

Exploration and Exploitation in Online Learning

Peter Auer

Chair for Information Technology, University of Leoben, Austria
auer@unileoben.ac.at

Online learning does not distinguish between a training and an evaluation phase of learning, but considers learning as an ongoing process, such that learning algorithms need to perform and make predictions while they learn. After reviewing the online learning model and some algorithms, I will consider variants of the model where only partial information is revealed to the learner, in particular the bandit problem and reinforcement learning. The uncertainty of the learner caused by receiving only partial information, leads to an exploration-exploitation dilemma: is further information needed, or can the available information already be exploited? I will discuss how optimism in the face of uncertainty can address this dilemma in many cases.