**Lab Assignment 5 and 6:**

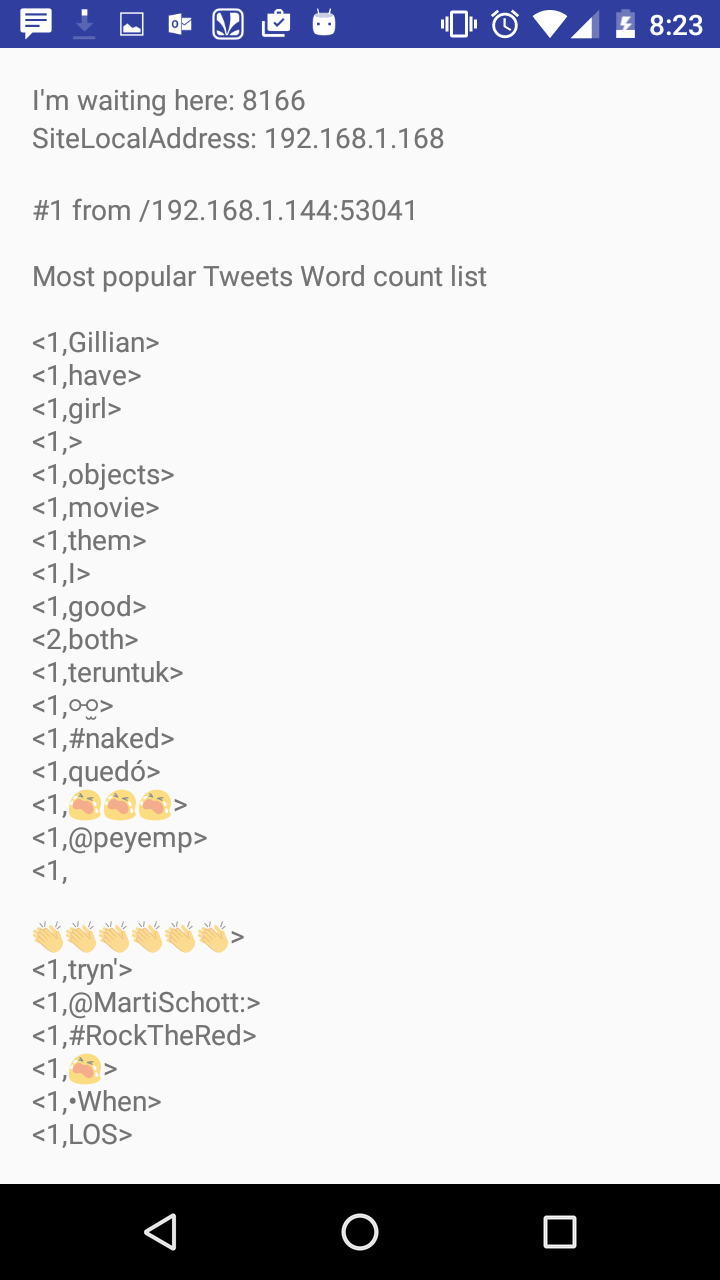
**Question 1: Spark and Smartphone/Watch Application:**

Implement a smart application with big data analytics related to your project showing the collaboration between Spark and Smart Apps. Implement Twitter Streaming and perform word count on it and publish the results and showcase it in your Smart Phone/Watch Application

**Description**:

I have created a streaming context to get the twitter stream data and performed Map Reduce framework on the twitter stream to determine the word count stream. By establishing the socket connection I connected spark to smartphone and shown the Twitter stream word count in the phone.

**Screenshots:**

****

**Question 2: Spark ML Lib Application**

Perform a machine learning algorithm with the Twitter Streaming data to categorize each Tweet

1) Training datasets: Collect different categories of Tweets related to your project.(Categories can be based on Hashtags /Subjects etc.)

2) Test data: the upcoming twitter stream.

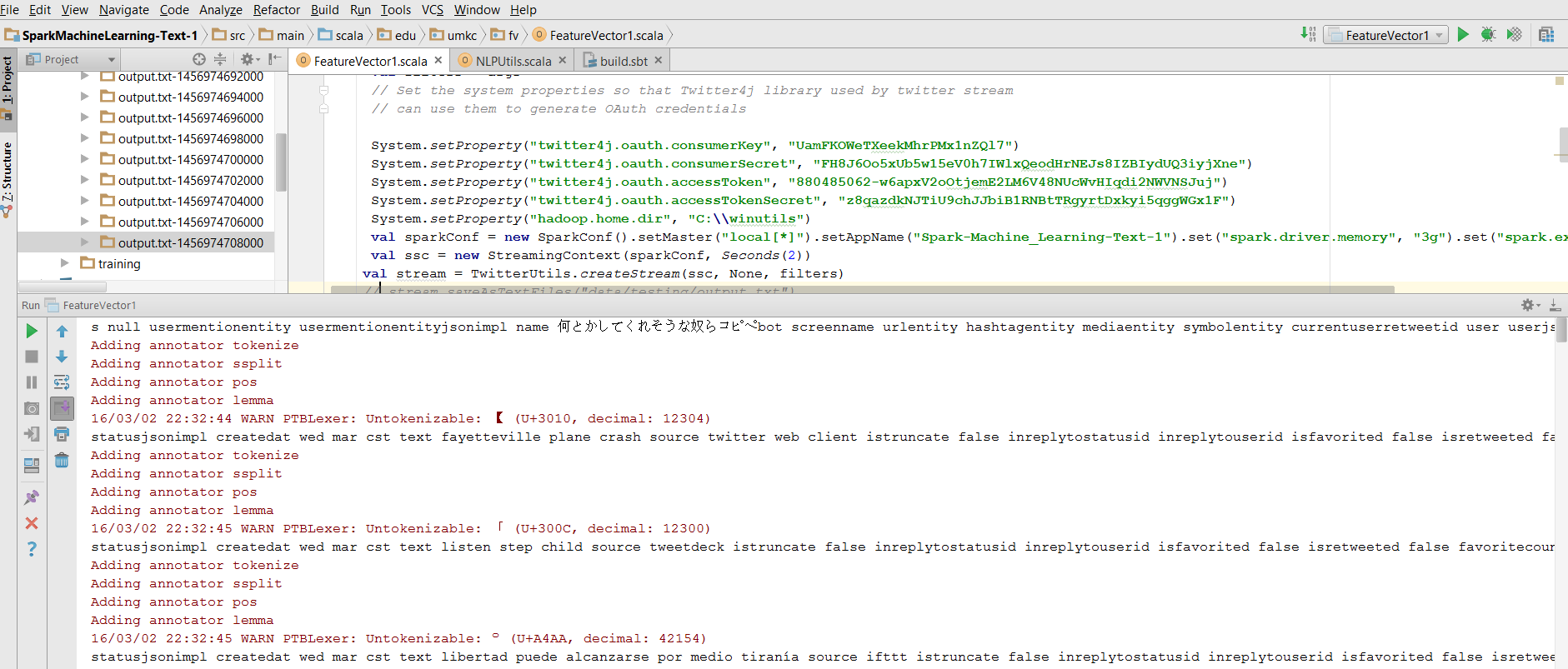
**Description:**

I have used *Naïve Bayes model* to categorize different tweets.

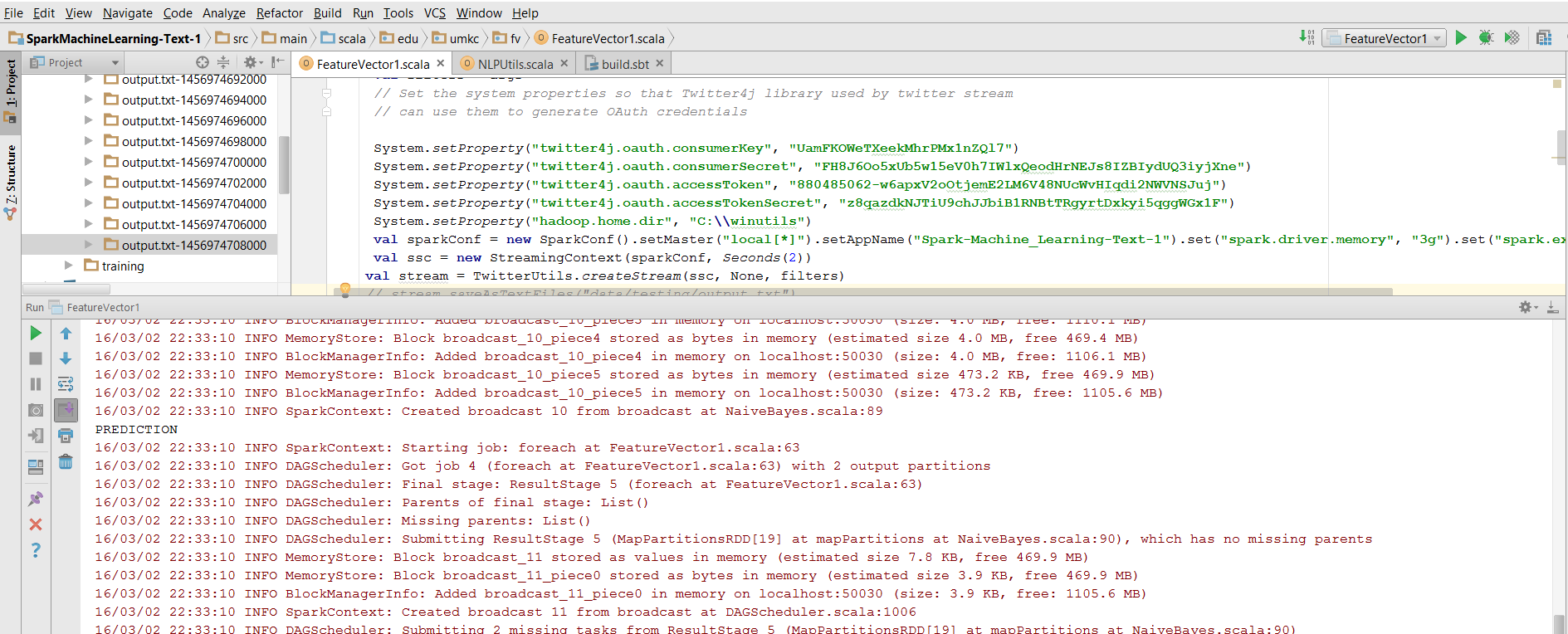
1. Training set: I collected the tweets based on 3 key words and are popular trends currently. Three keywords are:
2. “health”
3. “android”
4. “jobs”
5. I tested the data by sending the live twitter stream and the results are shown below in the screenshots

**Screenshots:**

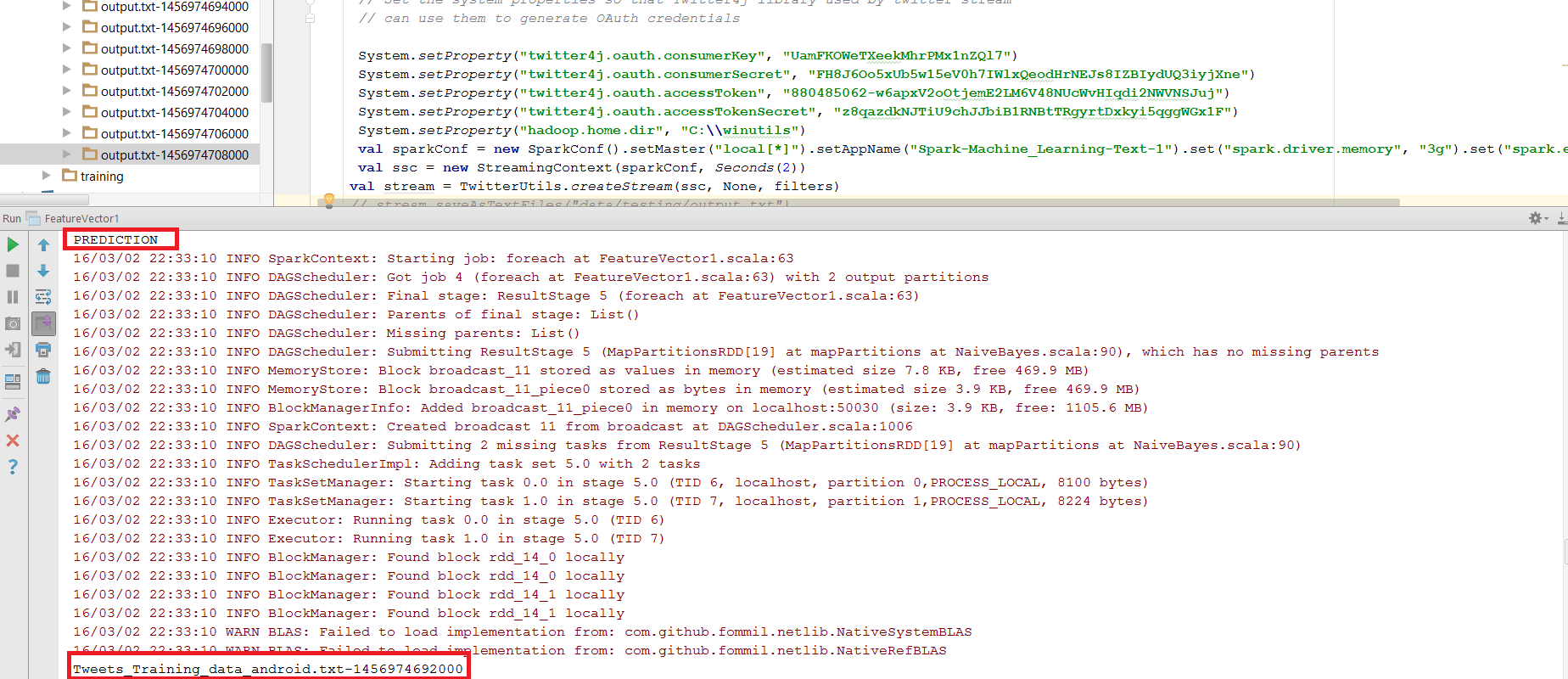
Tokenizing and lemmentizing the training data:

****

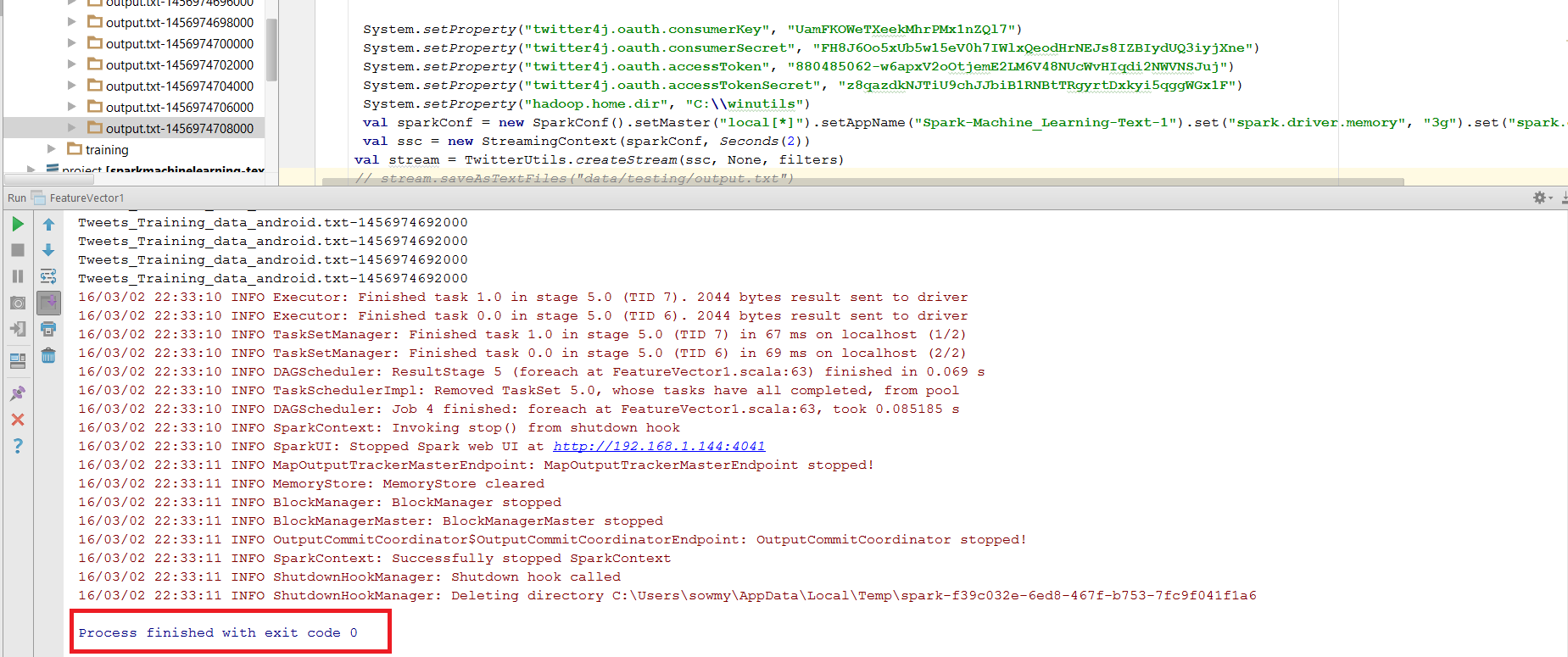
Predicting the Output:

****

Prediction of Output :

****

Execution completed successfully:

****