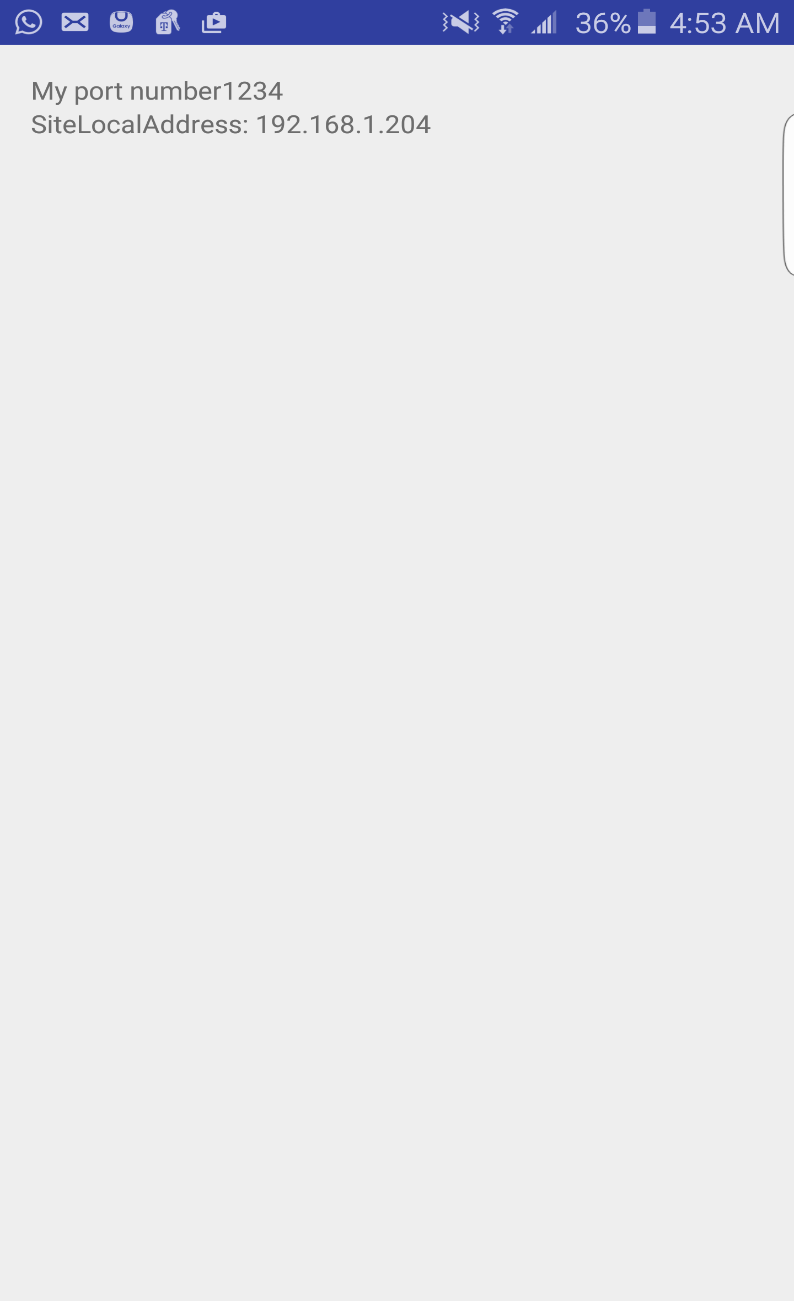
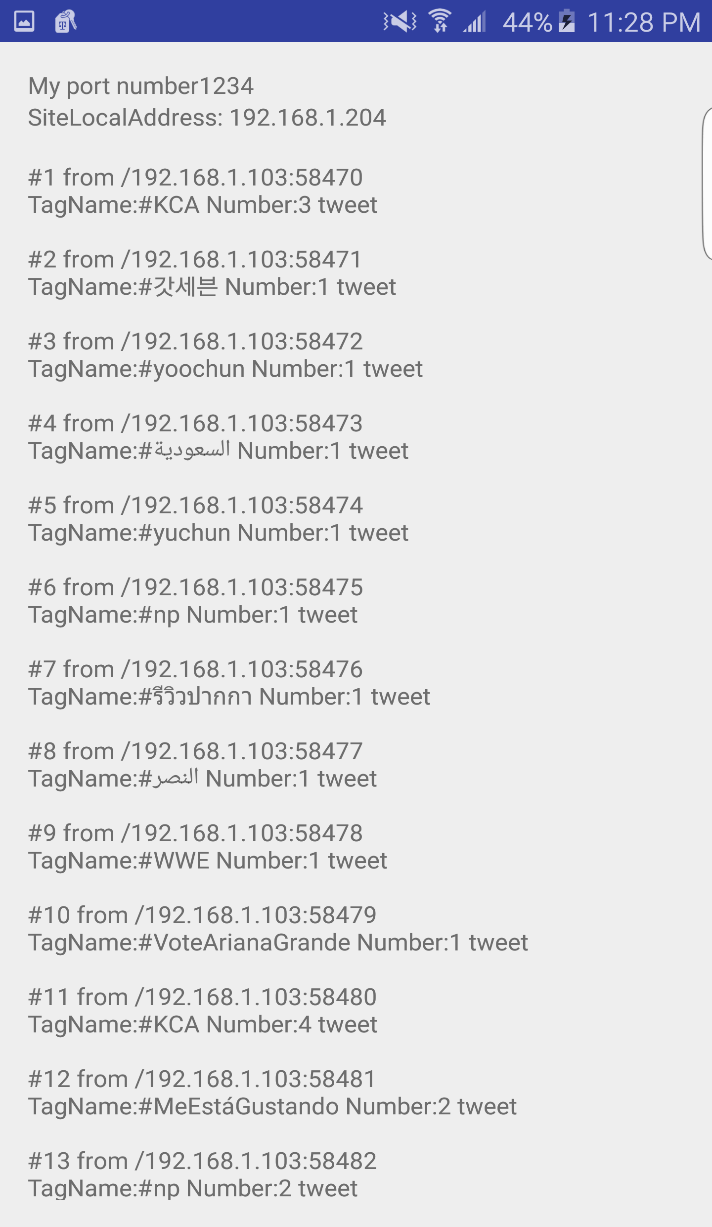
CS5542 Big Data Apps and Analytics  
LAB ASSIGNMENT #5 &6  
REPORT and SCREEN SHOTS

**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.**

1. We open a socket connection between the smart phone and spark . when we run the twitter streaming job, for every 10 seconds hashtags are sent to device.



1. Tweets Top hash tags are displayed on Smart Phone screen



**2. Spark ML Lib ApplicationPerform a machine learning algorithm with the Twitter Streaming data tocategorize each Tweet**

**1)Training datasets: Collect different categories of Tweets related to your project.(Categoriescan be based on HashTags /Subjects etc.)**

**2)Test data:theupcoming twitter stream**

**1)** I have collected the different categories of tweets such as Tweets related to restaurants, movies, sports.

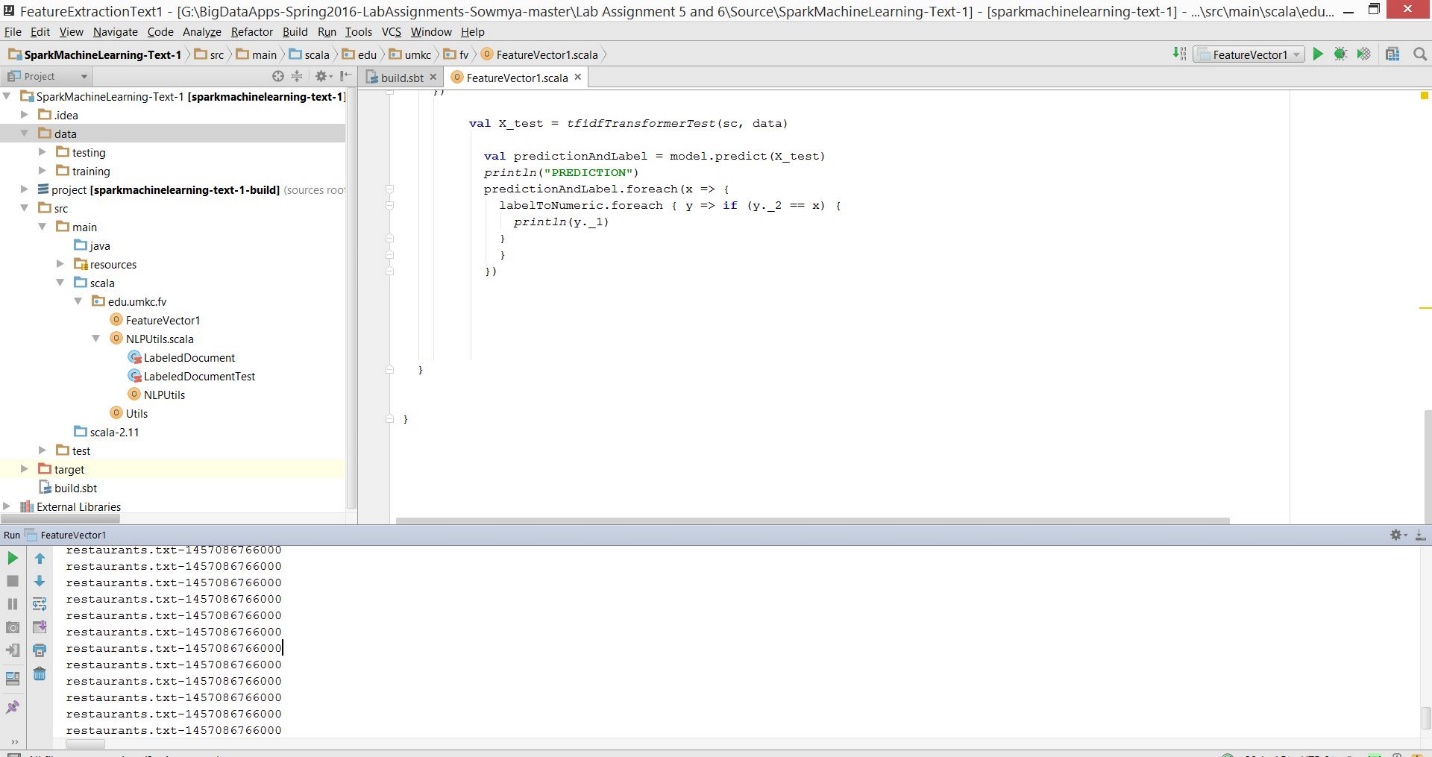
I trained my model to predict these categories using the Naïve Bayes Machine learning algorithm

The collected data is Divided into training data and Test data.

Training Data consists of Tweets related to three different categories. I collected the training data using the filters.

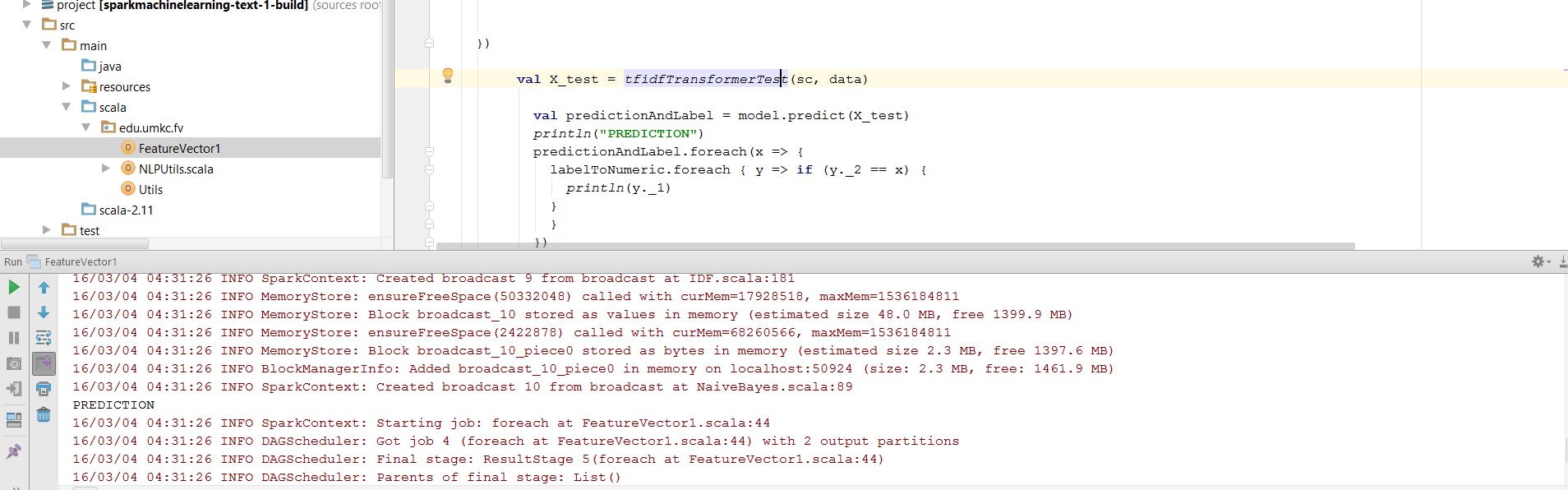
Testing data is collection of streaming tweets and we tried to predict the tweets which are more related to which category using classification algorithm.

Training Data and Test Data under Data tab



Tweets in test data are processed using the NLP

Prediction:



Predicting the given test data is more falls under the category of Sports Tweets.

