Reasoning about the learning process

João Gama
LIACC, University of Porto Portugal
jgama@fep.up.pt
http://www.liaad.up.pt/∼jgama/

Summary

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. In this talk we present a one-pass classification algorithm able to detect and react to changes. We present a framework that identify contexts using drift detection, characterize contexts using meta-learning, and select the most appropriate base model for the incoming data using unlabeled examples. 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.

Bibliography

João Gama is a researcher at LIAAD, the Laboratory of Artificial Intelligence and Decision Support of the University of Porto, working at the Machine Learning group. His main research interest is in Learning from Data Streams. He has published several articles in change detection, learning decision trees from data streams, hierarchical Clustering from streams, etc. Editor of special issues on Data Streams in Intelligent Data Analysis, J. Universal Computer Science, and New Generation Computing. He Co-chair the track on Data Streams in ACM Symposium in Applied Computing, a series of Workshops on Knowledge Discovery in Data Streams, and Knowledge Discovery from Sensor Data with ACM SIGKDD. He is the author of a recent book on Knowledge Discovery from Data Streams.