Abstract

MOOCs will come with an integrated course chat, for a more integrated cohesive system where students may post questions in real time (as people sometimes do in twitch or youtube live), and they are linked to the video lecture's timestamp. This can help them communicate with other students, e.g. as a student, seeing a classmate's emoji could help them feel less isolated, feel understood and optionally reading through students' concerns about the material is enlightening, without having to open another tool. As a teacher, this can provide valuable feedback to improve. This project will analyze real-time video-synced students' comments on any given lecture to provide insight on which phrases are confusing, helpful or make students feel in any given way to incorporate into the current communication or future lectures.

Project Proposal questions

1. **What are the names and NetIDs of all your team members? Who is the captain? The captain will have more administrative duties than team members.**
   Rodrigo Mendoza rm36@illinois.edu Only me
2. **What topic have you chosen? Why is it a problem? How does it relate to the theme and to the class?**
   Intelligent Learning Platform - Student Sentiment Analysis
   I'll implement a system that will let students comment at any time in the video and a system which analyzes the students' comments in real-time to provide sentiment analysis. I'll use word-based robust sentiment analysis tools for the analysis, and a web-based portal with the video playing for students at their own pace.
3. **Briefly describe any datasets, algorithms or techniques you plan to use**
   I'll use sentiment analysis datasets, which have positive or negative reviews associated with them and will use word-based sentiment analysis tools for analysis, then display the aggregated predicted strength of each comment to the teacher. I'm also planning to provide a way to get my own dataset info.
4. **How will you demonstrate that your approach will work as expected? Which programming language do you plan to use?**
   I will evaluate the system in two ways: 1. using a test portion of the dataset not used in training to measure how close the system predicts the comment individually, and 2. Testing the aggregated comments by doing a mix of different comments, all positive, all negative and half and half.
   I will use mostly python.
5. **Please justify that the workload of your topic is at least 20*N hours, N being the total number of students in your team. You may list the main tasks to be completed, and the estimated time cost for each task.**
   Researching and parsing dataset - 2 hours
   Implementing a sentiment analysis algorithm - 6 hours
   Training, iterating and fine-tuning - 4 hours
   Creating a website with a video playback and student commenting tool - 8 hours
   Integrating the algorithm in the website - 3 hours
   Creating the presentation slides and demo - 2 hours