Topic for a Bachelor/Master’s Thesis:

Multi-label classification with partial abstention

Intelligent Systems Group
Department of Computer Science
Paderborn University

In statistics and machine learning, classification with abstention (aka classification with a reject option) is an extension of the standard setting of classification, in which the learner is allowed to refuse a prediction for a given query instance $x$. For the learner, the main reason to abstain is a lack of certainty about the corresponding outcome—refusing or at least deferring a decision might then be better than taking a high risk of a wrong decision.

Nowadays, there are many machine learning problems in which complex predictions are sought (instead of scalar values, like in classification and regression). An important example is so-called multi-label classification (MLC), in which a response $y$ associated with an instance $x$ is a subset of an underlying reference set $\mathcal{Y}$ of class labels; that is, the output space is the power set $2^\mathcal{Y}$ [1].

For complex, structured predictions, the idea of abstaining from a prediction can be generalized toward partial abstention: Instead of predicting the entire structure, the learner predicts only parts of it, namely those for which it is certain enough. This idea has already been realized, for example, for the problem of label ranking, where predictions are rankings [2]. The goal of this thesis is to extend MLC toward partial abstention, which essentially means allowing the learner to refuse the prediction for some of the labels in $\mathcal{Y}$. Amongst others, this requires (i) a suitable characterization of uncertainty of a prediction, (ii) a generalization of MLC loss functions, and (iii) algorithmic solutions for the minimization of these generalized losses. As an MLC method, probabilistic classifier chains [3] appear to be specifically appealing in this context, as it yields probabilistic predictions.

Requirements: Formalization of MLC with partial abstention; development of learning methods for this setting; implementation and empirical evaluation of these methods.

Prerequisites: Background in machine learning; programming skills.

For more information, please contact Prof. Eyke Hüllermeier (eyke@upb.de).

References

