Reasoning about the Learning Process

João Gama
LIAAD-INESC Porto, FEP-University of Porto,
Rua de Ceuta 118, Porto 4050-190, Portugal
jgama@fep.up.pt

Abstract. Data Mining is faced with new challenges. In emerging applications (like financial data, traffic TCP/IP, sensor networks, etc) data continuously flow eventually at high speed. The processes generating data evolve over time, and the concepts we are learning change. Evolving data requires that learning algorithms must be able to monitor the learning process and the ability of predictive self-diagnosis. A significant and useful characteristic is diagnostics - not only after failure has occurred, but also predictive (before failure). These aspects require monitoring the evolution of the learning process, taking into account the available resources, and the ability of reasoning and learning about it.

In this work we present a one-pass classification algorithm able to detect and react to changes in the process that generates data. Instead of forgetting the previous learned models, they are stored for possible re-use if the context where they were learned re-appears. The system uses meta-learning techniques that characterize the domain of applicability of previous learned models. The meta-learner can detect re-occurrence of contexts and take pro-active actions by activating previous learned models. The main benefit of this approach is that the proposed meta-learner is capable of selecting similar historical concept, if there is one, without the knowledge of the true classes of examples.

Data streams; concept drift; meta-learning; recurrent concepts.
Join work with Petr Kosina.