**Articles Recommendation Engine**

**Execution Guide**

|  |  |
| --- | --- |
| **Version:** | **0.1** |
| **File Name:** | **Articles\_Recom\_Engine\_Exec\_Guide.docx** |
| **Scala Version:** | **2.10.5** |
| **SBT Version:** | **0.13.17** |
| **Spark Version** | **1.6.2** |
| **Date** | **01-APR-2018** |
| **Creator** | **Chiranjeevi Kolli** |

1. **CODE DEPLOYMENT – COMPLETE VERSION**

**Installation Prerequisites:**

Please make sure following softwares are installed in your machine.

1. JDK 1.8
2. Scala 2.10.5
3. Spark 1.6.2

**Execution Steps:**

1. Download the **DataEngineerHomeTest.tgz** file.
2. Extract the compressed file.
3. Find the **Articles\_Recom\_Engine\_Exec\_Guide.docx** documentunder **/ArticlesRecomEngine/Document/** directoryfor the steps to run the Article recommendation engine.
4. Copy below jar and dataset into your desired location

**/** **ArticlesRecomEngine/Deployment/teamrecomengine\_2.10-0.1.jar**

**/** **ArticlesRecomEngine/Deployment/ test-data-for-spark.json**

1. Login into windows/linux shell
2. Execute/run the Articles Recommendation Spark application (spark-submit) with 4 parameters/arguments apart from className and jar name separated by single space mentioned below.

Template:

spark-submit

--class <packagename>.<classname>

<absolute\_jar\_file\_name>

<spark\_execution\_mode>

<skuid>

<absolute\_input\_dataset\_name>

<absolute\_output\_file\_name>

|  |
| --- |
| For Example: spark-submit --class hackathon.Articles C:\Users\Chiru\IdeaProjects\RecomEngine\target\scala-2.10\recomengine\_2.10-0.1.jar local[\*] sku-402 C:\Users\Chiru\Desktop\Desktop\resource\test-data-for-spark.json C:\Users\Chiru\Desktop\Desktop\resource\results\output |

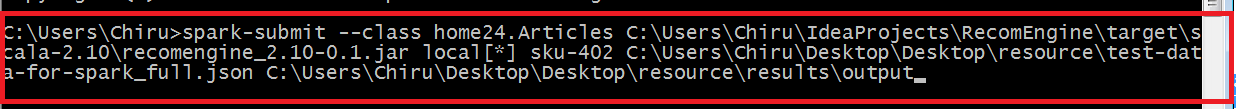
Argument 1: Pass the Spark Execution Mode (Example: local[\*])

Argument 2: Pass the SKU ID that you would like to identify similar SKU’s (Example: sku-777)

Argument 3: Pass the input dataset location that you placed in your machine (Step 3) (Example: /**ArticlesRecomEngine** /Deployment/ test-data-for-spark.json)

Argument 4: Pass the output resultant dataset location that you would like to generation after successful execution and finding the similar SKU’s on given/provide SKU ID(Example: /**ArticlesRecomEngine** /Deployment/Output/)

|  |  |  |
| --- | --- | --- |
| **S.No** | **command-line options** | **Examples** |
| 1 | packagename | hackathon |
| 2 | classname | Articles |
| 3 | absolute\_jar\_file\_name | C:\Users\Chiru\Desktop\Desktop\resource\recomengine\_2.10-0.1.jar |
| 4 | spark\_execution\_mode | local[\*] |
| 5 | skuid | Eg: sku-777 |
| 6 | absolute\_input\_dataset\_name | Eg: C:\Users\Chiru\Desktop\Desktop\resource\home24-test-data-for-spark.json |
| 7 | absolute\_output\_file\_name | Eg: C:\Users\Chiru\Desktop\Desktop\resource\results\output |



1. After successful execution, you can find the output/resultant dataset into mentioned location(Example: /**ArticlesRecomEngine**/Deployment/Output/)

