**UStudy**

**A complete educational portal**

**Arapally Ebenezer Anand**

**Bandaru Sai Kishore**

**Kommineni Sivakrishna**

**Ponnam Balakrishna**

**Summary:**

To brief out the implementation done in the prior increments we have completed the collection of data, designing of the gui pages and solr connectivity. In this increment we have made considerable progress in data mining and design of front end. The major completion is made in the solr connectivity of clustering of data. The data has been pruned and refined so as to fit the application and is being pushed into the solr. The retrieval of data from solr is the major issue which we are facing currently. The gui pages for the front end part are also been completed and the user validations were also made. The clustering of data is done on the college data and the recommendation are being done on the user profile.

**Framework Specifications:**

In this increment we have three major tasks which are being focused. The first task is to mine the data that is being collected. The second task is to design the various layouts for client interface and user validations. The third task is to pushing and retrieving data from the solr i.e. the repository which we are using for storing the data that is being mined. The data that is being collected has been successfully clustered. In the initial phases we focused on classifying the data based on the gre, toefl, tuition fee etc. But however due to the fuzziness in the data that is being collected we couldn’t classify the data and also apply a suitable algorithm in mahout. So we have used R for resolving this issue. The gui part has also been completed with user validations. The solr issue for pushing data has been completed but certain issues were being found while retrieving from solr in android. The further work is being continued on content recommendation in the application for registered users.

**Application Specification:**

* **Software Specification**
  + Tools: R, Solr, Android Development Kit
  + Operating System: Android
  + Development Operating System: Windows 8
  + Programming Language: Java 7.0
  + Databases: Oracle, Hadoop

**Implementation:**

The implementation part typically consisted of the following steps:

* Pruning and refining of data
* Clustering the data in R
* Pushing the data to R
* Parsing of output achieved from clustering into json format which has to be pushed to solr
* Designing of GUI and user validation.

The data that has been collected is pruned and refined into desired format for usage in clustering. Initially we couldn’t classify the data using Mahout so we found a solution for clustering the data using R. We used K means clustering for data clustering and were able to cluster the data various dimensions i.e. using gre, toefl scores, fees, aid, number of Asians, number of Europeans and various other dimensions. Currently we were able to successfully cluster data with three dimensions they are gre, toefl, fees.

Once the output is achieved from R we have stored it in a csv file and that has been stored on the disk. The output csv is being parsed into the json format. For this we have written a code for these parsing and been uploaded to github. The parsed json file has been pushed into the solr using the curl scripts. Then the pushed data is being retrieved through Android application. Currently we are facing while retrieving the data from solr which can be resolved in couple of days. Another point is we were able to retrieve data using a JavaEE project in eclipse but not with android.

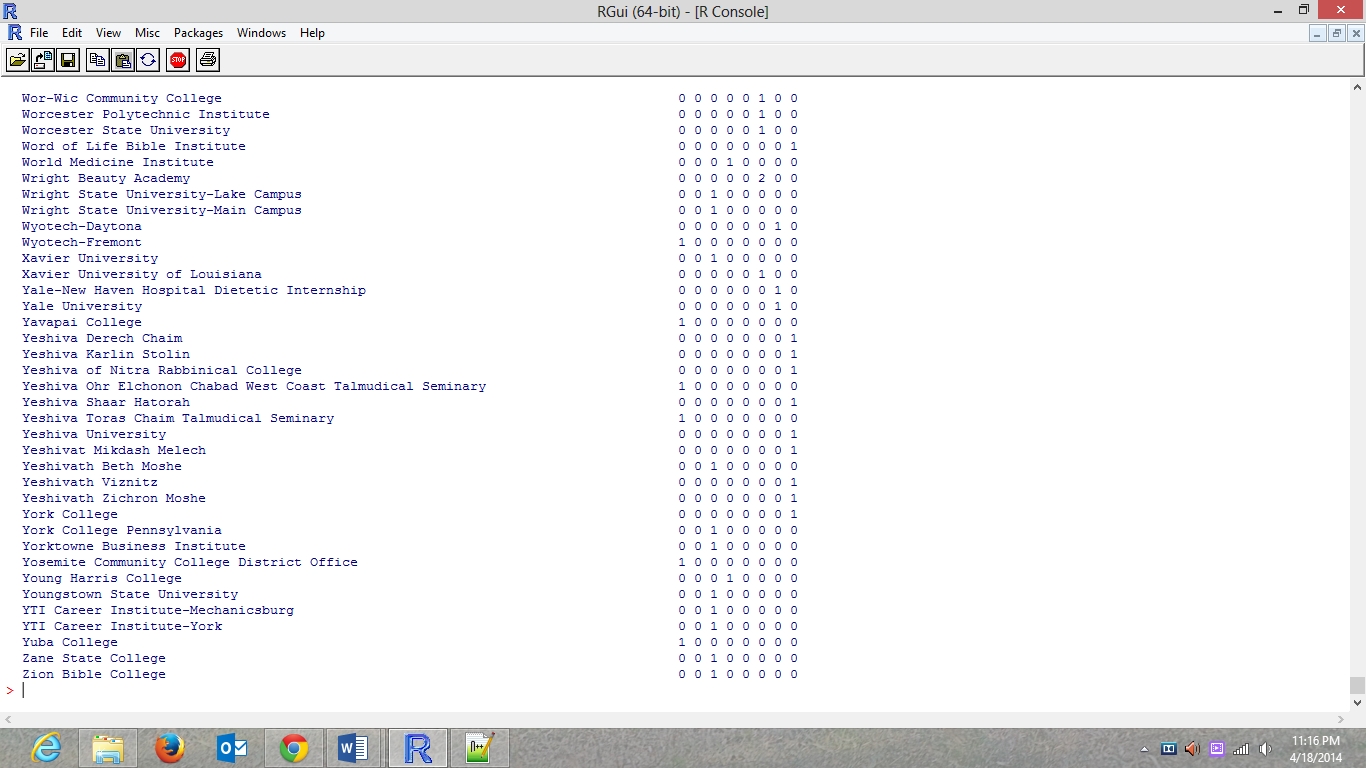


Figure R output after k means clustering with four dimensions

The above figure describes the output that is formed after using the k-means clustering on the available data.

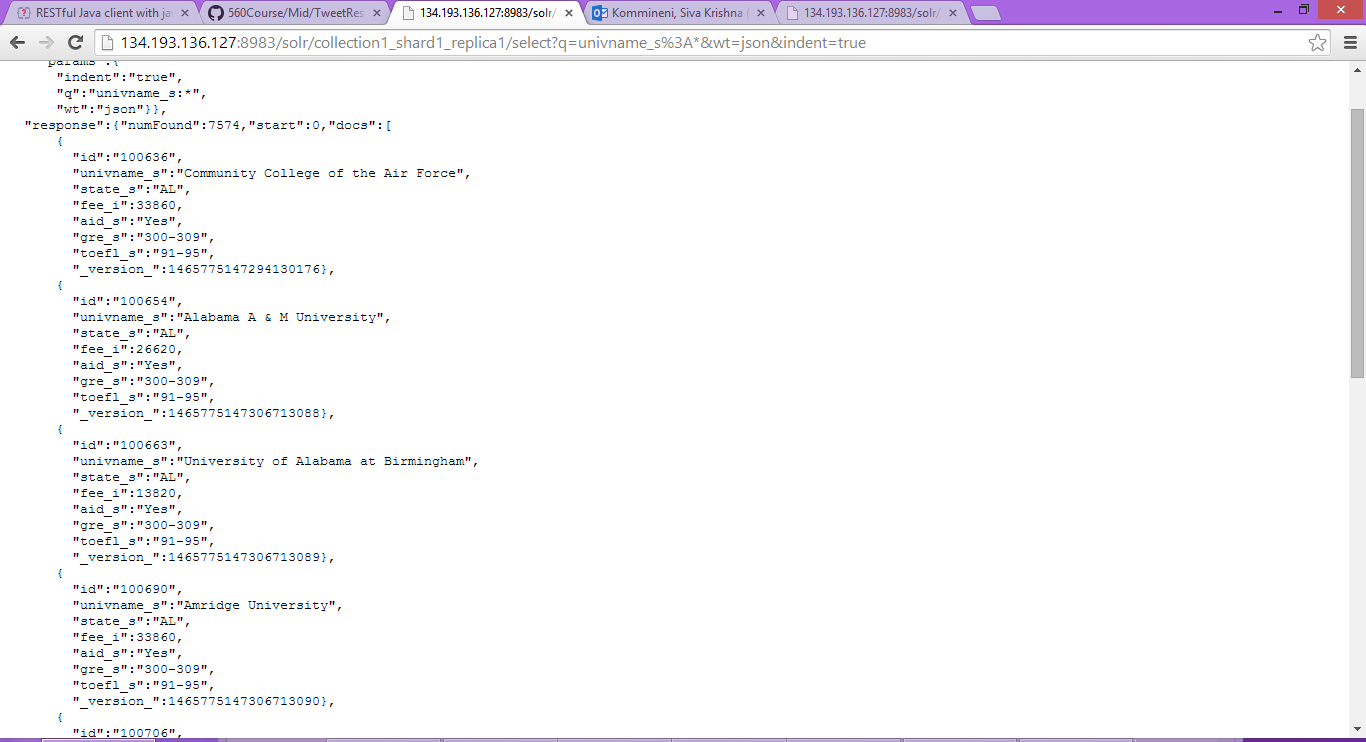


Figure Solr data

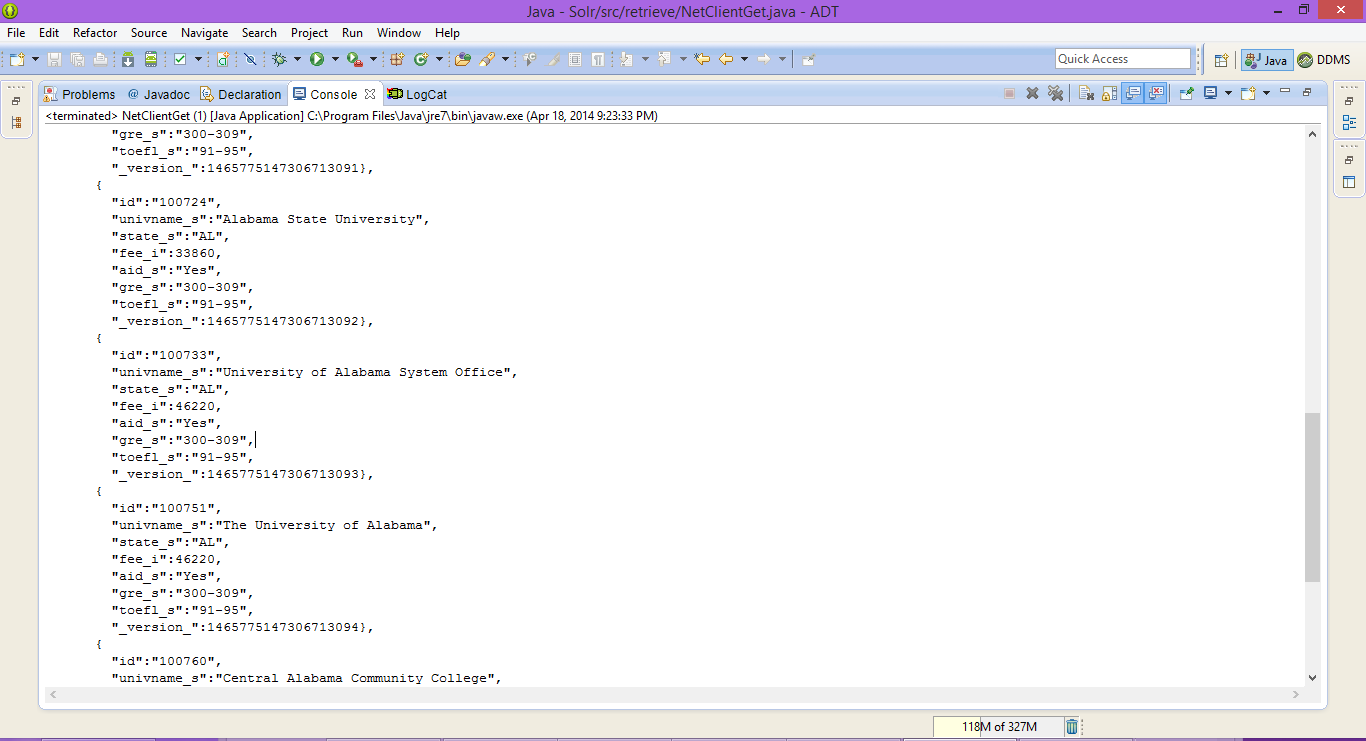


Figure Solr data visible in Java EE eclipse console

The above is the solr data that is being pushed after parsing of the output obtained from k means clustering in R.

The below snapshots shows the user GUI pages for logging, specifying the requirements and user output page.

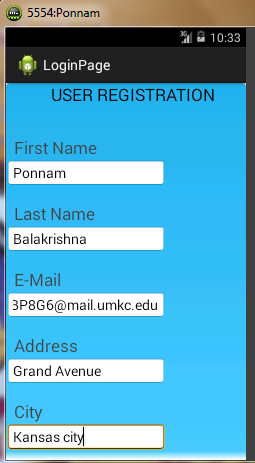


Figure User registration pages

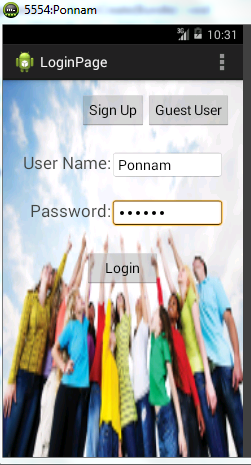


Figure Login page

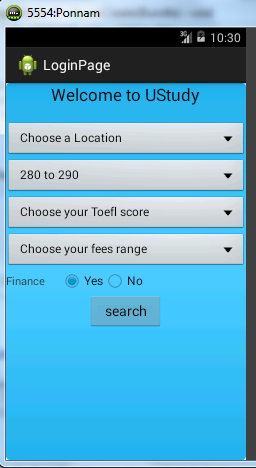


Figure Specifying the requirements

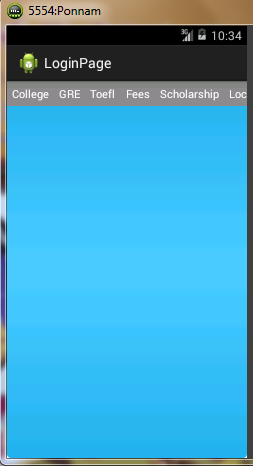


Figure Guest and user result page

**Project Management:**

All the tasks and their day today increments are being updated in the scrumdo.

* <https://www.scrumdo.com/organization/university-of-missouri-kansas-city5/dashboard>

**Third Increment:**

The tasks that will be included for third increment are:

1. Completion of the entire client web pages – Ponnam Balakrishna
2. Registration of user profiles onto NoSQL store and analyzing those data – Ebenezer Anand Arapally
3. Data Clustering using K-means clustering in R – Sai Kishore Bandaru, Ebenezer Anand Arapally
4. Retrieving the output from Solr based on the inputs specified by user and displaying on the result page in android – Kommineni Sivakrishna
5. Final Testing ­– Entire team
6. Recommendations using content personalization – Sai Kishore Bandaru.

The above mentioned tasks will be uploaded the scrumdo tool with specified timelines.