Identifying Struggling Students through a Clicker Data Application

Agnes Garcia, Tianheng Ma, & Joyce Yue
Advised by: Professor Leo Porter, Professor Bill Griswold, & Soohyun Nam Liao

Motivation
- Predicting student performance is important since it helps instructors to target student learning
- Researchers have created a lightweight machine learning model that can predict student performance with 70% accuracy using 3 weeks of clicker responses data
- Currently, such a model only exists in theory; there is a lack of research on applying these prediction models to real-life academic settings

Proposed Solution
We developed a computer application integrating Liao et al.'s model for computer science instructors to identify struggling students early in the term.

Application Workflow

User Interface Design

Step 1: User uploads XML files from a past term and current term with students' ID, final exam scores, clicker responses and the clicker question images.

Step 2: The system pre-fills the classification for each clicker question: individual, group, or ignore, and the user can change the classification if necessary.

Step 3: For the previous and current terms, the user selects the correct answer(s) for each clicker question that were classified as individual.

Step 4: For each past term clicker question, the system automatically finds a list of similar current term questions. The user can select the matching question from the given list.

Step 5: The user clicks "Process" to process the data and for the system to generate the predictions.

Step 6: The user sees the results that show each current student's probability of being in the bottom 40%.

User Feedback

User Study Round 1:
- User Study Round 1: Prof. Griswold & Soohyun Nam Liao
- Conducted screen and voice recordings during the study

Findings
- Total time of study: 3 hours
- Step 2 took up the most time and was tedious (shown below)
- Users needed instructions for each step

Changes Made
- Changed the workflow of Step 2: virtual scroll and pre-filled classification of questions as alternating individual and group
- Added progress indicators to Steps 2, 3, and 4
- Added instructions in the navigation page before each step and on the step's page itself

Future Work
- We plan to deploy this application to instructors in computer science and in other disciplines after conducting more user studies and determining if the workflow needs to be modified
- We plan to study whether instructor interventions after using this application are effective in improving students' final exam scores and grades in the course

Acknowledgements
We would like to thank Professor Christine Alvarado and Vignesh Gokul of the ERSP program for their mentorship. We also would like to thank Professors Leo Porter and Bill Griswold, and their team: Soohyun Liao, Ravi Sheth, and Xingyu Yang, for their support and guidance on this project.

Liao’s logistic regression model* will learn the relationships between the previous term's student’s final scores and clicker question performance, and use that to predict the current term student final performance.