**CS5542 Big Data Apps and Analytics**

**LAB ASSIGNMENT #5**

**Lab Submission before 2017/2/22(W) 11:59PM.**

* Follow the guidelines for uploading assignment to GitHub
* Submit the screenshots of the task under documentation folder in GitHub and push the code to Source folder in GitHub
* **No late submissions** (penalty or not) will be accepted.
* Create **wiki page** for report and include **images in wiki page**
* Submit your Github project for Lab Assignments with source code, input and output data and wiki page containing the report for your lab Assignment. [**https://goo.gl/forms/L5NVqbdwBwkwK2n13**](https://goo.gl/forms/L5NVqbdwBwkwK2n13)

1. **Spark Programming:**

**Write a spark program for the following Machine Learning Task.**

Create your own dataset for Image Classification Problem. Use the workflow as discussed in the Tutorial 4 Session using any classification algorithm (e.g., Random Forest, Naïve Bayes) excluding Decision Tree. Report the accuracy and confusion matrix obtained.

In the Wiki Page, include a brief description of your dataset and purpose behind image classification problem

1. **Client Application using Spark API**

Write a client application using the Spark API to connecting between Spark and your client. Your client can be either Web application or Android application. Refer to Tutorial 5 Spark API tutorial.

1. **Google Conversion Actions API**

Build a simple application to have a conversion using Google Conversation Actions API about the summary you had generated about your video. Refer to Tutorial 5 Conversion Actions API tutorial.