Editorial

Colin Fyfe

School of Computing,
University of the West of Scotland,
High Street, Paisley Pa1 2BE, Scotland
E-mail: colin.fyfe@uws.ac.uk

The six papers in this special issue are greatly extended versions of six of the diverse collection of papers which were presented at IDEAL2010, the 11th International Conference on Intelligent Data Engineering and Automated Learning which took place at the University of the West of Scotland, September 1st to 3rd 2010. The breadth of the submissions to that conference can be judged by the breadth of papers presented here.

The first paper is 'Approximating the covariance matrix of GMMs with low-rank perturbations' by Magdon-Ismail and Purnell. This deals with the difficult problem of approximating the full covariance matrix so that it can be calculated efficiently while meanwhile enabling the resulting approximation to be powerful enough that it does actually make a valid approximation to the full covariance matrix. Their work on the TIMIT speech database shows that their method improves accuracy compared to the standard diagonal approximation.

The second paper is 'Abstraction through clustering: complexity reduction in automated planning domains' by Dicken et al. Typical planning problems exhibit a combinatorial explosion and it is this problem which their clustering method attacks. They create a series of problems ranging from easy to hard to illustrate their method.

The third paper is 'Towards a hybrid NMF-based neural approach for face recognition on GPUs' by Lopes and Ribeiro. This paper discusses a practical approach to using the graphical processing units of modern PCs on the practical problem of automatic face recognition. Two commonly used face databases are used to illustrate the speed-ups possible with the use of the GPU.

The fourth paper is 'An efficient randomised sphere cover classifier' by Younsi and Bagnall. This paper describes a method for reducing a dataset's size without sacrificing accuracy compared to standard nearest neighbour classifiers. They demonstrate the effectiveness of their method on 24-datasets from the UCI repository and on six-gene evaluation datasets.

The final two-papers use the currently popular topic of Bregman divergences: the first paper is 'A generalisation of independence in statistical models for categorical distribution' by Fujimoto and Murata. This extends the previous work of the authors on the use of Bregman divergences and shows how to use their methods to e.g., estimate distributions from small samples. They finally illustrate the use of their methods with four-datasets from the UCI machine learning repository.

The final paper is 'Dual stream data exploration' by Wang et al. This considers two-extensions of the standard statistical technique of canonical correlation analysis (CCA) which can be used for extracting information from two-data streams simultaneously. This paper extends standard CCA in two-ways by using Bregman

divergences and with the recent technique of reservoir computing. The resulting algorithms are illustrated on a database of student exam marks.

The conference itself was very successful and we hope that you will enjoy reading these extended versions of the six-conference papers.