Just-in-Time Intervention

Hosuk Choi, Jianan Jin, Anusha Prasad, Manley Roberts, Ania Thomas, Matt Yang

Reminder of Motivation

- Primarily goal is to expand upon analysis done last semester.
  - Gaining more semesters of data
  - Adjusting timeframe of each snippet of data to be cumulative
  - Gaining more attributes
    - Selecting effective attributes
  - Improving predictions

Current Status Check

- Developing familiarity with tools
- Collecting data
- Transforming data
- Building a dataframe

MongoDB Subteam

- Members : Hosuk Choi, Matt Yang, and (kind of) Manley Roberts
- Will be examining clickstream and forum data held in a Mongo Database, which is a NoSQL database
  - Instead of tables and rows, Mongo is a nonrelational database and is made up of collections and documents
  - A document in MongoDB is similar to JSON objects
- Current Activity
  - Access MongoDB from Docker Container Image
  - Writing queries in Mongo to select attributes we need (next page)

MongoDB Data Sources

Clickstream

(Features that quantitate and count user interactions)

- Clicks
- Video plays
- Active days
- Edx.done.toggled
  - Says whether the selected event was completed by a certain student
- Navigation based (events that communicate navigation to a new page)
  - Next
  - Previous
  - Tab
  - Hyperlink

Cannot do forum data since we cannot attach it to a student

SQL Subteam

- Members: Ania Thomas and Anusha Prasad
- Identifying useful attributes from the Gardner & Brooks paper
  - Average Raw Score
  - Raw Points Per Submission
  - Correct Submissions Percent
  - Change In Weekly Average, etc.
- Gathering information from our previous SQL queries
  - Expanding the scope of the previous queries to support multiple semesters
  - Once completed, we will construct a Pandas dataframe
- Working to aggregate specific data for students on a week to week basis
  - Developing a way to check grade difference week to week

SQL Data Sources

Gardner and Brooks

- Attributes discussed in “Evaluating Predictive Models of Student Success: Closing the Methodological Gap” written by Gardner and Brooks
    - (Average Number of Quizzes, Total number of interactions, number of sessions, etc.)

GitHub and Previous Queries

- Previous exams
- Previous homeworks
- Previous quiz submissions
- Used cohort data to create table of active student days
- General queries (user and course IDs, cohort membership, course enrollment, active logins, etc.)
- Select Event queries (based on user and course IDs with weeks in semester)

Next Steps

- Merge data sources into common dataframe
- Create more complex attributes
- Begin developing models and evaluating attribute success