Data Mining: Learning From Large Data Sets

Fall Semester 2015

{usamuel, adavidso, kurmannn}@student.ethz.ch

December 23, 2015

Extracting Representative Elements

Problem formulation.

Approach and Results. We first implemented LinUCB, which turned out to be too slow when run on the cluster, we tried UCB1, ignoring all features of users and data. While this ran faster, the score had severely deteriorated, so we switched back to using LinUCB. Using the Sherman-Morrison formula to prevent re-computing inverses, we managed to get LinUCB to terminate in time, which resulted in a score of 0.06567. attempts to improve upon this result using a hybrid approach, random feature transformation proved both fruitless.

Workload distribution. Samuel wrote the implementation of LinUCB and UCB1. ALexander improved it with the Sherman-Morrison formula. Alexander tried out Thompson sampling.