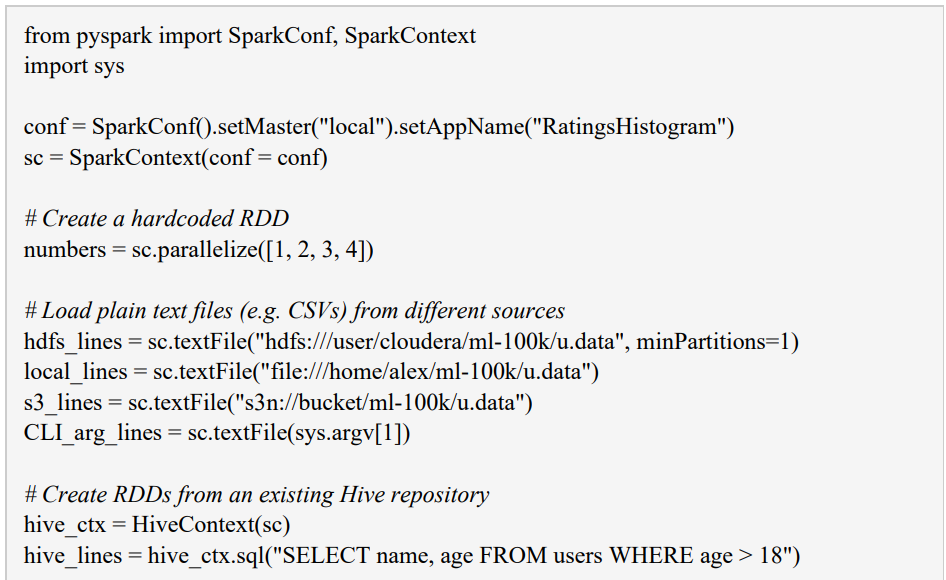
**PRACTICAL-11**

**AIM:**

To perform Sentiment Analysis using Twitter data, Scala and Spark.

**CODE:**

**Reading data:**

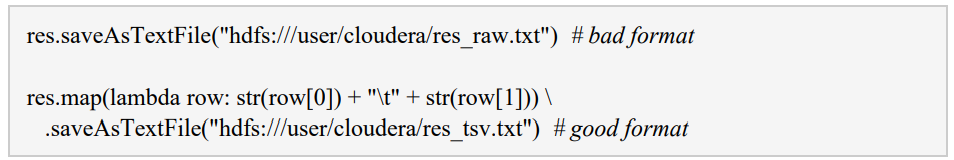
****

Notice that this can not be run with the standard Python interpreter. Instead, you use the spark-submit to submit it as a batch task, or name the Shell Pyspark

Writing of data:

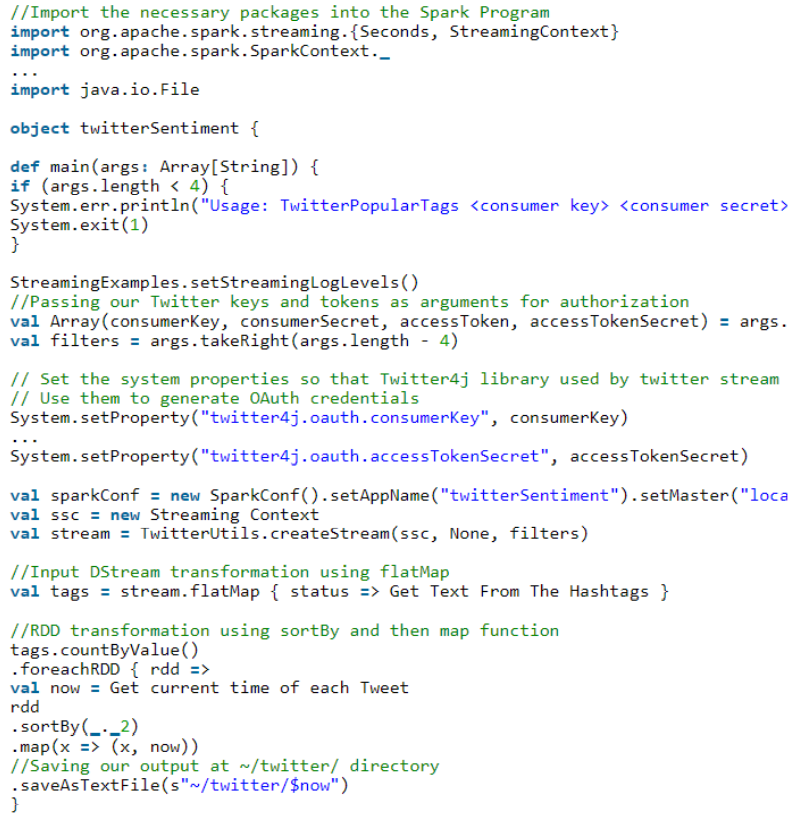
● The RDD class has the method saveAsTextFile. However, this saves the representation of each variable in a string. In Python, the resulting text file will contain lines such as (1949, 111).

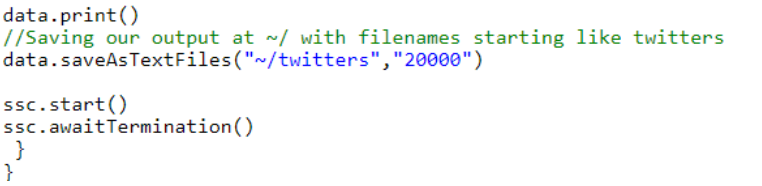
● If you want to save the data in CSV or TSV format, you can either use Python's StringIO and csv modules or simply map each variable (vector) into a single string, e.g. as follows:

****

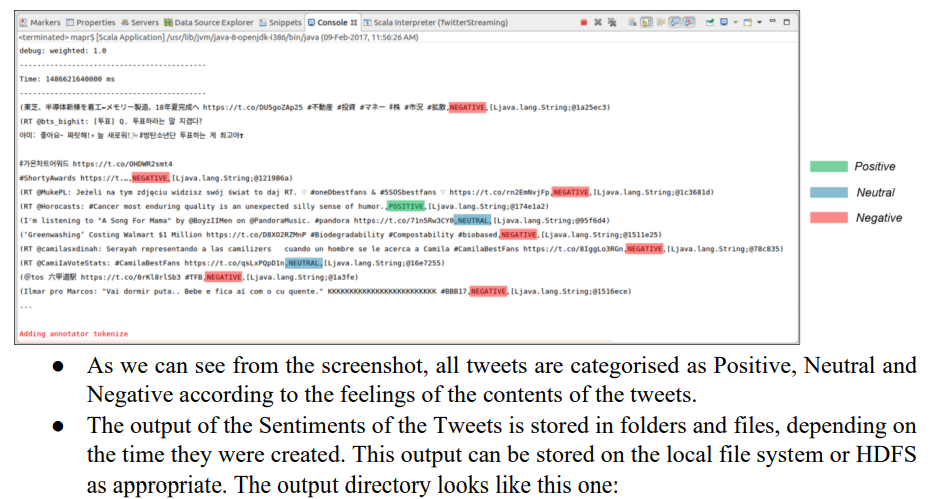
**Sentiment Analysis using Scala:**

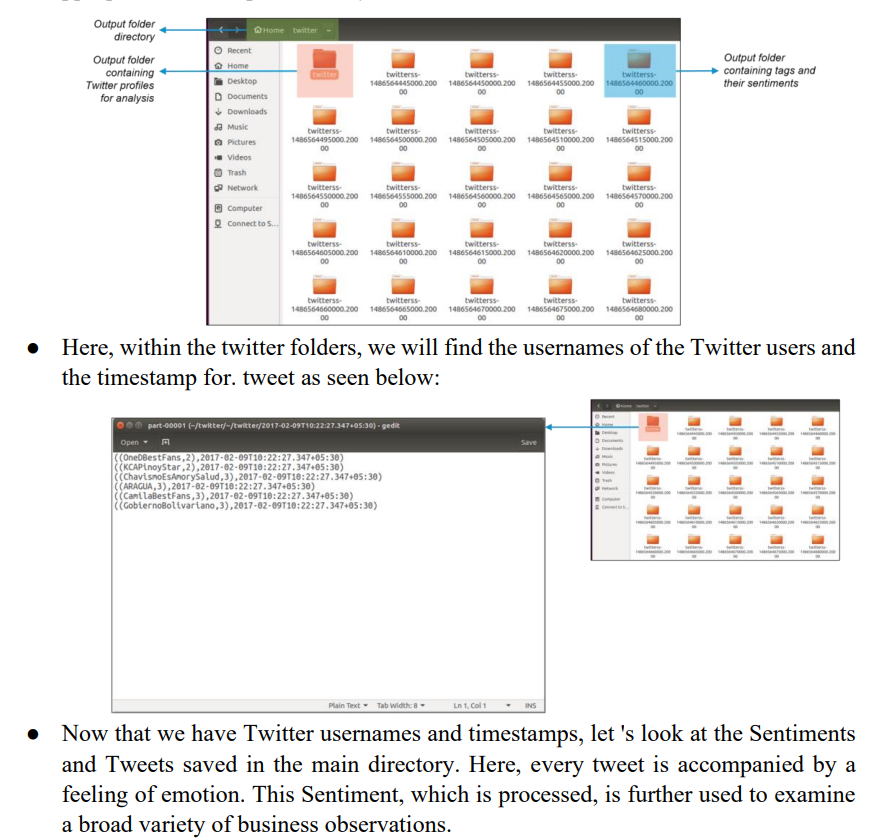
**Spark Streaming Implementation:**

****

****

**OUTPUT:**

****

****

****

#### **CONCLUSION:**

In this practical, we learnt about spark and installed it and configured it. We explored the spark shell as well.