ReadMe

Unzip the delivered assignment submission folder

1. Create a maven project
2. Create a package with name mr.assignment4
3. Copy paste the provided class files
4. Update the pom.xml with the one provided
5. Add the provided the Makefile in the same location as pom.xml
6. Update the maven project
7. Add Scala Nature to the project
8. Change the compiler to Scala compiler
9. Create a directory by name input
10. Place all the required input files

To run locally:

Open the terminal from the folder containing Makefile

Run command – make alone

Final result with top 100 pages and their ranks, will be generated in the output folder

To run on AWS

Open the terminal from the folder containing Makefile

Create folder input in the AWS bucket specified in Makefile

Upload input files using the command - make upload-input-aws

Run Command – make cloud

Final result with top 100 pages and their ranks, will be generated in the output folder

Once the processing completes, output file can be downloaded using the command - make download-output-aws