Standardized Partition Spaces

Ursula Garczarek

Roche Diagnostics GmbH,
Nonnenwald 2, 82372 Penzberg, Germany

Abstract. In this presentation, a general approach to the comparative presentation of classification rules is introduced: an adequate scaling into the so-called standardized partition space. This offers a unifying framework for the representation of a wide variety of classification rules.

Various theoretical frameworks of classification techniques exist in statistics and machine learning. By a schematic representation of their proceedings, it becomes clear that the differences in their basic approaches can be understood in terms of their different definitions of what is expected – the "expected loss". The main impact these varying expectations have is on the functional space in which potential classification rules lie.

Irrespective of the dissimilarity of the approaches on the theoretic level, on the technical level their proceedings are quite similar: based on training data membership functions are learnt that enable to quantify the membership of some object in classes on the basis of its predictor values.

A unified representation of these membership functions is possible in the so-called standardized partition space. Linchpin is the adequate scaling of the values from different membership functions into this space. Our scaling results in a presentation of membership assessments that reflect the classifiers characteristic of classification and give a realistic impression of the classifiers performance on a test set. The scaling can be understood in terms of a bayesian binomial-beta model to estimate the success probability of an assignment.

To give an impression of the possibilities of the procedure, we exemplify their use for comparing and interpreting classification rules, and for combining evidence on class membership from varying sources.

References

Garczarek, Ursula (2002),
"Classification Rules in Standardized Partition Spaces", url = `http://eldorado.uni-dortmund.de:8080/FB5/ls7/forschung/2002/Garczarek`,
Dissertation, Department of Statistics, University of Dortmund