

The main assignment and bonus part is done in Scala 2.10.2 and Spark 2.0.1. Following are details on how to execute the program

1. Main assignment

- a. **Community.Scala**

Sample command to execute

```
bin\spark-submit --class Community Community.jar ratings.csv  
output\community.txt
```

You can specify any output path and the output will be written to that.
Program takes around 10 minutes to complete

- b. **Betweenness.Scala**

```
bin\spark-submit --class Betweenness Betweenness.jar ratings.csv  
output\betweenness.txt
```

Program takes around 2 to 3 minutes to complete.

For both tasks, the output files Vijayakumar_Gedigeri_communities.txt and Vijayakumar_Gedigeri_betweenness.txt are included in the submission directory

2. Bonus part: Using Sparkling Graph – **SparklingGraph.Scala**

For this task, external library ml.sparkling:sparkling-graph-operators_2.10:0.0.6 is loaded using Maven.

```
bin\spark-submit --class SparklingGraph SparklingGraph.jar ratings.csv  
output\spark_communities.txt
```

Sample output file Vijayakumar_Gedigeri_spark_communities.txt is also included in the submission.