**Documentation (TestBed2)**

**Job recommendation system**

**Drexel University**

**X (input)**

Matrix = Jobs (Tasks) X features (Metrics)

This file is imported from the Metric Task generator system.

**Y (input)**

Matrix = Jobs X users

This file is imported from the Metric Task generator system.

**R (input)**

Matrix = Jobs X user

* The R should just show 1 or 0 (binary approach) according with Y. It means that if Y has some measured job, the R will be 1 in that cell. 0 Otherwise.
* This file is imported from the Metric Task generator system.

**Job\_names (input)**

* Just a column vector with the job’s name.
* This file is imported from the Metric Task generator system.

**User\_table (input)**

* Name for each user.

**Results (output)**

* Column 1 is a Job from the TOP 10(or less if TOP 10 is not applicable) jobs compared with each job (column 3) that has a similarity equal or greater than 70%.
* Column 2 is the predicted rate for our collaborative filtering for recommended jobs.
* Column 3 is the job that is being compared with the job at column 1.
* Column 4 is the original rating by employer for each job in column 4.
* Column 6 is the similarity in percentage. Just for jobs with similarity equal or greater than 70%.
* Column 7 is the user number (id).

**Average (output)**

* Here we have the top 10 jobs (if applicable) with an average between the job and the others jobs with similarity greater than 70%
* Column 1 is the job name from the TOP 10 jobs (or less if TOP 10 is not applicable).
* Column 2 is the predicted rating for the recommended job.
* Column 3 is the average of the ratings among the TOP 10 jobs which have similarity equal or greater than 70% to the recommended job.
* Column 4 is the average of the percentage among the TOP 10 jobs which have similarity equal or greater than 70% to the recommended job.
* RATING AVGS TOTAL row shows the total average from the similarities from the top 10 jobs for the actual user.
* PERCENTAGE AVGS TOTAL row shows the average from all the similarities.

\* In the last two rows, we have the total system statistics for all users in the dataset.

**IDandAVG (output)**

Here we write the users ID and their calculated rating.

Column 1 is the user name.

Column 2 is the calculated rating by our software.

**\*Output Jobs (output-folder)**

This folder contains sets of the output data divided by 30 jobs in different folders. Inside of each folder are all the 3 output files of the recommended system.

**\*Output Workers (output-folder)**

This folder contains sets of the output data divided by 100 users in different folders. Inside of each folder are all the 3 output files of the recommended system.

\*These folders are not generated automatically. What we did was a copy and paste of every running.