THE ART OF BUILDING DECISION TREES IN THE MEDICAL DOMAIN

Peter Kokol

Many real-world medical problems are nowadays being handled with tools for automatic intelligent data analysis. Various methods have been developed to improve the quality of analysis for specific domains. Application of any method in a specific domain has special requirements. While medical experts are not “very good with numbers” we as informaticians must focus on methods, that are capable of extracting knowledge in a form closer to human perception (white box methods), e.g. methods that induce decision trees, classification rules, etc. For the same reason instance methods based on artificial neural networks (black box methods) that are nevertheless capable of generalization of nonlinearly separable problems, but have poor explanatory power are not suitable to be used in the medical domain.

Knowing the “no free lunch theorem” and the fact that normally medical experts do not have enough time and knowledge to find the best possible method for their specific problem we developed a Multimethod machine learning paradigm that can be used to automatically analyze various machine learning approaches, compare and combine them in a hybrid decision tree. In the talk we will present the new paradigm and the results obtained with using it. We will compare various decision tree approaches, purity measures, ensemble methods and finally hybrid decision trees.