Milestone 3 Progress Evaluation

Academic Behavior Reccomendation System

Shreyas Ugemuge `sugemuge2014@my.fit.edu`

Yaqeen AlKathiri `yalkathiri2013@my.fit.edu`

Mohammed AlHabsi `malhabsi2013@my.fit.edu`

Shiru Hou `shou2015@my.fit.edu`

Faculty Sponsor: Dr. Phillip Chan `pkc@cs.fit.edu`

April 24, 2017

Contents

1 Progress of current Milestone

Task	Completed	Shreyas	Yaqeen	Shiru	Mohammed	Remarks
Finish Extracting Behaviors	100%	25%	25%	25%	25%	N/A
Program to Merge and Format Data for Data mining	95%	100%	-	-	-	File names are hard-coded at this point. Need to write a shell script to run program.
Compile Information	100%	25%	25%	25%	25%	N/A
Change Behavior Exrtaction to weekly	100%	25%	25%	25%	25%	N/A
Progress Evaluation and Presentation	100%	25%	25%	25%	25%	N/A

1.1 Discussion of each task

1.1.1 Finish Extracting Behaviors

When each different item (like quizzes and tests) are considered separately the team implemented 25+ behaviors. However, for the sake of simplicity and to get a better understanding of how much data was extracted several behaviors will be grouped within this document. For instance, seperate behaviors like time taken for test 1 and time taken for test 7 will be grouped as time taken for tests. Here are all the behaviors that have been implemented

1. Total number of Logins

2. Average activity per session

3. Number of days where student did not login

4. Time taken for each test and average time taken to do each test

5. Number of study Guide materials accessed

6. Number of ASRs submitted

7. Number of Procedure quizzes submitted

8. Number of Cyberrat assignments submitted

9. Number of surveys done

10. Activity count for reviewed material

11. Number of online meetings attended

12. Number of optional assignments submitted

13. Number of activities for accessing supplemental material

14. Average Number of activities before each test

15. Number of tests A and Tests B submitted

16. Days between when the material was available and when it was submitted

1.1.2 Program to merge and format data

Program was used to merge all the csv files, using java and openCSV. The program uses the ID as a primary key and merges data over multiple csv files into src/merged.csv. The limitation is that the filenames have to be hardcoded and there is only one kind of output format.
A solution to tell the program of the expected format has to be implemented.

1.1.3 Compile Information

the grades.csv and all behaviors located in src/docs/ are used to create a final compiled file Behaviors_merged.csv .

1.1.4 Change behavior extraction with the number of weeks as parameter

4 of the behaviors have been modified to include this functionality. The same program was moved from python to java because pandas library abstracts much of functionality and doesn't provide a way to iteratively go over the table.

1.2 Discussion of team member contribution

1.2.1 Shreyas

Reimplemented 4 behaviors in java instead of python. Decided format for individual csv files. Wrote merge program. Modified bahaviors to extract weekly information.

- Total number of logins counted the amount of times the student logged in over given weeks. Abandoned sessions. i.e. sessions with a login but no logout were handled

- Average activity per session. A session is time between two logins or between a login and a logout, where each row in the log entry is an activity.

- Time taken for each test. The start time is marked by the 'delivered' activity and the end time is marked by 'submitted'. The result is output in minutes as this was the most readable metric for the available significant figures.

- Number of days where the student did not login.

1.2.2 Shiru

Finished 4 implementations:

- Time between when unit was available and ASR submission Calculate how many days of each student submitted the ASR after each unit materials were available. Use Python pandas and numpy library to extract data from csv file and calculate it. Then write the result to a new csv file.

- Study guide video or podcast and guide materials Calculate how many activities of each student accessed to study guide video or podcast and guide materials. G1, G2... means how many activities of student access each guide materials.

- Total number of each student reviewed activities Use the search key Reviewed in the raw data from students log file, calculate how many activities of each student do reviewed actives include the reviewed assignments and quiz after they are graded.

- Total number of each student study supplemental materials activities

1.2.3 Yaqeen

- Implemented and Extracted 7 behaviors:

  - Number of ASR Unit Assignments Each Student Submitted: Located all ASR Units in the log files and counted the number of submissions for each one.
  - Number of Optional CyberRat Assignment Submitted: Followed the same concept from the first behavior to extract the OCA. In this behavior, the only difficulty was that only about %5 of all the students acutely attempted to do the OCA. Where many of them didnt even click it.
  - Number of Tests B Submitted: Located all tests B that the students submitted and counted them.
  - Number of Late ASR Unit Assignments Submission:
    Located ASR Units and submission dates. Compared the submission dates of ASR Unit with the deadline for each Unit. Then calculated the total number of ASR Units submitted after the deadline.

- Duration Time for Final A, ASR Units, and Surveys:
  Calculating the time from the starting of the ASR/Survey until its submitted in minutes.
  The start time is Delivered or Resumed depending on if the student stopped and them resumed. Submitted is the end time.

- Time between Tests B submissions and availability:
  How early each student took the test after the test was available in days. Located submission date and compared it to the date when the test was available. The comparison was done by taking the difference between the submission date and the date of test availability.

- Time between Surveys submissions and availability:
  Followed the same concept from the above behavior to extract how early the surveys were submitted after it was available.

1.2.4   Mohammed

Conceptualized 7 new behaviors, implemented and extracted 5 of them. Used Java as main language for the behavior implementation. Used opencsv library to read and extract the information from the students log files, and write csv output for each behavior separately. Assigned to contribute in implementing the GUI. Contributed in the progress evaluation document. Behaviors Implemented:

- Online meetings attended:
  We can see from the syllabus that attending online meetings is a mandatory part of the course. From the students log files, the program goes through the row of the activities description and simply calculate the number of Online Meeting boxes and writes the results. Following are some screenshots of the activity description and output.

- Number of procedure quizzes:
  This is one of the optional activities for bonus points in this course. The program reads the log files, and calculates how many optional procedures quiz submitted and outputs the extraction to a csv file for all students.

- Number of surveys:
  This is another optional activity for bonus points in this course. The program reads the log files, and calculates how many optional surveys submitted and outputs the extraction to a csv file for all students.

- Number of CyberRat Assignments

- Average activities for each test:
  Here the program calculates the total number of activities before each test, and then calculates the average for the 7 tests in this course, then it outputs the extraction to a csv file for all students. This helps us know how much did the student access the materials for exams, as exams are worth most portion of the final scores.

- How early does the student take the exams:
  For each of the seven exams, the program calculates the number of days between when the exam was available and when did the student submit it.

- Changed behaviors for weekly extraction

2   Plan for the next milestone

Task	Shreyas	Yaqeen	Shiru	Mohammed
Make All behavior extraction weekly	-	33	33	33
Use decided data mining tools and packages to correlate behaviors with grades	25	25	25	25
Design and Implement GUI	-	50	-	50
Test correlations using the Train/Test Model, dividing the semester into 2 parts.	25	25	25	25

3 Sponsor feedback on each task for current milestone

4 Sponsor Evaluation

Sponsor: Please detach this page and return to Dr. Shoaff

Score (0-10) for each member: circle a score (or circle two adjacent scores for .25 or write down a real number between 0 and 10)

Shreyas Ugemuge	0	1	2	3	4	5	6	6.5	7	7.5	8	8.5	9	9.5	10
Yaqeen AlKathiri	0	1	2	3	4	5	6	6.5	7	7.5	8	8.5	9	9.5	10
Mohammed AlHabsi	0	1	2	3	4	5	6	6.5	7	7.5	8	8.5	9	9.5	10
Shiru Hou	0	1	2	3	4	5	6	6.5	7	7.5	8	8.5	9	9.5	10

Faculty Sponsor

_____

Signature

_____

Date