

Real-time Tweet Analysis

Project set up instructions:

1. Install Hadoop, HBase, Hive, Pig, Scala and Play framework
2. In project 'TweetStreamToHBase'
 - a. Go to Object 'HBaseUtils' and edit the path to hbase-site.xml of your own machine
 - b. Go to 'TwitterStreaming' class and edit the paths to paths compatible with your machine
3. In project 'BigData'
 - a. Go to class 'TweetUtils.java' and edit the path 'HBASE_CONF' to hbase-site.xml of your own machine
 - b. Go to 'TweetUtils.java' class and edit the path 'OUTPUT_PREFIX' to paths compatible with your machine
4. In project 'BigDataVisualization'
 - a. Go to 'routes' and edit '/Visualization/*file' path to the one specified in OUTPUT_PREFIX

Project is all set to be run.

Project run instructions:

1. Run project 'TweetStreamToHBase' on one port (for example, 8900). Below command can be used:
Command - ./activator "run 8900"
This will start the streaming where tweet go to HBase
2. Run each mapreduce class in project 'BigData'. These classes are in the package MapReduce
This will save the data to be visualized in the path you specified in OUTPUT_PREFIX
3. Run project 'BigDataVisualization' on another port (for example, 8901). Below command can be used:
Command - ./activator "run 8901"
This has all the analysis. Click on any analysis and you will see the visualization of the analysis.