# The effect of feature extraction on MRS brain tumour classification

**J. Luts[1], J-B. Poullet[1], J. M. Garcia-Gomez[2], A. Heerschap[3], and S. Van Huffel[1]**

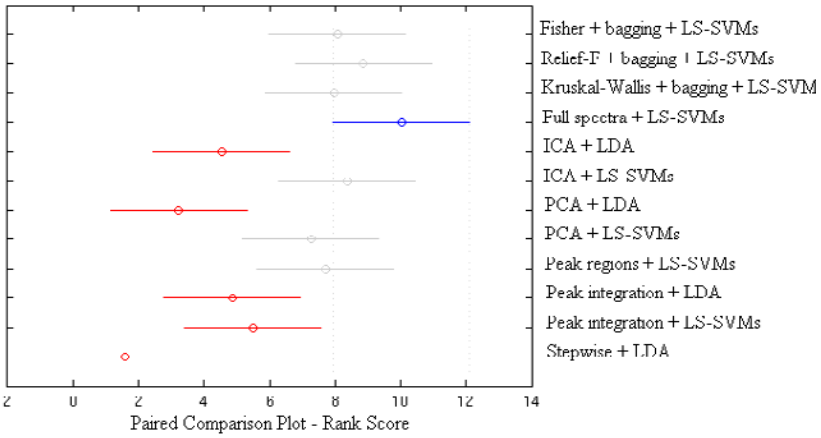
[1]Department of Electrical Engineering, Katholieke Universiteit Leuven, Leuven-Heverlee, Belgium, [2]BET-IBM, Universidad Politecnica de Valencia, Valencia, Spain, [3]University Medical Center St. Radboud, Department of Radiology, University of Nijmegen, Nijmegen, Netherlands

**Introduction –** In recent years, the use of Magnetic Resonance Spectroscopy (MRS), which provides metabolic information, has gained a lot of interest for a more detailed and specific non-invasive evaluation of brain tumours. As individual viewing and analysis of the multiple spectral patterns is time-consuming and often requires specific spectroscopic expertise, it is not practical in a clinical environment. Automatic processing and evaluation of the data and easy and rapid display of the results is needed for routine clinical interpretation of an exam. If particular pattern recognition techniques can be automated and integrated into a clinical Decision Support System (DSS), this will strongly support the clinical use of MRS. Lately, projects like eTUMOUR [1] and HealthAgents [2] attempt to introduce DSSs into clinical practice. Further, to restrict the computation time and maximize the applicability of a DSS, the preprocessing protocol of MRS should be kept as concise as possible. Consequently, before building these systems, the importance of feature selection should be addressed. Some studies have already presented results and conclude that the extraction of features out of spectra, prior to classification, is not critical [3,4]. In this work, the feature reduction problem is analyzed. Seven feature selection methods are used in a thorough study to examine the importance of good dimensionality reduction for MRS spectra, prior to feeding them to a DSS.

**Material –** From the INTERPRET database 662 multi-voxel short echo time spectra, encompassing 6 types of brain tissue from 25 patients and 4 volunteers acquired at University Medical Center Nijmegen, are selected. These include normal tissue, CSF, grade II and III glial tumours, meningiomas and glioblastomas. All spectra are semi-automatically preprocessed as described in [5].

**Methods –** To perform pattern recognition, Linear Discriminant Analysis (LDA) and standard Least Squares Support Vector Machines (LS-SVMs) are used [6]. Regarding dimensionality reduction, Principal Component Analysis (PCA), Independent Component Analysis (ICA), simple peak integration of the single peaks, peak regions, stepwise variable selection, Fisher criterion, Kruskal-Wallis test and Relief-F [7] are utilized. When using the last three methods, respectively, an extra bagging step is introduced. As Support Vector Machine (SVM) based methods can deal with a high input dimension, the full spectra (0.5 to 4.0 ppm) were also included in the analysis, in combination with LS-SVMs.

**Results –** To compare the different methods a repeated stratified sampling procedure is performed (100 repetitions) for each pair of classes. The performance of the methods is assessed by the mean Area Under the Curve (AUC) of the Receiver Operating Characteristic (ROC) curve. Using Friedman's nonparametric two-way analysis of variance test ($p < 0.001$) and Tukey's honestly significant difference criterion for multiple comparison on the mean AUC values, one observes in the figure below that using the full spectra in combination with LS-SVMs achieves the best result. A higher rank score means that the AUC value for that method is larger. Also, as noted in [8], the kernel based method performs slightly better than LDA. In the paired comparison plot below, the methods, for which the confidence intervals are indicated in red, differ significantly from LS-SVMs with full spectra. In contrast to LDA, where there is not really a clear trend concerning dimension reduction, one can observe a tendency when using LS-SVMs. Lowering the number of input features seems to decrease the performance of the LS-SVM classifier. Using peak regions improves the performance compared to peak integration, using the full spectra seems to increase, although not significantly, the performance further.




**Discussion & Conclusion** – In general, over all pairwise problems, one observes that using the complete spectra with a kernel based method achieves the best performance. This result has important practical consequences for the design of DSSs. The observations suggest simplifying the feature reduction step and preprocessing protocol of spectra when using a DSS purely for classification purposes. Additionally, this has the important advantage to speed up the execution time of a classification task in a DSS. This fact can accelerate the introduction of a DSS into clinical practice. Also, these findings stroke with the conclusion of the study on long echo time spectra, carried out in [9].

**References** – [1] http://www.etumour.net/; [2] http://www.healthagents.net/; [3] A.W. Simonetti et al., NMR Biomed., 18(1):34-43; [4] A. Devos et al., J. Magn. Res., 170 :164-175; [5] A.W. Simonetti et al., Anal. Chem., 75(20):5352-5361; [6] J.A.K. Suykens et al., ISBN 981-238-151-1; [7] M. Robnik-Sikonja et al., ML, 53: 23-69; [8] A. Devos et al., J. Magn. Res., 173:218-228; [9] L. Lukas et al., AIM, 31(1) :73-89.