

Seminário em Engenharia Matemática

Jorge Santos

Título:

"On the Use of the Adjusted Rand Index as a Metric for Evaluating Supervised Classification"

Resumo:

The Adjusted Rand Index (ARI) is frequently used in cluster validation since it is a measure of agreement between two partitions: one given by the clustering process and the other defined by external criteria. We investigate the usability of this clustering validation measure in supervised classification problems by two different approaches: as a performance measure and in feature selection. Since ARI measures the relation between pairs of dataset elements not using information from classes (labels) it can be used to detect problems with the classification algorithm specially when combined with conventional performance measures. Instead, if we use the class information, we can apply ARI also to perform feature selection. We present the results of several experiments where we have applied ARI both as a performance measure and for feature selection showing the validity of this index for the given tasks.