

# Challenges in Learning from Streaming Data

João Gama<sup>1,2</sup>

<sup>1</sup> LIAAD-INESC TEC, University of Porto

<sup>2</sup> Faculty of Economics, University Porto  
jgama@fep.up.pt

**Abstract.** Nowadays, there are applications in which the data are modeled best not as persistent tables, but rather as transient data streams. In this article, we discuss the limitations of current machine learning and data mining algorithms. We discuss the fundamental issues in learning in dynamic environments like continuously maintain learning models that evolve over time, learning and forgetting, concept drift and change detection. Data streams produce a huge amount of data that introduce new constraints in the design of learning algorithms: limited computational resources in terms of memory, cpu power, and communication bandwidth. We present some illustrative algorithms, designed to taking these constraints into account, for decision-tree learning, hierarchical clustering and frequent pattern mining. We identify the main issues and current challenges that emerge in learning from data streams that open research lines for further developments.