**CS5542 Big Data Apps and Analytics**

**LAB ASSIGNMENT #3**

**Lab Submission through the following link before 2017/2/8(W) 11:59PM.**

[**https://goo.gl/forms/L5NVqbdwBwkwK2n13**](https://goo.gl/forms/L5NVqbdwBwkwK2n13)

1. **Spark Programming:**

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

A group of primatologists wants to study the details of the daily movement, activities, and interactions of a group of 6 chimpanzees living on "chimp island" - a natural, though somewhat open habitat about 50 meters in diameter, bounded on all sides by water, in the San-Diego zoo. Since they don't want to sit all day every day recording the second-by second positions and activities of the chimps, they have come to you, a computer vision expert, for automated assistance. They are interested in both compiling statistics about the movement and location of individuals, and in the frequency and locations of different interactions and activities (feeding, sleeping, grooming, fighting, etc.) They are willing to help in labeling relevant activities, even to the point of answering a few hundred quick questions per day of data (what's she doing here?), but they don't want to sit through 24 hours of video to do it. Ultimately they want an automated database that they can use to find out how many hours a day chimp Jane sleeps and where, histogram preferred eating locations, obtain statistics on who grooms whom, etc.

1. Implement to build a linear regression model for selected two parameters for chimpanzee’s daily movement, activities and interaction. Define your own datasets.
2. Implement K-Means clustering for the clusters of the chimpanzee’s activities. Define your own data sets.
3. **Video Annotation:**

Build a simple application to give the summary of a video by using Clarifai API. Using OpenImg Library to the key-frame images from the clarifai API.

**Summary**

**Collect all the annotation, create a summary of the video**

**Key Frame Detection using Open IMG**

**Annotation for each key frame image**

**Video**

Submit your Github project for Lab Assignments with source code, input and output data and wiki page containing the report for your lab Assignment.