Seeding Course Forums with the Teacher-in-the-Loop

Einat Shusterman[1], Hyunsoo Gloria Kim[2], Marc Facciotti[2], Michele Igo[2], Kamali Sripathi, David Karger[3], Avi Segal[1], & ***Kobi Gal***[1]
[1] *Ben Gurion University at the Negev, Israel*
[2] *University of California, David, USA*
[3] *Massachusetts Institute of Technology, USA*

Online forums are an integral part of modern day courses, but motivating students to participate in educationally beneficial discussions can be challenging. Our proposed solution is to initialize (or "seed") a new course forum with comments from past instances of the same course that are intended to trigger discussion that is beneficial to learning. In this work, we develop methods for selecting high-quality seeds and evaluate their impact over one course instance of a 186-student Biology class. We designed a scale for measuring the "seeding suitability" score of a given thread (an opening comment and its ensuing discussion). We then constructed a supervised machine learning (ML) model for predicting the seeding suitability score of a given thread. This model was evaluated in two ways. First, by comparing its performance to the expert opinion of the course instructors on test/holdout data. Second, by embedding it in a live course, where it was actively used to facilitate seeding by the course instructors. For each reading assignment in the course, we presented a ranked list of seeding recommendations to the course instructors, who could review the list and filter out seeds with inconsistent or malformed content. We then ran a randomized controlled study, in which one group of students was shown seeds that were recommended by the ML model, and another group was shown seeds that were recommended by an alternative model that ranked seeds purely by the length of discussion that was generated in previous course instances. We found that the group of students that received posts from either seeding model generated more discussion than a control group in the course that did not get seeded posts. Furthermore, students who received seeds selected by the ML-based model showed higher levels of engagement, as well as greater learning gains, than those who received seeds ranked by length of discussion.