The Knowledge Engineering Review, Vol. 26:3, 283–301. © Cambridge University Press, 2011
doi:10.1017/S0269888911000129

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]*Grupo de Informática Biomédica, Instituto de Aplicaciones de las Tecnologías de la Información y de las Comunicaciones Avanzadas (ITACA), Universidad Politécnica de Valencia, Edificio 8G, Camino de Vera s/n, 46022 Valencia, España;*
e-mail: carsaesi@upv.es, juanmig@upv.es, javirob@upv.es, vesaltor@upv.es, 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@esat.kuleuven.be, 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 by using non-invasive methods. HealthAgents is a European Union-funded research project that 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 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[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 data. The combination of the CF with a distributed agent system constitutes the basis of the brain tumour dDSS developed in HealthAgents.

[1] Extensible Markup Language (XML), `http://www.w3.org/XML/`

https://doi.org/10.1017/S0269888911000129 Published online by Cambridge University Press

1 Introduction

The purpose of the HealthAgents project (2006–2008; González-Vélez *et al.*, 2009)[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. the applied research of pattern recognition (PR) techniques to develop automatic classifiers for brain tumour diagnosis by means of *in vivo* magnetic resonance spectroscopy (MRS), as well as *ex vivo* high-resolution magic angle spinning (HR-MAS) and gene-expression microarray data;
2. the design and implementation of a classification framework (CF) in order to make the automatic brain tumour classification functionality available to the HealthAgents system.

In González-Vélez *et al.* (2009) the main attributes of the HealthAgents distributed decision support system (dDSS) were introduced. From the research on PR carried out in HealthAgents, a generic and extensible CF has been produced, which can be adapted to different problems and fields. This work describes the full design and capabilities of such a CF and its integration in the HealthAgents dDSS in depth. 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.* (1972) developed a CDSS for the support of diagnosis and surgery of acute abdominal pain based on a naive Bayesian approach. In the early 1980s, Shortliffe *et al.*(1981) 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 (Lisboa *et al.*, 2003), gastrointestinal (Lucas *et al.*, 1998; Zheng *et al.*, 2005), haematologic (Foran *et al.*, 2000; Tung & Quek, 2005; Ratei *et al.*, 2007), oral (Nayak *et al.*, 2006), lung (Coppini *et al.*, 2003), bladder (Spyridonos *et al.*, 2002) or prostate (Kelm *et al.*, 2007) 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 (García-Gómez *et al.*, 2004). The HealthAgents project is one of the first projects, in which agent technology has been incorporated in a CDSS to provide a multidisciplinary network for clinical decision support (González-Vélez *et al.*, 2009). There are many software tools based on PR methodologies that are mainly focused on the design and building of predictive models. In 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; Tate *et al.*, 2006), 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 in which 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

[2] http://www.healthagents.net/

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

https://doi.org/10.1017/S0269888911000129 Published online by Cambridge University Press

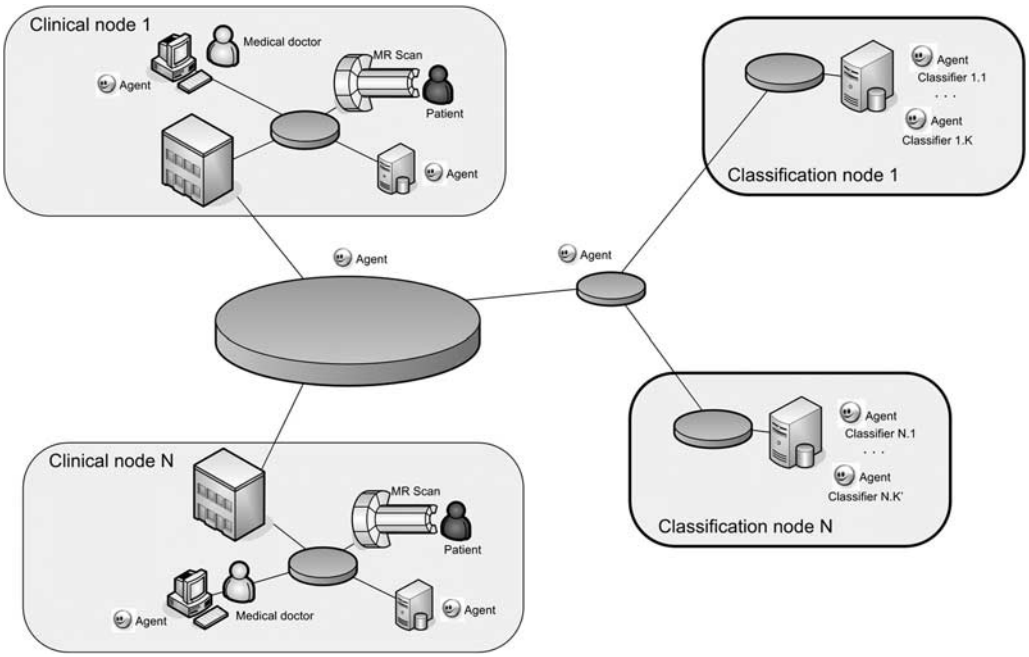


Figure 1 HealthAgents network displaying clinical and classification nodes

DSS for brain tumour diagnosis. The HealthAgents approach takes advantage of the benefits that both the CF and the distributed design provide:

1. The system ensures—because of the CF's 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 Extensible Markup Language (XML) specification, as shown in García-Gómez *et al.* (2004). The use of the XML standard enables new classifier specifications to be extended when the framework is improved with new methods. It also enables the inclusion of additional information regarding 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 of the test case to the contextual information in the training data. This is useful to the user as 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 a similar performance, but those that are built from two different data sets with populations centred 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 paediatric 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 used 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

https://doi.org/10.1017/S0269888911000129 Published online by Cambridge University Press

provide user interactions and classification nodes. Clinical nodes have a double role: to provide data and to 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 function simultaneously as a clinical and classification node. This indicates that a clinical centre can classify its data locally with local classifiers.

The multiagent 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 can obtain a set of results that can be combined, ranked or displayed via a visualization on the CDSS. However, in current development and research, clinical nodes can assign permissions to their anonymized 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 Xiao *et al.* (2008). The multiagent 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. The clinical nodes are located at Birmingham Children's Hospital, the 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 (Fisher, 1925); K-nearest neighbour (KNN; Cover & Hart, 1967); least square support vector machines (LS-SVM; Suykens and Vandewalle, 1999); principal component analysis (PCA; Hastie *et al.*, 2001); and independent component analysis (ICA; Cardoso & Souloumiac, 1993; Comon, 1994; Hyvärinen *et al.*, 2001). 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

The 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 proton magnetic resonance spectroscopy (^{1}H-MRS) technique, and *ex vivo*, by means of HR-MAS on a tissue sample extracted from the biopsy. In addition, the gene-expression profile, determined using deoxyribonucleic acid (DNA) microarrays, may enhance the classification of the tumour types that are 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 (Castells *et al.*, 2009) 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 regarding various metabolites characterizing brain tumours makes this biomedical signal suitable for the *in vivo* diagnosis of

[4] http://www.esat.kuleuven.be/

[5] http://www.ibime.upv.es/bmg/

https://doi.org/10.1017/S0269888911000129 Published online by Cambridge University Press

brain tumours (Howe & Opstad, 2003). 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 that are useful for tumour classification—but whose signals have a stronger sensitivity to artifacts. Long TE (132–135 ms) is less informative, but easier to analyse when compared to short TE. In addition, 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 to obtain better results in the classification of brain tumours (García-Gómez *et al.*, 2008). Prior to analysis, spectra are pre-processed according to INTERPRET protocol[6].

Histopathological classification of brain tumours is described in the World Health Organization's (WHO) classification of tumours of the central nervous systems (Louis *et al.*, 2007). This taxonomy is the base of the classes defined in the HealthAgents classifiers. In addition, groups based on tumour characteristics or grades are also predicted by these classifiers.

2.2 Pattern recognition techniques for brain tumour diagnosis

PR 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 have 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 other methods such as k-random sampling, have also been applied in HealthAgents (García-Gómez *et al.*, 2008). 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 in which 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 (García-Gómez *et al.*, 2008; Luts *et al.*, 2008). 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 (García-Gómez *et al.*, 2009). A brief description is given of the PR techniques that are currently included in the CF, which is 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 maximizes 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: This is a widely used algorithm in pattern classification. It is a non-parametric, instance-based learning method. On the basis of 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 in which 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) are (non)linear techniques 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

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

https://doi.org/10.1017/S0269888911000129 Published online by Cambridge University Press

original samples into a high-dimensional feature space in which 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 one of solving a set of linear equations. This characteristic makes LS-SVM attractive for solving the problems of high dimensionality, and it has been extensively applied to ^{1}H-MRS-based data.

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 onto 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 and so on. 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: This is 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 (Cardoso & Souloumiac, 1993) 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 Information and communication 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 (Alexander *et al.*, 1977) 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 (Gamma *et al.*, 1995). 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 the pieces of a puzzle that can be assembled in the order specified by the classifier developers. The multiagent 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 Foundation of Intelligent Physical Agents (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

[7] JADE, Telecom Italia, `http://jade.tilab.com/`

[8] FIPA, IEEE Computer Society, `http://www.fipa.org/`

[9] RDF, `http://www.w3.org/RDF/`

https://doi.org/10.1017/S0269888911000129 Published online by Cambridge University Press

facilitates the creation of the messages that allow agents to communicate. 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; for example, 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 the network in order to improve the general efficiency of the HealthAgents system—and to make the directory service robust when faced with network dropouts.

XML is used as a 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 necessary steps to take 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 decision 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 the 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, showed that Java was the obvious choice of programming language for the present CF.

3.1 The classification framework

The CF is a software module, or library, which provides classification functionality. Although 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, excluding the training process of 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.

The 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 to create a new action. Therefore, the developer only needs to be concerned with the implementation of the method algorithms when implementing new classifier actions.

https://doi.org/10.1017/S0269888911000129 Published online by Cambridge University Press

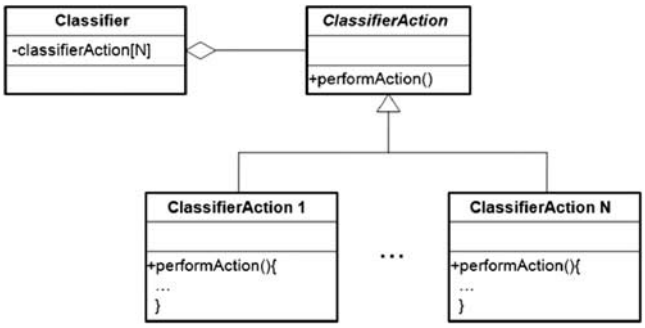


Figure 2 UML class diagram of the design of classifier actions in the CF. Using the 'strategy' design pattern, classifier action algorithms are encapsulated by the `performAction()` method

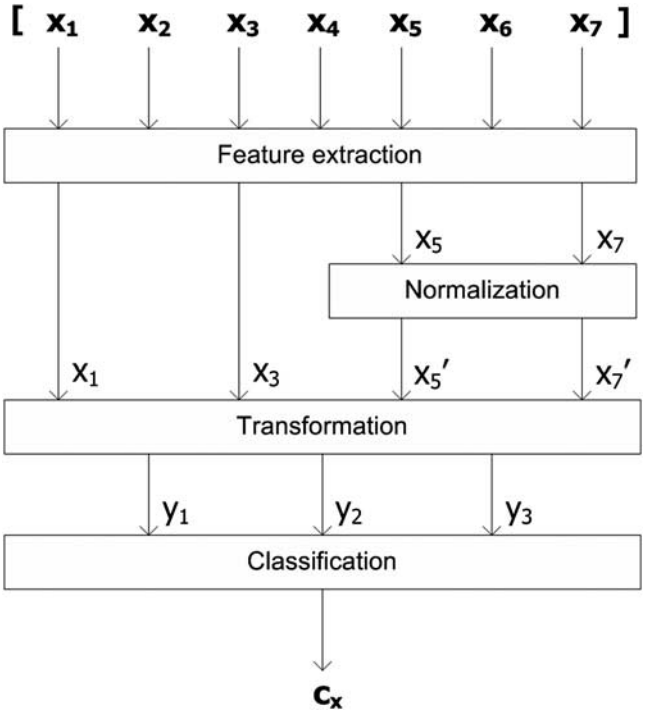


Figure 3 Example of a set of classifier actions pipelined to obtain a classification of an array of features. From the whole set of features (from x_1 to x_7), four are selected by the feature extraction filter (x_1, x_3, x_5 and x_7). Normalization is applied to x_5 and x_7. A transformation is then applied to x_1, x_3, and the normalized features x_5' and x_7'. Finally, the output of the transformation is sent to the classification module in order to obtain the posterior probabilities of each class

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, for example, medical and biological data. Even if classifiers use distinct combinations of different data, this array could be formed as a concatenation of the two original arrays. The CF can receive these arrays, and from them select the set of features a classifier needs 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. As the CF is offered as a generic tool, the classifiers are able to deal with most kinds of data and to 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

https://doi.org/10.1017/S0269888911000129 Published online by Cambridge University Press

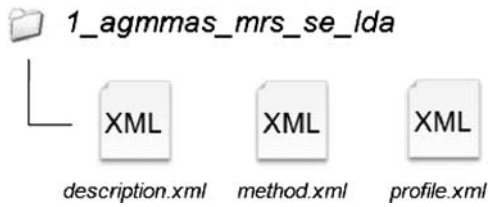


Figure 4 A classifier is represented by three XML specification files

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. On the basis of this file, the classifier gives information to the users so that they know whether it is suitable for 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 regarding the use of the classifier—such as a measure of the correctness of the classifier answers.
`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, in which 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 it 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. The 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 pass through the pipeline of classifier actions and generate 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. As several classifiers could coexist in the system, clients may want to have the classifiers ranked in order to identify those that are more suitable for a particular case to diagnose. This tool is also used to rank the obtained results from a set of classifiers; therefore, 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 occur 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 BAR (balanced accuracy rate) obtained after training, during the evaluation stage of the classifier. It is an initial approach to assess 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.

https://doi.org/10.1017/S0269888911000129 Published online by Cambridge University Press

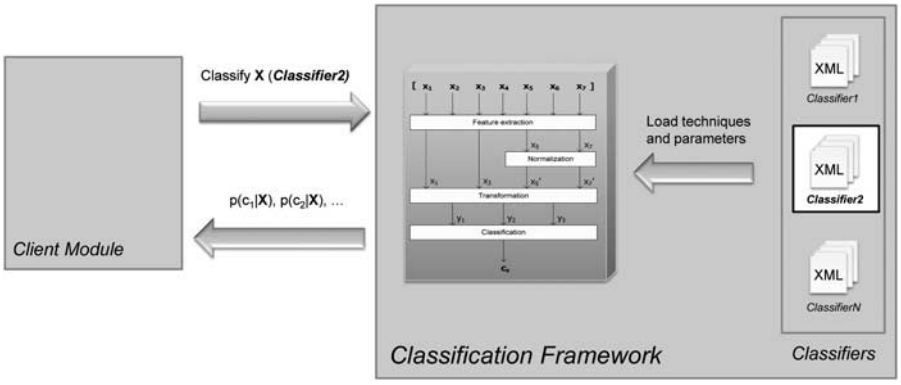


Figure 5 A client module requests the classification of X using the `Classifier2` classifier. The classifier actions specified in the `Classifier2` `method.xml` file are loaded with their corresponding parameters to create the pipeline. Probabilities are sent back to the client module once classification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

The CF stores the results that the classifiers provided for a case. When a case is labelled 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 the $\mathbf{Z}_{\text{test}}$.

The combined use of the CF with the ranking tool enables the intelligent management of sets of classifiers. Owing 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 penalises 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 databases. Moreover, the contextual component of the ranking tool will emphasize the classifiers that are most suitable for a case of particular contextual information. Therefore, the ranking model can make intelligent recommendations over a large set of classifiers.

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

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.

https://doi.org/10.1017/S0269888911000129 Published online by Cambridge University Press

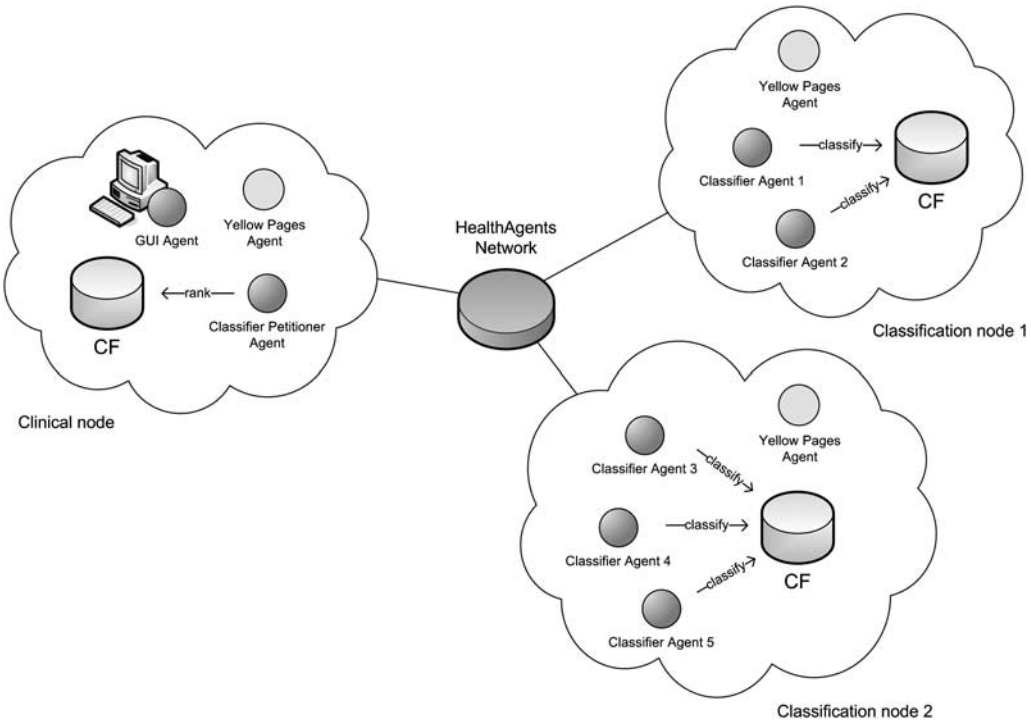


Figure 6 HealthAgents network integrating the CF. Classifier agents use the CF to perform classifications and the classifier petitioner agent uses this to rank classifiers

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; Tate *et al*., 2006). 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 re-sent 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 to 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 graphical user interface (GUI) agent, which runs on a

https://doi.org/10.1017/S0269888911000129 Published online by Cambridge University Press

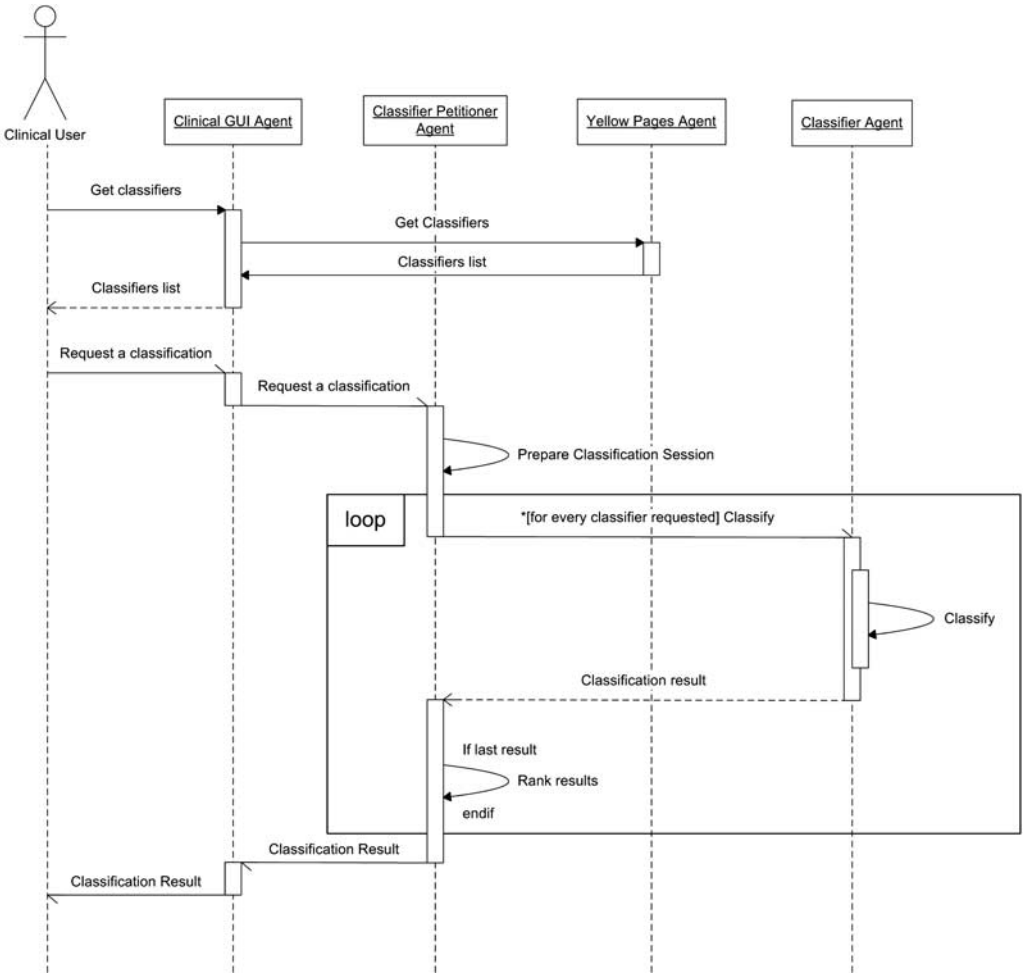


Figure 7 UML sequence diagram showing the sequence of messages sent between the agents involved in a classification started from a clinical node by a clinical user

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 a web-based GUI, which is connected to the HealthAgents network through a GUI agent and 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 the classifiers that are 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 on 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 are displayed on a specialized results screen (Figure 8). The physician can navigate through the results obtained from each classifier and display them in a tabular form. As the results are sorted by the ranking tool, the table shows the results from the classifiers that are best suited to the case at the top, which is similar to an Internet search performed by most search engines. By selecting a result in the

https://doi.org/10.1017/S0269888911000129 Published online by Cambridge University Press

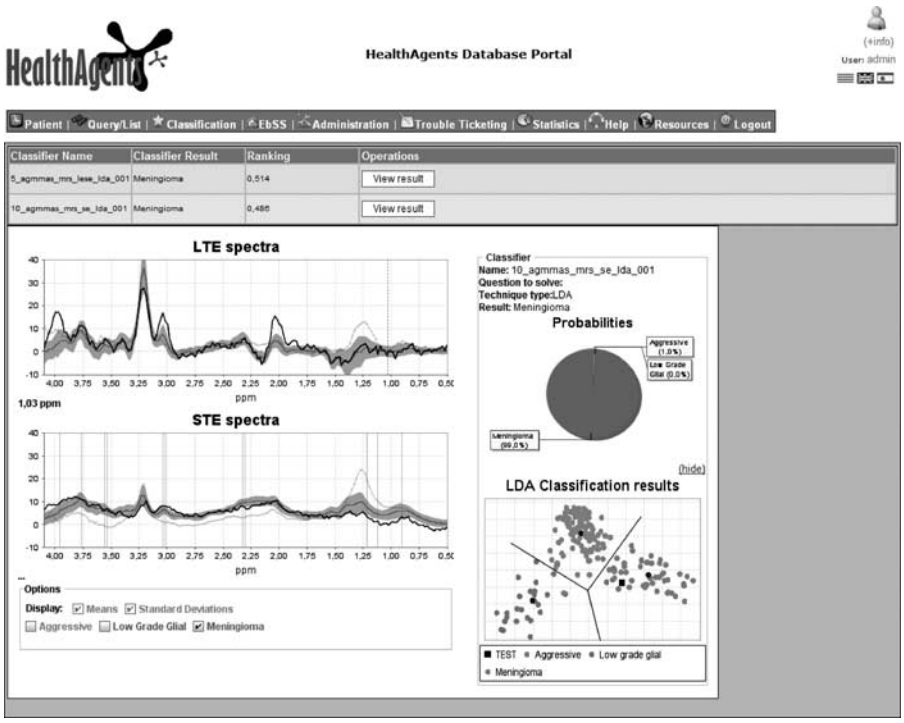


Figure 8 Classification results screen of the HealthAgents dDSS

table, further detail about this result is displayed in a user-friendly GUI. This view is divided into three components:

1. The first displays the resulting posterior probabilities in a pie chart together with information related to the description of the classifier.
2. The second, 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, whereas the within-class scattering is minimized.
3. The third 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, thus 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 PR 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

https://doi.org/10.1017/S0269888911000129 Published online by Cambridge University Press

https://doi.org/10.1017/S0269888911000129 Published online by Cambridge University Press

Table 1 Twenty-five brain tumour classifiers for solving the different clinical problems currently integrated into HEALTHAGENTS

Question	Features	Number of 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

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), long time of echo (LTE), short time of echo (STE), independent component analysis (ICA), least square support vector machines (LS-SVM), principal component analysis (PCA), K-nearest neighbour (KNN)). 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)).

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

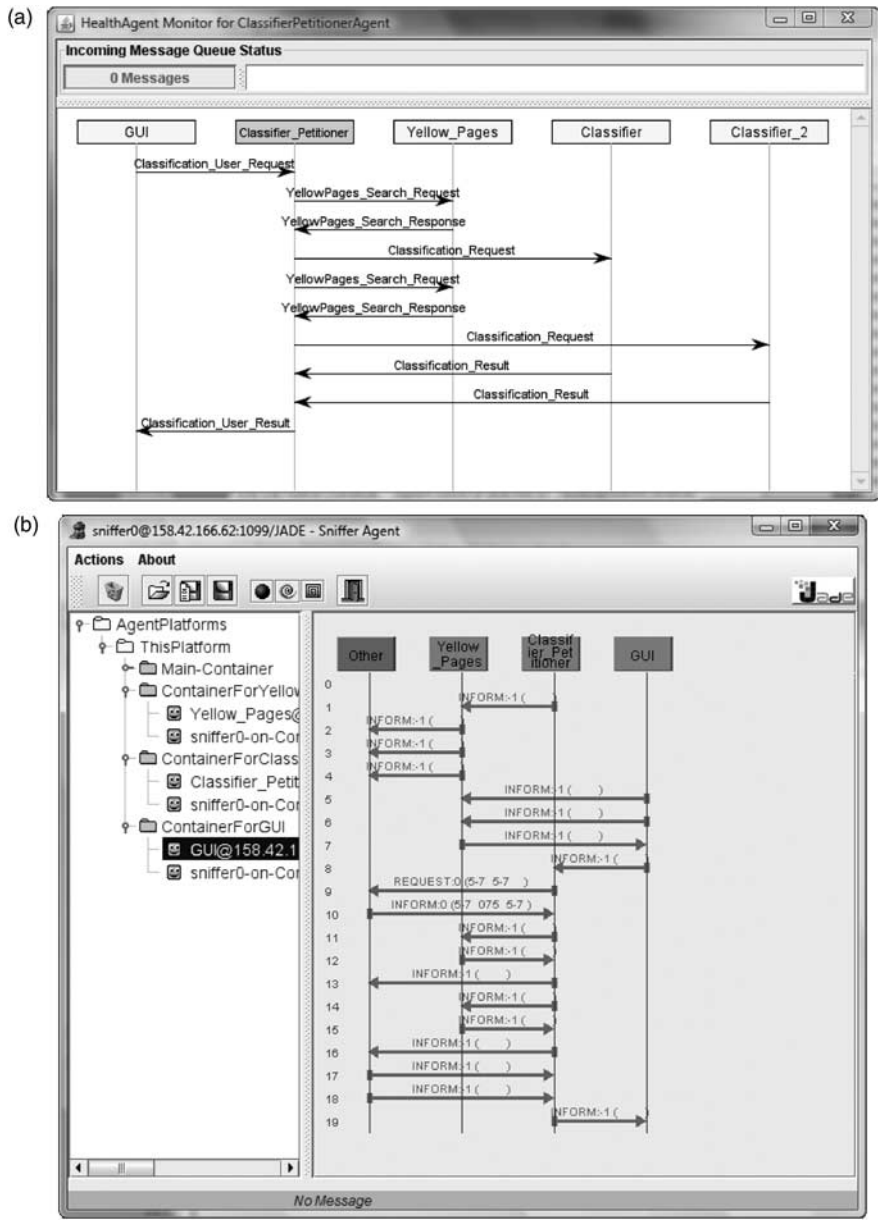


Figure 9 (Figure Continued)

https://doi.org/10.1017/S0269888911000129 Published online by Cambridge University Press

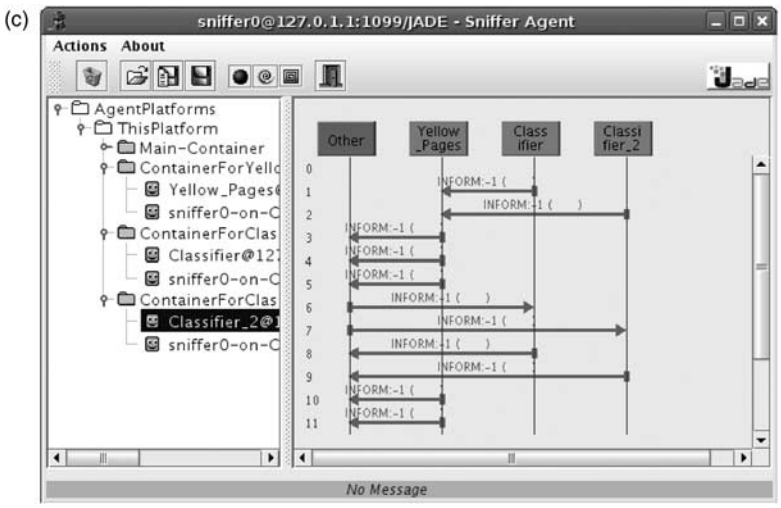


Figure 9 These three screenshots show the monitoring of a classification performed from a Clinical Node using two classifiers located in a Classification Node. (a) shows the HealthAgents ' monitor of the CPA. This monitor shows the messages sent to and from the Classifier Petitioner Agent in the network. 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. The (b) and (c) screens show the JADE Sniffer on the Clinical and the Classification Nodes, respectively. 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 and from OTHER agent

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 preliminary interpretation. In addition, the physician can further investigate the set of results by obtaining a detailed view of a particular classification result.

In terms of performance, the CF has been shown to function 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 have been carried out on a dual core PC running at 2.2 GHz with 2 GB RAM. The tests consisted of 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 PR methods included in the CF. The tests revealed an average of 125 ms 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 by the GUI agent.

To study the added value the HealthAgents dDSS provides, an evaluation following the Technology Acceptance Model methodology (Davis, 1989) was carried out. This evaluation was 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

https://doi.org/10.1017/S0269888911000129 Published online by Cambridge University Press

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. In addition, the tests performed show the robustness of the CF while solving multiple classifications using various methods. The CF and its 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 Tortajada *et al*. (2008) 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 database from which 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 027214). 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 the 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 and for providing the data used for training some of the classifiers included in the HealthAgents 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. This work has been partially supported by the Health Institute Carlos III through the RETICS Combiomed, RD07/0067/2001.

References

Alexander, C., Ishikawa, S., Silverstein, M., Jacobson, M., Fiksdahl-King, I. & Angel, S. 1977. *A Pattern Language*. Oxford University Press.

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

Castells, X., García-Gómez, J. M., Navarro, A., Acebes, J. J., Godino, Ó., Boluda, S., Barceló, A., Robles, M., Ariño, J. & Arús, C. 2009. Automated brain tumor biopsy prediction using single-labelling cDNA microarrays-based gene expression profiling. *Diagnostic Molecular Pathology* **18**, 206–218.

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

Coppini, G., Diciotti, S., Falchini, M., Villari, N. & Valli, G. 2003. Neural networks for computer-aided diagnosis: detection of lung nodules in chest radiograms. *IEEE Transactions on Information Technology in Biomedicine* **7**(4), 344–357.

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

Davis, F. D. 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly* **13**(3), 319–339.

https://doi.org/10.1017/S0269888911000129 Published online by Cambridge University Press

Fisher, R. A. 1925. *Statistical Methods for Research Workers*. Oliver and Boyd, Edinburgh, UK. http://psychclassics.yorku.ca/Fisher/Methods/.

Foran, D. J., Comaniciu, D., Meer, P. & Goodell, L. A. 2000. Computer-assisted discrimination among malignant lymphomas and leukemia using immunophenotyping, intelligent image repositories, and telemicroscopy. *IEEE Transactions on Information Technology in Biomedicine* **4**(4), 265–273.

Gamma, E., Helm, R., Johnson, R. & Vlissides, J. 1995. *Design Patterns: Elements of Reusable Object-oriented Software*. Pearson Addison-Wesley.

García-Gómez, J. M., Vidal, C., Martí-Bonmatí, L. & Robles, M. 2004. Distributed decision support architecture for soft tissue tumor diagnosis. In *MEDINFO 2004: 11th World Congress on Medical Informatics*, San Francisco, California, USA, 1608.

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

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

González-Vélez, H., Mier, M., Julià-Sapé, M., Arvanitis, T. N., García-Gómez, J. M., Robles, M., Lewis, P. H., Dasmahapatra, S., Dupplaw, D., Peet, A., Arús, C., Celda, B., Van Huffel, S. & Lluch-Ariet, M. 2009. HealthAgents: distributed multi-agent brain tumor diagnosis and prognosis. *Applied Intelligence* **30**(3), 191–202.

Hastie, T., Tibshirani, R. & Friedman, J. H. 2001. *The Elements of Statistical Learning*. Springer.

Howe, F. A. & Opstad, K. S. 2003. 1H MR spectroscopy of brain tumours and masses. *NMR in Biomedicine* **16**(3), 123–131.

Hyvärinen, A., Karhunen, J. & Oja, E. 2001. *Independent Component Analysis*. John Wiley & Sons.

Kelm, B. M., Menze, B. H., Zechmann, C. M., Baudendistel, K. T. & Hamprecht, F. A. 2007. Automated estimation of tumor probability in prostate magnetic resonance spectroscopic imaging: pattern recognition vs quantification. *Magnetic Resonance in Medicine* **57**(1), 150–159.

Leaper, D. J., De Dombal, F. T., Horrocks, J. C. & Staniland, J. R. 1972. Computer-assisted diagnosis of abdominal pain using estimates provided by clinicians. *The British Journal of Surgery* **59**(11), 897–898.

Lisboa, P. J. G., Wong, H., Harris, P. & Swindell, R. 2003. A Bayesian neural network approach for modelling censored data with an application to prognosis after surgery for breast cancer. *Artificial Intelligence in Medicine* **28**(1), 1–25.

Louis, D., Ohgaki, H., Wiestler, O., Cavenee, W., Burger, P., Jouvet, A., Scheithauer, B. & Kleihues, P. 2007. The 2007 WHO classification of tumours of the central nervous system. *Acta Neuropathologica* **114**(2), 97–109.

Lucas, P. J., Boot, H. & Taal, B. G. 1998. Computer-based decision support in the management of primary gastric non-Hodgkin lymphoma. *Methods of Information in Medicine* **37**(3), 206–219.

Luts, J., Poullet, J.-B., Garcia-Gomez, J. M., Heerschap, A., Robles, M., Suykens, J. A. K. & Van Huffel, S. 2008. Effect of feature extraction for brain tumor classification based on short echo time 1H MR spectra. *Magnetic Resonance in Medicine* **60**(2), 288–298.

Nayak, G. S., Kamath, S., Pai, K. M., Sarkar, A., Ray, S., Kurien, J., D'Almeida, L., Krishnanand, B. R., Santhosh, C., Kartha, V. B. & Mahato, K. K. 2006. 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.

Ratei, R., Karawajew, L., Lacombe, F., Jagoda, K., Del Poeta, G., Kraan, J., De Santiago, M., Kappelmayer, J., Bjorklund, E., Ludwig, W.-D., Gratama, J. W. & Orfao, A. 2007. 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.

Shortliffe, E. H., Scott, A. C., Bischoff, M. B., Campbell, A. B., Melle, W. & Jacobs, C. D. 1981. ONCOCIN: an expert system for oncology protocol management. In *Seventh International Joint Conference on Artificial Intelligence*, Vancouver.

Spyridonos, P., Cavouras, D., Ravazoula, P. & Nikiforidis, G. 2002. A computer-based diagnostic and prognostic system for assessing urinary bladder tumour grade and predicting cancer recurrence. *Medical Informatics and the Internet in Medicine* **27**(2), 111–122.

Suykens, J. A. K. & Vandewalle, J. 1999. Least squares support vector machine classifiers. *Neural Processing Letters* **9**(3), 293–300.

Tate, A. R., Underwood, J., Acosta, D. M., Julià-Sapé, M., Majós, C., Moreno-Torres, Á., Howe, F. A., van der Graaf, M., Lefournier, V., Murphy, M. M., Loosemore, A., Ladroue, C., Wesseling, P.,

https://doi.org/10.1017/S0269888911000129 Published online by Cambridge University Press

Luc Bosson, J., Cabanas, M. E., Simonetti, A. W., Gajewicz, W., Calvar, J., Capdevila, A., Wilkins, P. R., Bell, B. A., Remy, C., Heerschap, A., Watson, D., Griffiths, J. R. & Arús, C. 2006. Development of a decision support system for diagnosis and grading of brain tumours using in vivo magnetic resonance single voxel spectra. *NMR in Biomedicine* **19**(4), 411–434.

Tortajada, S., García-Gómez, J., Vicente, J. & Robles, M. 2008. Dynamic learning of brain tumour classifiers based on 1H-MRS. In *Book of Abstracts ESMRMB 2008 – Supplement 1*, *Journal Magnetic Resonance Materials in Physics, Biology and Medicine*, Springer Verlag, **21**, 14–15.

Tung, W. L. & Quek, C. 2005. Genso-fdss: a neural-fuzzy decision support system for pediatric all cancer subtype identification using gene expression data. *Artificial Intelligence in Medicine* **33**(1), 61–88.

Xiao, L., Lewis, P. & Gibb, A. 2008. Developing a security protocol for a distributed decision support system in a healthcare environment. In *ICSE'08: 30th International Conference on Software Engineering*, ACM, Leipzig, Germany, 673–682. .

Zheng, M. M., Krishnan, S. M. & Tjoa, M. P. 2005. A fusion-based clinical decision support for disease diagnosis from endoscopic images. *Computers in Biology and Medicine* **35**(3), 259–274.

https://doi.org/10.1017/S0269888911000129 Published online by Cambridge University Press