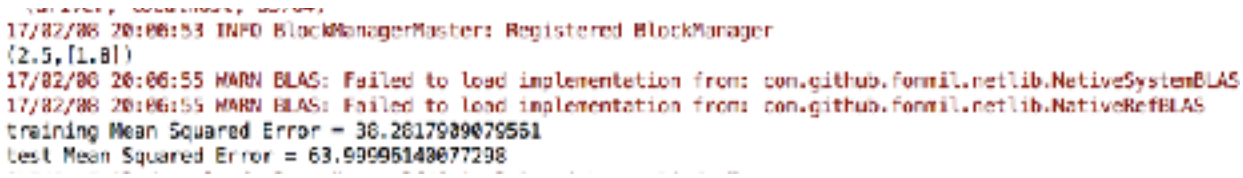
CS5542 big data apps and analytics lab assignment

1.spark programming:

write a spark programming for the chimpanzee's activities machine learning tasks.

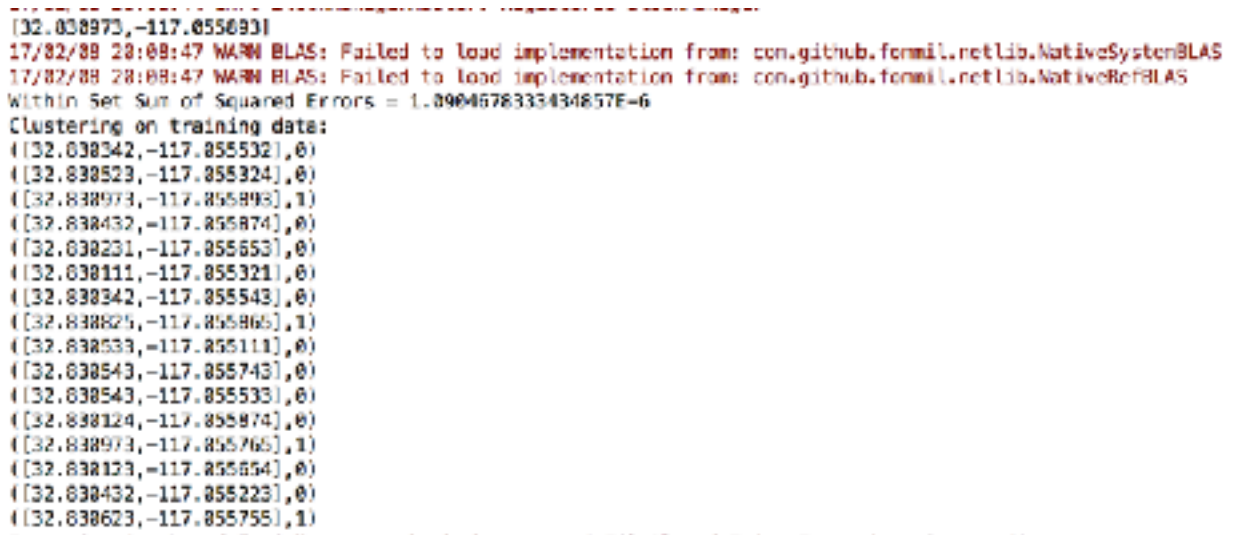
- Implement to build a linear regression model for selected two parameters for chimpanzee's daily movement, activities and interaction. Define your own datasets.

I design a dataset which defined the feeding time and the corresponding weight of food, the former number stands how many hours past from last feeding time, the second number of each line means the weight of the food that the chimpanzee ate that during meal, so that I got following result:



- Implement K-Means clustering for the clusters of the chimpanzee's activities. Define your own data sets.

I designed a dataset of the location where the chimpanzee is at 13:00 pm everyday. and the dataset totally contains two weeks' data, the result shows below:



2.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. and this is the work flow:

i follow this work flow and built a which can give the summary of the video, and here is the summary result:

```
155.0
122.0
205.0
115.0
107.0
139.0
276.0
157.0
273.0
170.0
250.0
259.0
306.0
250.0
402.0
320.0
208.0
255.0
372.0
indoors  1.9305356

Process finished with exit code 0
```

```
1  the summary word of this video is : indoors
```