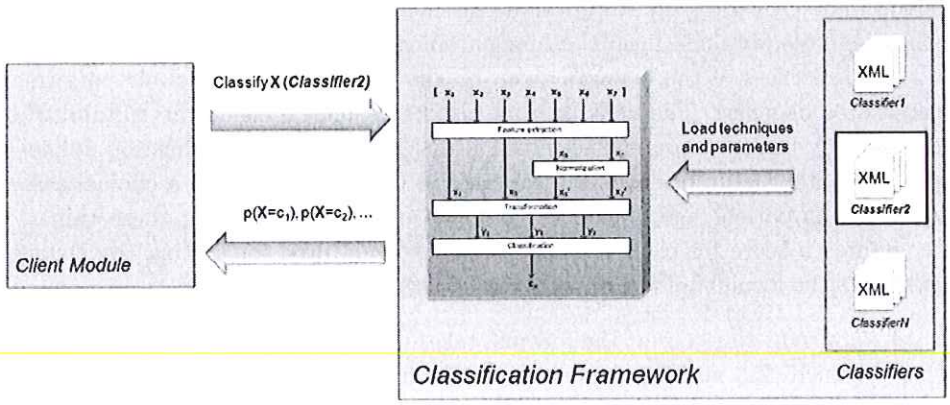

Figure 5 A client module requests the classification of X using the classifier `Classifier2` classifier. The classifier actions specified in the `Classifier2` `method.xml` file are loaded with their corresponding parameters to create the pipeline. Probabilities are sent back to the client module once classification is complete.

`method.xml`: contains the specification of the classification process. The XML specification of the classification process enables the inclusion of the actions required by the classifier, where the set of parameters used for its execution and the selection of features for its input are defined for each classifier action. Consequently, when a new classifier action is included in the framework, a description of the structure of its XML specification should be provided. The `method.xml` file describes which outputs will be returned by the classifier, and these can be any combination of output from any of the actions. In addition, extra elements such as plots or data visualizations can be included in this file.

A simple classification scenario using the CF can be described as follows: the client module requests the description of the available classifiers to check which fit its needs. Once a valid classifier is selected, it is instantiated by indicating its name. The classifier parses the content of its `method.xml` specification and creates an array of classifier actions. Parameters specified in the XML are loaded by these actions. Now that the classifier is ready to perform a classification, the data sample to be classified can be received from the client. Data passes through the pipeline of classifier actions, and generates the corresponding outputs at each step. Once the final action is complete, the classifier generates the corresponding output which is sent back to the client module. This procedure is illustrated in Figure 5.

A ranking tool is also provided as a module in the framework. Since several classifiers could coexist in the system, clients may want to have the classifiers ranked in order to identify which are more suitable for a particular case to diagnose. This tool is also used to rank the obtained results from a set of classifiers, so the higher the position in the ranking, the more reliable the result. In addition, it can also be useful to solve possible conflicts between classifiers by giving contradictory answers: which can happen when a test case is close to a decision boundary in one or more classifiers.

The ranking algorithm is based on a statistical method to compute a score for classifiers. Two types of classifier performance are taken into account: static and dynamic. Static performance is the balanced accuracy rate (BAR) obtained after training, during the evaluation stage of the classifier. It is an initial approach in assessing how good the classifier is according to the evaluation stage — and is usually based on an independent test set, or other techniques such as k-random sampling or cross validation. Dynamic performance is a measure that represents the performance of the classifier over time, based on user validations of previous classifications carried out in the system. To work properly, feedback of the diagnosis from the test cases is needed from the clinicians. The CF stores the results that classifiers provided for a case. When a case is labeled with the diagnosis obtained from the histopathological analysis, the CF is notified and, depending on the correctness of the answers given by the classifiers, the dynamic performance values will increase or decrease. The ranking tool also takes into account the suitability of a case for the classifier by using its contextual information. The classifier specification defines the distributions of the contextual information of the cases in the training set. In a clinical case, for example, values such as patient age or gender could be used. By combining these values, the ranking tool computes a score for each classifier, which is then used for sorting the classifiers. Equation 1 represents the formula of the ranking model used in the ranking tool when comparing two classifiers $\mathcal{M}_m$ and $\mathcal{M}_l$. Notice that the formula takes into account the data with which each classifier has been trained ($\mathbf{Z}_m$ and $\mathbf{Z}_l$, respectively). The first term of the ratio, $\frac{P(k^\kappa|\mathcal{K}_m)}{P(k^\kappa|\mathcal{K}_l)}$, refers to the contextual score, where $\mathcal{K}_m$ and $\mathcal{K}_l$ represent the contextual information associated with each classifier (i.e. the mean age of the training cases); and k^κ is the contextual information of the current test case (namely, the age of the patient being tested); mid term, $\frac{P(\mathbf{Z}|\mathbf{Z}_m,\mathcal{M}_m)}{P(\mathbf{Z}|\mathbf{Z}_l,\mathcal{M}_l)}$, refers to the evaluation of the classifier while 'in lab' and before releasing it into the system. Finally, the rightmost term represents the dynamic performance, that is, the measurement of how well each classifier is predicting over time when working in the system. The cases being predicted are represented in the formula with $\mathbf{Z}_{test}$.

The combined use of the CF with the ranking tool enables an intelligent management of sets of classifiers. Due to several factors, such as a poor design, overtraining, or an unrepresentative training set, the level of classifier trust may vary over time. The dynamic component of the ranking tool penalizes, or rewards, classifiers depending on their accuracy over time. The more reliable classifiers earn better scores. Therefore, in a large set of classifiers, even for the same question, the best classifiers will be identifiable, so leaving to the application using the CF the task of selecting which classifiers to show to the final user. This approach enables the use of multiple classifiers for the same question, each built from different PR methods or trained from different data bases. Moreover, the contextual component of the ranking tool will emphasize those classifiers most suitable for a case of particular contextual information. Therefore, the ranking model can make intelligent recommendations over a large set of classifiers.

$$\frac{P(\mathcal{M}_m|\mathbf{Z},\mathbf{Z}_{test})}{P(\mathcal{M}_l|\mathbf{Z},\mathbf{Z}_{test})} = \frac{P(k^\kappa|\mathcal{K}_m)}{P(k^\kappa|\mathcal{K}_l)} \frac{P(\mathbf{Z}|\mathbf{Z}_m,\mathcal{M}_m)}{P(\mathbf{Z}|\mathbf{Z}_l,\mathcal{M}_l)} \frac{P(\mathbf{Z}_{test}|\mathbf{Z}_m,\mathcal{M}_m)}{P(\mathbf{Z}_{test}|\mathbf{Z}_l,\mathcal{M}_l)} \tag{1}$$

3.2 Integration with HEALTHAGENTS

Classifiers for brain tumour diagnosis studies have been built for ^{1}H-MRS. They have successfully been integrated into the HEALTHAGENTS system using the CF presented in this study.

The CF classifies ^{1}H-MRS data in HEALTHAGENTS. ^{1}H-MRS data signals are pre-processed by a pre-processing module, which for the HEALTHAGENTS project is the Data Manipulation Software (DMS) (29). This is a MRS pre-processing tool developed during the INTERPRET project. After pre-processing, the output signal consists of an array of 512 real numbers, which is the array of features dealt by the PR system.

HEALTHAGENTS nodes communicate via the HEALTHAGENTS agent platform. Clinical nodes are able to make secure petitions for classifications across all of the classifiers in the classification nodes.

Two kinds of agents have been developed for the classification scenario in HEALTHAGENTS: the classifier petitioner agent and the classifier agent. The Classifier petitioner agent is responsible for managing classification requests from clinical nodes. Once a clinical node has determined an appropriate set of classifiers to be used for the case in hand, this set of suitable classifiers, together with the case to be classified, is sent to the classifier petitioner agent. The classifier petitioner agent has been designed to handle different classification requests concurrently because a clinical node might be running several dDSS clients simultaneously. When the classifier petitioner Agent receives a request, it is resent to every classifier in the list. The results obtained by each classifier arrive asynchronously to the classifier petitioner agent; therefore, results are mapped to the corresponding classification request on arrival. When the last result is received, the classifier petitioner agent uses the ranking tool provided in the CF to rank the set of classification results. Finally, the ranked set of results is sent back to the requesting agent on the clinical node.

The classifier agent represents a classifier in the HEALTHAGENTS system. This agent is the link between the CF and the HEALTHAGENTS dDSS. Classifier agents utilize the API provided by the CF to perform classifications and access classifier information. Each classifier agent has an associated classifier in the CF. When a classifier agent is started, it loads its classifier description to advertise on the network the main features of the classifier it is running. This notification is sent to the local Yellow Pages agent during its registration on the network (see Figure 6). An agent requiring services from a classifier agent will be able to find the classifier agent by sending a query to a Yellow Pages agent. For example, the GUI agent, which runs on a clinical node, can query the nearest Yellow Pages to obtain the location of all the classifier agents that can discriminate a particular type of tumour and/or classify using a particular kind of data. The Classifier Agent will then receive a message containing the data and classify it using its associated classifier implementation in the CF. Once the classifier has provided the result, a reply message will be created, containing the classification result, in order to send the response to the agent that requested the classification (see Figure 7 for further details).

Decision support in the HEALTHAGENTS dDSS is provided to physicians by the web-based graphical user interface (GUI). This GUI is connected to the HEALTHAGENTS network through a GUI agent, which is responsible for requesting classifications from the classifier petitioner agent. The dDSS provides a means for physicians to classify their cases using the classifiers running on the network. Only classifiers compatible with the data of their case will be shown via the GUI. In principle, physicians will decide which classifier they wish to use — depending their questions. But other factors, such as the suitability of a classifier for the case, are important. A physician can make a specific selection of classifiers, or can let the ranking tool provide hints about which classifiers best fit the case in hand. Once the classification is complete, results are sent to the dDSS and displayed in a specialized results screen (Figure 8). The physician can navigate through the results obtained from each classifier — and displayed in a tabular form. Since results are sorted by the ranking tool, the table shows the results from classifiers that are best suited to the case at the top, similarly to an Internet search performed by most search engines. By selecting a result in the table, further detail about this result is displayed in a user-friendly GUI. This view is divided into three components:

(1) the first component displays the resulting posterior probabilities in a pie chart together with information related to the description of the classifier.

(2) the second component, depending on the technique used, enables a dynamic visualization of the results obtained from the classifier. For instance, Fisher LDA classifiers can display the latent space of the samples used in the training where the between-class scattering is maximized while the within-class is minimized.

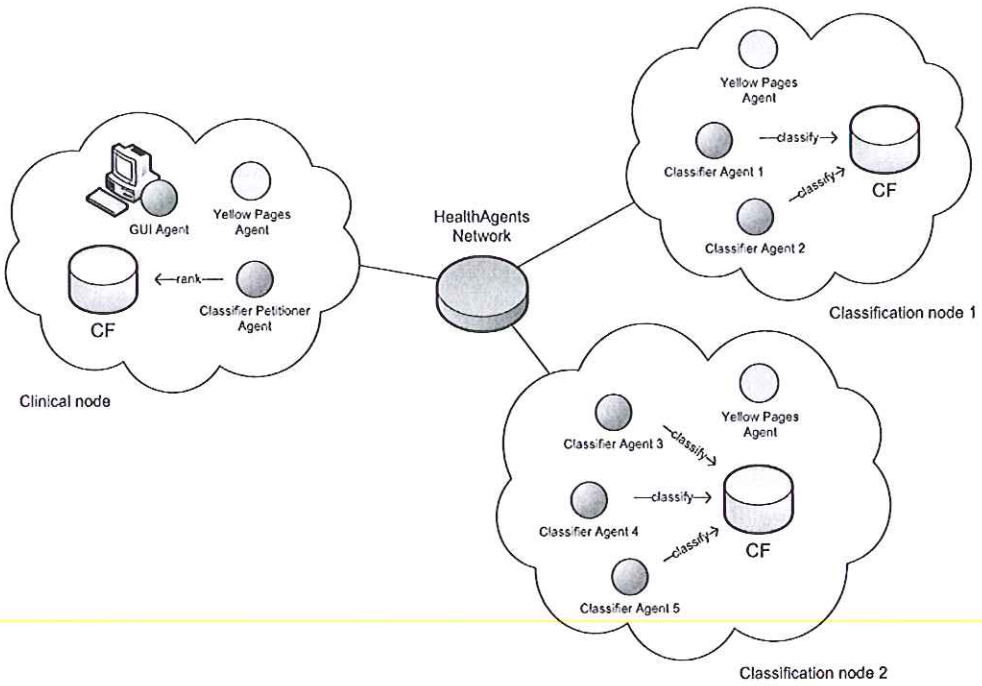

Figure 6 HEALTHAGENTS network integrating the CF. Classifier agents use the CF to perform classifications and the classifier petitioner agent uses this to rank classifiers.

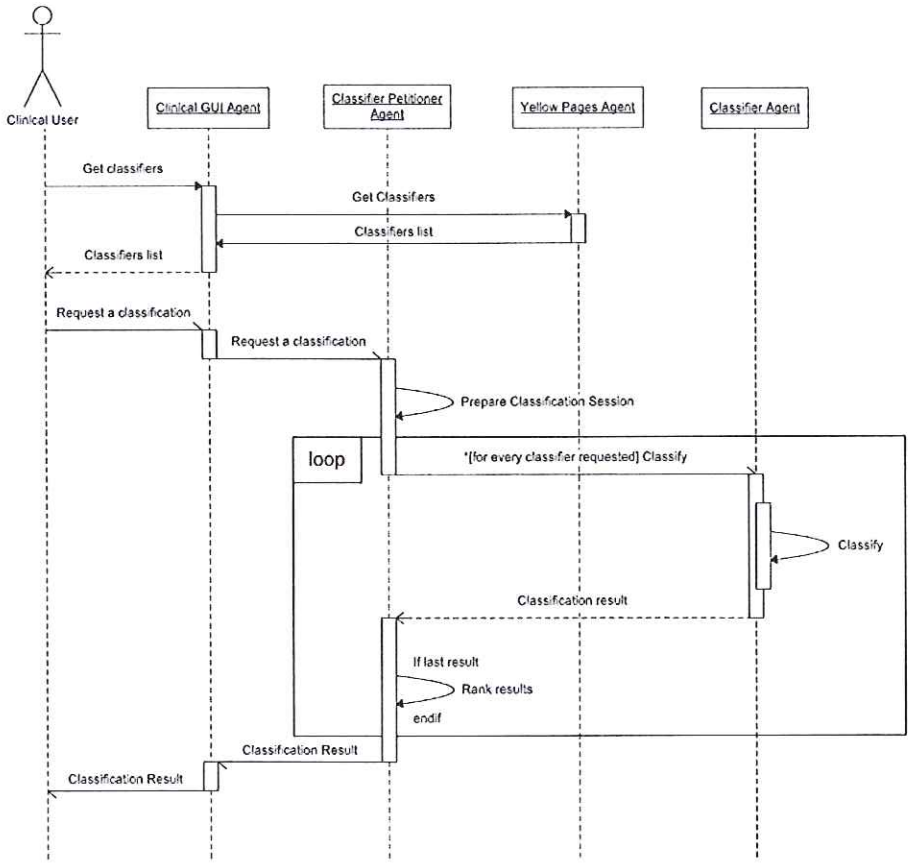

Figure 7 UML sequence diagram showing the sequence of messages sent between the agents involved in a classification started from a clinical node by a clinical user.

(3) the third component visualizes the ^{1}H-MRS data means and standard deviations per class of the ^{1}H-MRS data used in the training of the classifier — together with the spectra being classified —, so enabling a visual comparison. Such 'mean spectra classification' reinforces the results obtained from the classifier. The visualization can also display which points-of-interest of the spectra are being used for prediction by the classifier. In addition, combining this component with classifier visualization enables the spectra of the selected data to be overlaid on the MRS visualization provided that this is permitted by the security policy of the origin database.

4 Results and evaluation

A generic pattern recognition framework has been developed for providing classification services to software platforms, such as clinical and biomedical applications. The framework has been included in the dDSS developed within the HEALTHAGENTS project in order to provide classification functionality. The framework includes feature extraction, data transformation and classification methods developed by two partners from the HEALTHAGENTS project. Using these methods, experiments were carried out for each of the problems proposed by physicians in order to obtain a set of brain tumour classifiers based on ^{1}H-MRS. The best evaluated classifiers have been included in the CF and these are shown in Table 1. A ranking tool based on the ranking model designed in HEALTHAGENTS is provided in the module. To identify which are more suitable for a particular case, the tanking tool is being used to sort the results obtained from different classifiers.

As a proof of principle, two classification nodes have been set up for use in the HEALTHAGENTS project: one in the Katholieke Universiteit Leuven, Belgium, and another in the ITACA Institute of the Universidad Politécnica de Valencia, Spain. The classifiers at these nodes are being accessed from three clinical nodes: at Birmingham Children's Hospital, in the United Kingdom; the Universidad de Valencia, Spain; and the Universitat Autònoma de Barcelona, Spain. The classification nodes are running two HEALTHAGENTS agent platforms in dedicated servers connected to the HEALTHAGENTS network. The nodes are sharing the current set of classifiers, which are available in the HEALTHAGENTS network through their respective classifier agents published by each node local Yellow Pages agent.

The HEALTHAGENTS dDSS enables both adult and child brain tumour cases to be classified using the provided classifiers (Table 1). Classifiers can be sorted according to their suitability for an individual case by means of the ranking tool. This enables physicians to be notified about which answers are more reliable for a specific case. Classification results are displayed by means of the interactive results GUI (Figure 8) and this helps the clinical user obtain rapid interpretation. In addition, the physician can further investigate the set of results by obtaining detailed view of a particular classification result.

In terms of performance, the CF has been shown to behave correctly. Performance is directly related to the processing speed of the equipment the system is running on, so the faster the CF host, the better the performance the CF will provide. To obtain a measure of the prediction time of the CF, a set of tests has been carried out on a dual core PC running at 2.2Ghz with 2GB RAM. Tests consisted in performing 1000 classifications of brain tumour cases while randomly using the 25 developed classifiers. This test enabled the execution of the different routines carried out by the combination of the pattern recognition methods included in the CF. The tests revealed an average of 125 miliseconds per classification.

To illustrate the distributed use of the CF in the HEALTHAGENTS dDSS, Figure 9 shows the tools used for monitoring the agents involved in a distributed classification. Messages pass from the GUI agent at clinical nodes through the classifier petitioner agent to each classifier agent in the classification nodes. The result is received in the GUI agent.

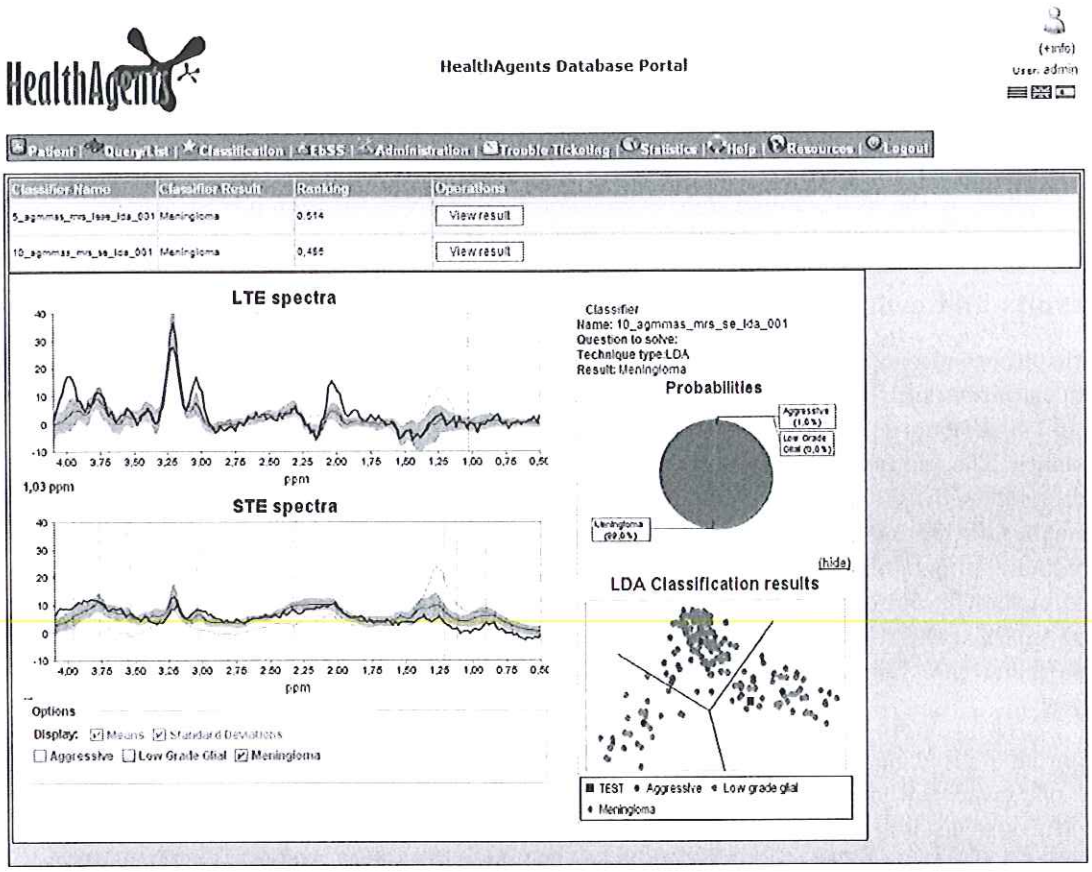

Figure 8 Classification results screen of the HEALTHAGENTS dDSS.

To study the added value the HEALTHAGENTS dDSS provides, an evaluation following the TAM methodology (6) was carried out. This evaluation was been led by a HEALTHAGENTS partner — the University of Birmingham — in the United Kingdom. In this study, 26 expert physicians were interviewed. As an overall response, they believed that the use of the dDSS would be beneficial for improving the quality of their brain tumour diagnoses. In addition, they considered the system easy to use, which is an important point in a DSS, especially in a clinical environment.

5 Conclusions and future work

A new framework for providing classification functionalities for industrial or research applications has been developed and tested in a real environment using the HEALTHAGENTS dDSS. The CF presented in this paper has illustrated that classifiers developed in laboratories can be successfully transferred and integrated into prototypes or research applications for clinical use such as the HEALTHAGENTS dDSS. Brain tumour classification is offered by the CF in a multidisciplinary research platform. Different centres can create and include their classifiers in the platform and clinical centres can access the classifiers to provide decision support for their day-to-day work, as well as for research. The CF is the core of the HEALTHAGENTS dDSS and because of the continuous improvement of PR techniques; the CF needs to be scalable and maintainable. This scalability has been illustrated with the incorporation of classifiers based on various classification methods developed by different research laboratories. Additionally, the tests performed show the robustness of the CF while solving multiple classifications using various methods. The CF and its

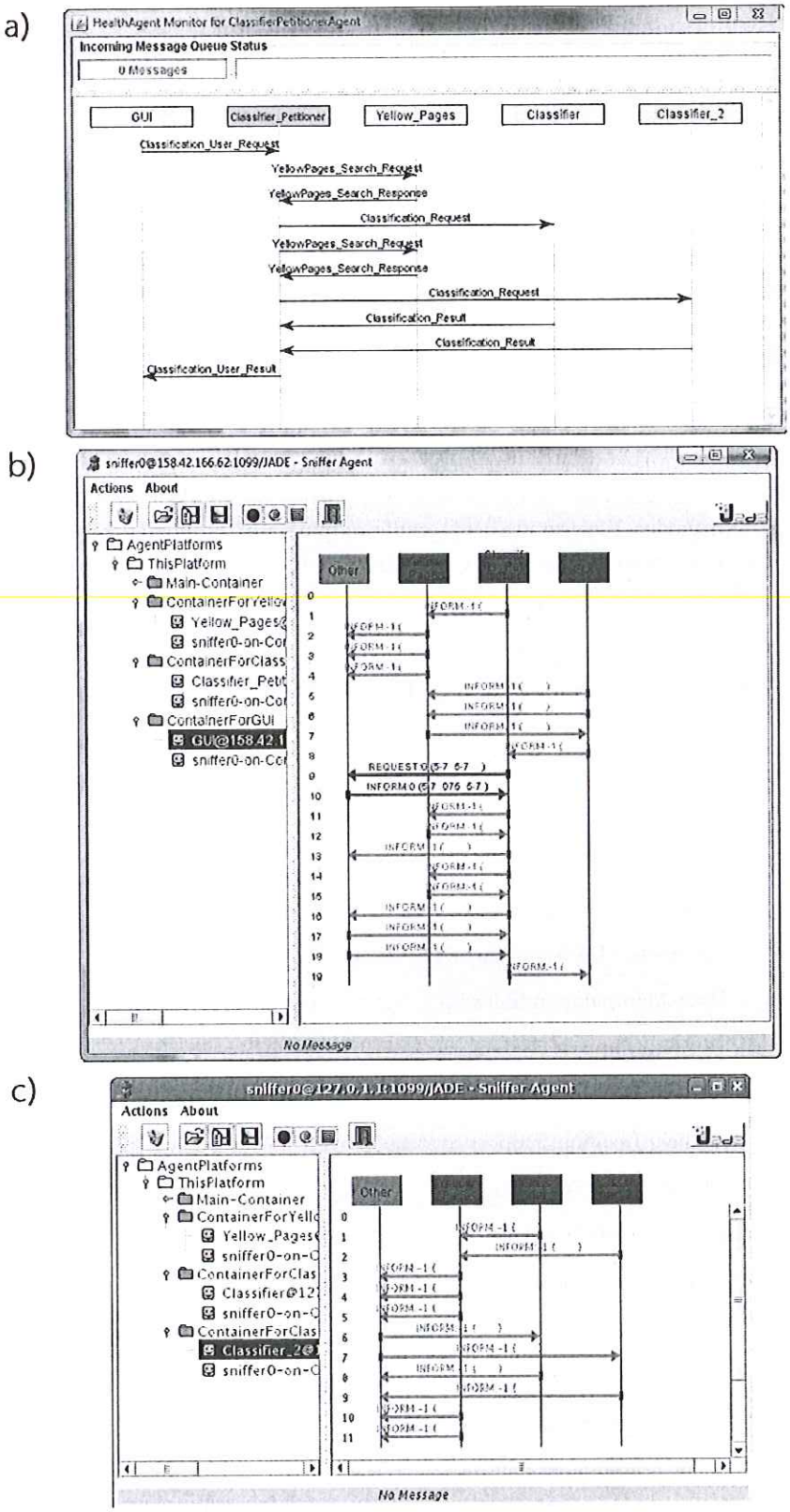

Figure 9 These three screenshots show the monitoring of a classification performed from a Clinical Node using two classifiers located in a Classification Node. *a*) shows the HEALTHAGENTS monitor of the CPA. This monitor shows the messages sent to and from the Classifier Petitioner Agent in the network. The GUI agent requests a classification to the CPA which asks the YP about the remote Classifier Agents. The CPA forwards the request to the Classifier Agents which send back their results. Finally the CPA collects the results and sends them to the requester GUI Agent. *b*) and *c*) screen show respectively the JADE Sniffer on Clinical and Classification Node. The JADE Sniffer shows the low level FIPA agent messages sent between the agents in a platform. Messages sent between remote platforms, such as messages sent between the CPA at the Clinical Node and the Classifier Agents at the Classification Node, are displayed on the sniffer as sent or received to/from OTHER agent.

use in the HEALTHAGENTS dDSS can help improve diagnostic systems — particulary for brain tumour diagnosis.

Current improvements focus on the automatic creation of new classifiers and the re-training of existing classifiers. It has been shown in (30) that the use of newly-observed data to update classifiers can outperform the behaviour of static classifiers. The advantage of a distributed DSS is that many clinical nodes can join the efforts and this can facilitate the task of compiling a good quality data base from where trusted dynamic classifiers can be updated over time. Future work focuses on the inclusion of PR techniques that will improve the CF skills, such as artificial neural networks and rule-based techniques. Furthermore, the capacity of the CF to include classifiers for more types of data will enable the inclusion of predictive models based on other biomedical data such as HR-MAS or DNA microarrays.

Acknowledgements

This work was partially funded by the European Commission: HEALTHAGENTS (contract no. FP6-2005-IST 027213). Jan Luts is a PhD student supported by an IWT grant. Carlos Sáez, Salvador Tortajada and Javier Vicente are PhD students partially supported by the Programa Torres Quevedo from Ministerio de Educación y Ciencia, co-founded by the European Social Fund (PTQ-08-01-06817, PTQ-08-01-06802, PTQ05-02-03386). We thank INTERPRET partners for their support providing data used for training some of the classifiers included in the HA network, in particular we thank C. Majós (IDI-Bellvitge), John Griffiths (SGUL), Arend Heerschap (RU), Witold Gajewicz (MUL), Jorge Calvar (FLENI), Margarida Julià-Sapé (UAB) and Carles Arús (UAB). The language revision of this document was funded by the Universidad Politécnica de Valencia.

List of abbreviations

API	application program interface
BAR	balanced accuracy rate
dDSS	distributed Decision Support System
DMS	Data Manipulation Software
DSS	Decision Support System
CDSS	Clinical Decision Support System
CDSSs	Clinical Decision Support Systems
CF	Classification Framework
GUI	graphical user interface
HR-MAS	high resolution magic angle spinning spectroscopy
ICA	independent component analysis
ICT	Information and Communication Technologies
KNN	k-nearest neighbour
LDA	linear discriminant analysis
LS-SVM	least square support vector machines
MRI	magnetic resonance imaging
MRS	magnetic resonance spectroscopy
PCA	principal component analysis

PR	Pattern Recognition
RDF	Resource Description Framework
SVM	support vector machines
TE	time of echo
XML	Extensible Markup Language

References

[1] C. Alexander, S. Ishikawa, M. Silverstein, M. Jacobson, I. Fiksdahl-King, and A. S. A pattern language. *Oxford University Press*, 1977.

[2] X. Castells, J. M. García-Gómez, A. Navarro, J. J. Acebes, Ó. Godino, S. Boluda, A. Barceló, M. Robles, J. Ariño, and C. Arús. Automated brain tumor biopsy prediction using single-labelling cDNA microarrays-based gene expression profiling. *Diagnostic Molecular Pathology*, 2009.

[3] P. Comon. Independent component analysis, a new concept? *Signal Process.*, 36(3):287–314, 1994.

[4] G. Coppini, S. Diciotti, M. Falchini, N. Villari, and G. Valli. Neural networks for computer-aided diagnosis: detection of lung nodules in chest radiograms. *IEEE Trans Inf Technol Biomed*, 7(4):344–357, 2003.

[5] T. Cover and P. E. Hart. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1):21–27, 1967.

[6] F. D. Davis. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, (Sept.):319–339, September 1989.

[7] R. A. Fisher. *Statistical methods for research workers.* Oliver and Boyd, Edinburgh, UK, 1925.

[8] D. J. Foran, D. Comaniciu, P. Meer, and L. A. Goodell. Computer-assisted discrimination among malignant lymphomas and leukemia using immunophenotyping, intelligent image repositories, and telemicroscopy. *IEEE Trans Inf Technol Biomed*, 4(4):265–273, 2000.

[9] E. Gamma, R. Helm, R. Johnson, and J. Vlissides. *Design Patterns: Elements of Reusable Object-Oriented Software.* Pearson Addison-Wesley, San Francisco, 1995.

[10] J. García-Gómez, S. Tortajada, C. Vidal, M. Julià-Sapé, J. Luts, S. Van Huffel, C. Arús, and M. Robles. The influence of combining two echo times in automatic brain tumor classification by magnetic resonance spectroscopy. *NMR in Biomedicine*, 21(10):1112–1125, 2008.

[11] J. M. García-Gómez, J. Luts, M. Julià-Sapé, P. Krooshof, S. Tortajada, J. V. Robledo, W. Melssen, E. Fuster-García, I. Olier, G. Postma, D. Monleón, Á. Moreno-Torres, J. Pujol, A.-P. Candiota, M. C. Martínez-Bisbal, J. Suykens, L. Buydens, B. Celda, S. Van Huffel, C. Arús, and M. Robles. Multiproject-multicenter evaluation of automatic brain tumor classification by magnetic resonance spectroscopy. *MAGMA*, 22(1):5–18, 2009.

[12] J. M. García-Gómez, C. Vidal, L. Martí-Bonmatí, and M. Robles. Distributed decision support architecture for soft tissue tumor diagnosis. In *MEDINFO 2004: 11th World Congress on Medical Informatics*, page 1608, 2004.

[13] H. González-Vélez, M. Mier, M. Julià-Sapé, T. N. Arvanitis, J. M. García-Gómez, M. Robles, P. H. Lewis, S. Dasmahapatra, D. Dupplaw, A. Peet, C. Arús, B. Celda, S. V. Huffel, and M. Lluch i Ariet. HealthAgents: Distributed multi-agent brain tumor diagnosis and prognosis. *Applied Intelligence*, 30(3):191–202, June 2009. Thomson Scientific Impact Factor (2007): 0.500.

[14] T. Hastie, R. Tibshirani, and J. H. Friedman. *The Elements of Statistical Learning*. Springer, August 2001.

[15] F. A. Howe and K. S. Opstad. 1H MR spectroscopy of brain tumours and masses. *NMR Biomed*, 16(3):123–131, May 2003.

[16] A. Hyvrinen, J. Karhunen, and E. Oja. *Independent Component Analysis*. John Wiley & Sons, 2001.

[17] J.-F. Cardoso and A. Souloumiac. Blind beamforming for non Gaussian signals. *IEEE Proceedings-F*, 140(6):362–370, 1993.

[18] B. M. Kelm, B. H. Menze, C. M. Zechmann, K. T. Baudendistel, and F. A. Hamprecht. Automated estimation of tumor probability in prostate magnetic resonance spectroscopic imaging: pattern recognition vs quantification. *Magn Reson Med*, 57(1):150–159, 2007.

[19] D. J. Leaper, F. T. De Dombal, J. C. Horrocks, and J. R. Staniland. Computer-assisted diagnosis of abdominal pain using estimates provided by clinicians. *Br J Surg*, 59(11):897–898, 1972.

[20] P. J. G. Lisboa, H. Wong, P. Harris, and R. Swindell. A Bayesian neural network approach for modelling censored data with an application to prognosis after surgery for breast cancer. *Artif Intell Med*, 28(1):1–25, 2003.

[21] D. Louis, H. Ohgaki, O. Wiestler, W. Cavenee, P. Burger, A. Jouvet, B. Scheithauer, and P. Kleihues. The 2007 WHO Classification of Tumours of the Central Nervous System. *Acta Neuropatol*, pages 114:97–109, 2007.

[22] P. J. Lucas, H. Boot, and B. G. Taal. Computer-based decision support in the management of primary gastric non-hodgkin lymphoma. *Methods Inf Med*, 37(3):206–219, 1998.

[23] J. Luts, J.-B. Poullet, J. M. Garcia-Gomez, A. Heerschap, M. Robles, J. A. K. Suykens, and S. Van Huffel. Effect of feature extraction for brain tumor classification based on short echo time 1H MR spectra. *Magn Reson Med*, 60(2):288–298, 2008.

[24] G. S. Nayak, S. Kamath, K. M. Pai, A. Sarkar, S. Ray, J. Kurien, L. D'Almeida, B. R. Krishnanand, C. Santhosh, V. B. Kartha, and K. K. Mahato. Principal component analysis and artificial neural network analysis of oral tissue fluorescence spectra: classification of normal premalignant and malignant pathological conditions. *Biopolymers*, 82(2):152–166, 2006.

[25] R. Ratei, L. Karawajew, F. Lacombe, K. Jagoda, G. Del Poeta, J. Kraan, M. De Santiago, J. Kappelmayer, E. Bjorklund, W.-D. Ludwig, J. W. Gratama, and A. Orfao. Discriminant function analysis as decision support system for the diagnosis of acute leukemia with a minimal four color screening panel and multiparameter flow cytometry immunophenotyping. *Leukemia*, 21(6):1204–1211, 2007.

[26] E. H. Shortliffe, A. C. Scott, M. B. Bischoff, A. B. Campbell, W. va. Melle, and C. D. Jacobs. Oncocin: An expert system for oncology protocol management. In *Seventh International Joint Conference on Artificial Intelligence*, Vancouver, 1981.

[27] P. Spyridonos, D. Cavouras, P. Ravazoula, and G. Nikiforidis. A computer-based diagnostic and prognostic system for assessing urinary bladder tumour grade and predicting cancer recurrence. *Med Inform Internet Med*, 27(2):111–122, 2002.

[28] J. A. K. Suykens and J. Vandewalle. Least squares support vector machine classifiers. *Neural Process. Lett.*, 9(3):293–300, 1999.

[29] A. R. Tate, J. Underwood, D. M. Acosta, M. Julià-Sapé, C. Majós, Á. Moreno-Torres, F. A. Howe, M. van der Graaf, V. Lefournier, M. M. Murphy, A. Loosemore, C. Ladroue, P. Wesseling, J. Luc Bosson, M. E. Cabanas, A. W. Simonetti, W. Gajewicz, J. Calvar, A. Capdevila, P. R. Wilkins, B. A. Bell, C. Remy, A. Heerschap, D. Watson, J. R. Griffiths, and C. Arús. Development of a decision support system for diagnosis and grading of brain tumours using in vivo magnetic resonance single voxel spectra. *NMR Biomed*, 19(4):411–434, Jun 2006.

[30] S. Tortajada, J. García-Gómez, J. Vicente, and M. Robles. Dynamic learning of Brain Tumour classifiers based on 1H-MRS. In SpringerLink, editor, *Book of Abstracts ESMRMB 2008 - Supplement 1, Journal Magnetic Resonance Materials in Physics, Biology and Medicine*, volume 21, pages 14–15, 2008.

[31] W. L. Tung and C. Quek. Genso-fdss: a neural-fuzzy decision support system for pediatric all cancer subtype identification using gene expression data. *Artif Intell Med*, 33(1):61–88, 2005.

[32] L. Xiao, P. Lewis, and A. Gibb. Developing a security protocol for a distributed decision support system in a healthcare environment. In *ICSE '08: Proceedings of the 30th international conference on Software engineering*, pages 673–682, New York, NY, USA, 2008. ACM.

[33] M. M. Zheng, S. M. Krishnan, and M. P. Tjoa. A fusion-based clinical decision support for disease diagnosis from endoscopic images. *Comput Biol Med*, 35(3):259–274, 2005.

Question	Features	No. Cases	FE	Classifier	ACC	BAR	Developer
AGG vs MEN vs LGG	LTE + STE	125 (71,13,41)	PPM	KNN	0,80 (±0, 07)	0,78	UPV-ITACA
AGG vs MEN vs LGG	LTE + STE	125 (71,13,41)	PPM	LDA	0,75 (±0, 07)	0,65	UPV-ITACA
AGG vs MEN vs LGG	STE	217 (124,58,35)	PPM	LDA	0,94 (±0, 03)	0,93	UPV-ITACA
AGG vs MEN vs LGG	STE	217 (124,58,35)	PKI	KNN	0,86 (±0, 04)	0,84	UPV-ITACA
AGG vs MEN vs LGG	STE	217 (124,58,35)	PCA	LDA	0,91 (±0, 04)	0,90	UPV-ITACA
AGG vs MEN vs LGG (Ad.)	STE	209 (121,55,33)	PKI	LDA	0,90 (±0, 04)	0,88	UPV-ITACA
MEN vs MET	STE	96 (58,38)	PCA	KNN	0,91 (±0, 05)	0,91	UPV-ITACA
MEN vs MET	STE	94 (57,37)	ICA	LS-SVM	0,88 (±0, 07)	0,89	KUL
MEN vs MET	STE	94 (57,37)	ICA	LS-SVM	0,90 (±0, 06)	0,91	KUL
HG vs LG (Ad.)	STE	237 (134,103)	PKI	LDA	0,87 (±0, 05)	0,87	UPV-ITACA
HG vs LG	STE	262 (147,115)	PKI	LDA	0,85 (±0, 04)	0,84	UPV-ITACA
HG vs LG (Ch.)	STE	70 (28,42)	PCA	LDA	0,69 (±0, 10)	0,68	UPV-ITACA
AFF vs NON-AFF	STE	304 (22,282)	PKI	KNN	0,99 (±0, 01)	0,94	UPV-ITACA
TUM vs NON-TUM	STE	282 (274,8)	PKI	KNN	0,97 (±0, 02)	0,64	UPV-ITACA
AGG vs NON AGG	STE	244 (135,109)	PKI	KNN	0,81 (±0, 05)	0,81	UPV-ITACA
AGG vs NON AGG (Ch.)	STE	79 (30,49)	PCA	KNN	0,72 (±0, 10)	0,71	UPV-ITACA
GLM vs EMB (Ch.)	STE	58 (43,15)	PCA	KNN	0,72 (±0, 11)	0,62	UPV-ITACA
GLM vs EMB (Ch.)	STE	48 (33,15)	PCA	KNN	0,71 (±0, 13)	0,66	UPV-ITACA
GBM vs LGG	STE	117 (84,33)	ICA	LS-SVM	0,83 (±0, 07)	0,83	KUL
GBM vs LGG	STE	117 (84,33)	ICA	LS-SVM	0,84 (±0, 07)	0,87	KUL
GBM vs MEN	STE	141 (84,57)	ICA	LS-SVM	0,84 (±0, 06)	0,79	KUL
GBM vs MEN	STE	141 (84,57)	ICA	LS-SVM	0,91 (±0, 07)	0,89	KUL
MEN vs LGG	STE	90 (57,33)	ICA	LS-SVM	0,84 (±0, 08)	0,83	KUL
MEN vs LGG	STE	90 (57,33)	ICA	LS-SVM	0,92 (±0, 06)	0,92	KUL
MET vs LGG	STE	70 (37,33)	ICA	LS-SVM	0,85 (±0, 08)	0,87	KUL

Table 1 The table shows the 25 brain tumour classifiers for solving the different clinical problems currently integrated into HealthAgents. First column indicates the tumour types/groups the classifier discriminates (aggressive tumour (AGG), meningioma (MEN), low grade glial (LGG), metastasis (MET), high grade tumour (HG), low grade tumour (LG), affected tissue (AFF), non affected tissue (NON-AFF), tumour (TUM), non tumour (NON-TUM), non aggressive tumour (NON-AGG), glioma (GLM), embryonal tumour (EMB), glioblastoma (GBM)). It is indicated whether the classifier is specific for adults (Ad.) or children (Ch.). The second column indicates the signals used by the classifier. The third column indicates the size of the training set and the distribution of cases per label. The fourth column indicates the feature extraction (FE) method employed by the classifier (parts-per-million (PPM), peak integration (PKI)). The fifth column indicates the accuraccy (ACC) of the classifier. The sixth column indicates the BAR. The last column indicates the developer group of the classifier (Katholieke Universiteit Leuven, Belgium (KUL); ITACA Institute at the Universidad Politécnica de Valencia, Spain (UPV-ITACA)).