

Supervised and hierarchical unsupervised neural networks for clustering both gene expression profiles and samples

[Alvaro Mateos](#)¹, [Javier Herrero](#)¹, [Javier Tamames](#)², [Joaquín Dopazo](#)¹

¹Bioinformatics Unit, CNIO; ²ALMA bioinformatics SA

In this paper, we describe several applications of unsupervised and supervised neural networks to the analysis of DNA microarrays gene expression profiles. Hierarchical clustering and Self Organising Maps (SOM) are among the most extensively used tools for clustering gene expression patterns from microarray experiments, in part due to the availability of software implementations. Nevertheless, both procedures, as well as other similar ones, present different problems of efficiency, accuracy and runtime. Here we show how the Self Organising Tree Algorithm (SOTA) is a fast (linear runtime) and efficient method for clustering gene expression profiles. In addition, the average values of the clusters, provided by SOTA, can be used for obtaining a classification of samples either by unsupervised or supervised methods.