|  |
| --- |
| /\* |
|  | \* |
|  | \* Application to find out the views of different people on the demonetization by analysing the tweets from twitter. |
|  | \* |
|  | \*/ |
|  |  |
|  |  |
|  | package sentimentAnalysis |
|  |  |
|  | import org.apache.spark.SparkContext |
|  | import org.apache.spark.SparkConf |
|  | import org.apache.spark.sql.SQLContext |
|  | import org.apache.spark.sql.SQLImplicits |
|  | import org.apache.spark.sql.types.\_ |
|  | import org.apache.spark.sql.types.StringType |
|  | import org.apache.spark.sql.\_ |
|  |  |
|  | object DemonitizationAnalysis { |
|  |  |
|  | def main(args: Array[String]): Unit = { |
|  |  |
|  | //specify the configuration for the spark application using instance of SparkConf |
|  | val config = new SparkConf().setAppName("Assignment 20.2").setMaster("local") |
|  |  |
|  | //setting the configuration and creating an instance of SparkContext |
|  | val sc = new SparkContext(config) |
|  |  |
|  | //Entry point of our sqlContext |
|  | val sqlContext = new org.apache.spark.sql.SQLContext(sc) |
|  |  |
|  | //to use toDF method |
|  | import sqlContext.implicits.\_ |
|  |  |
|  | /\* |
|  | \*create an rdd from textfile and get the id and array of words and convert the same to dataframe using toDF method |
|  | \*/ |
|  | val tweets = sc.textFile("/home/acadgild/sridhar\_scala/assignment/demonetization-tweets.csv").map(x => x.split(",")).filter(x=>x.length>=2).map(x => (x(0).replaceAll("\"",""),x(1).replaceAll("\"","").toLowerCase)).map(x => (x.\_1,x.\_2.split(" "))).toDF("id","words") |
|  |  |
|  | /\* |
|  | \* create a temporary table tweets which will be used for using sql to query from |
|  | \*/ |
|  | tweets.registerTempTable("tweets") |
|  |  |
|  | /\* |
|  | \* create dataframe by exlpoding the array of words and register the dataframe to a temporary table tweet\_word |
|  | \* |
|  | \*/ |
|  |  |
|  | val explode = sqlContext.sql("select id as id,explode(words) as word from tweets").registerTempTable("tweet\_word") |
|  |  |
|  | /\* |
|  | \*create an rdd from textfile and get the words and their ratings and convert the same to dataframe using toDF method |
|  | \* and register the same to temporary table afinn |
|  | \* afinn contains words and their ratings |
|  | \*/ |
|  |  |
|  | val afinn = sc.textFile("/home/acadgild/sridhar\_scala/assignment/AFFIN.txt").map(x => x.split("\t")).map(x => (x(0),x(1))).toDF("word","rating").registerTempTable("afinn") |
|  |  |
|  | /\* |
|  | \* join tweet\_word and afinn table and get the rating for each users tweet |
|  | \* |
|  | \*/ |
|  | val join = sqlContext.sql("select t.id,AVG(a.rating) as rating from tweet\_word t join afinn a on t.word=a.word group by t.id order by rating desc").show |
|  |  |
|  |  |
|  | } |

