

**Description:****Versions used:**

Spark- 2.2.1

Python- 2.7

- Please install “networkx” before running the following programs.

**1. Commands to execute the program: (Task-1) – (Betweenness):**

```
bin\spark-submit Prashanth_Manja_Betweenness.py < path_of_the_ratings.csv_file >
```

**Example:**

```
bin\spark-submit Prashanth_Manja_Betweenness.py \data\ratings.csv
```

Runtime: 75s

**2. Commands to execute the program: (Task-2) – (Community):**

```
bin\spark-submit Prashanth_Manja_Community.py < path_of_the_ratings.csv_file >
```

**Example:**

```
bin\spark-submit Prashanth_Manja_Community.py \data\ratings.csv
```

Runtime: 430s

**3.Bonus :****Python:**

```
bin\spark-submit Prashanth_Manja_Bonus_Community.py < path_of_the_ratings.csv_file >
```

**Example:**

```
bin\spark-submit Prashanth_Manja_Bonus_Community.py \data\ratings.csv
```

**Scala:**

**\*Please include the library “ml.sparkling.graph.operators”**

```
bin\spark-submit --class Prashanth_Manja_Bonus_Communities <jar-file-path> <inputpath-including 'ratings.csv'>
```