Move your lamp post: Recent data reflects learner knowledge better than older data

April Galyardt
University of Georgia
galyardt@uga.edu

Ilya Goldin
Pearson
ilya.goldin@pearson.com

ABSTRACT

In educational technology and learning sciences, there are multiple uses for a predictive model of whether a student will perform a task correctly or not. For example, an intelligent tutoring system may use such a model to estimate whether or not a student has mastered a skill. We analyze the significance of data recency in making such predictions, i.e., asking whether relatively more recent observations of a student's performance matter more than relatively older observations. We investigate several representations of recency, such as the count of prior practice in the AFM model, and the proportion of recent successes with exponential and box kernels. We find that an exponential decay of a proportion of successes provides the summary of recent practice with the highest predictive accuracy over alternative models. As a secondary contribution, we develop a new logistic regression model, Recent-Performance Factors Analysis, that leverages this representation of recent performance, and has higher predictive accuracy than existing logistic regression models.