

# How to measure uncertainty in Uncertainty Sampling for Active Learning

Eyke Hüllermeier

University Potsdam, Germany  
eyke@uni-paderborn.de

**Abstract.** Various strategies for active learning have been proposed in the machine learning literature. In uncertainty sampling, which is among the most popular approaches, the active learner sequentially queries the label of those instances for which its current prediction is maximally uncertain. The predictions as well as the measures used to quantify the degree of uncertainty, such as entropy, are traditionally of a probabilistic nature. Yet, alternative approaches to capturing uncertainty in machine learning, alongside with corresponding uncertainty measures, have been proposed in recent years. In particular, some of these measures seek to distinguish different sources and to separate different types of uncertainty, such as the reducible (epistemic) and the irreducible (aleatoric) part of the total uncertainty in a prediction. This talk elaborates on the usefulness of such measures for uncertainty sampling and compares their performance in active learning. To this end, uncertainty sampling is instantiated with different measures, the properties of the sampling strategies thus obtained are analyzed and compared in an experimental study.