**HW5 Report**

**Size of the fixed length sample    Overlap (0-X%)    K- value  Classifier    Accuracy**

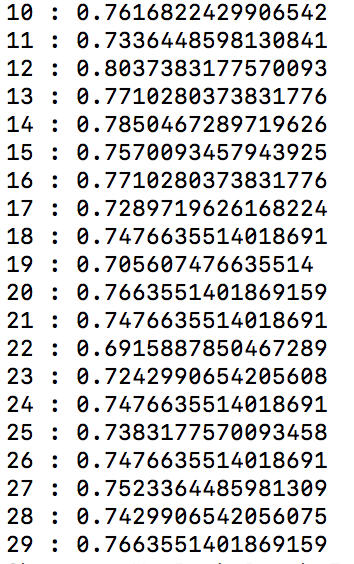
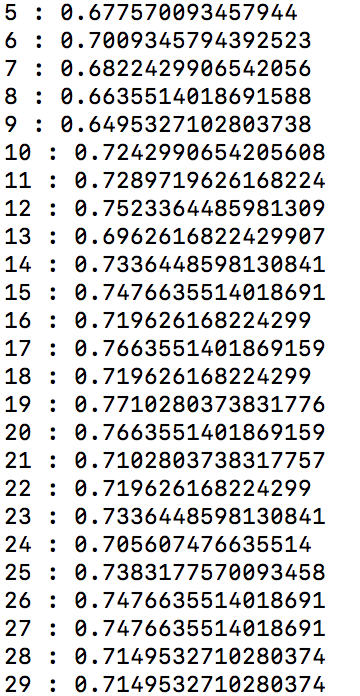
**12 \* 3 0 19 RandomForest 0.77-0.8**

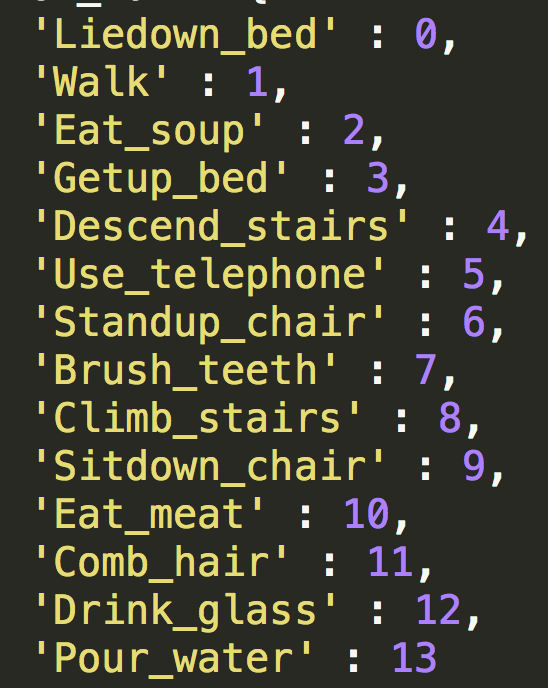
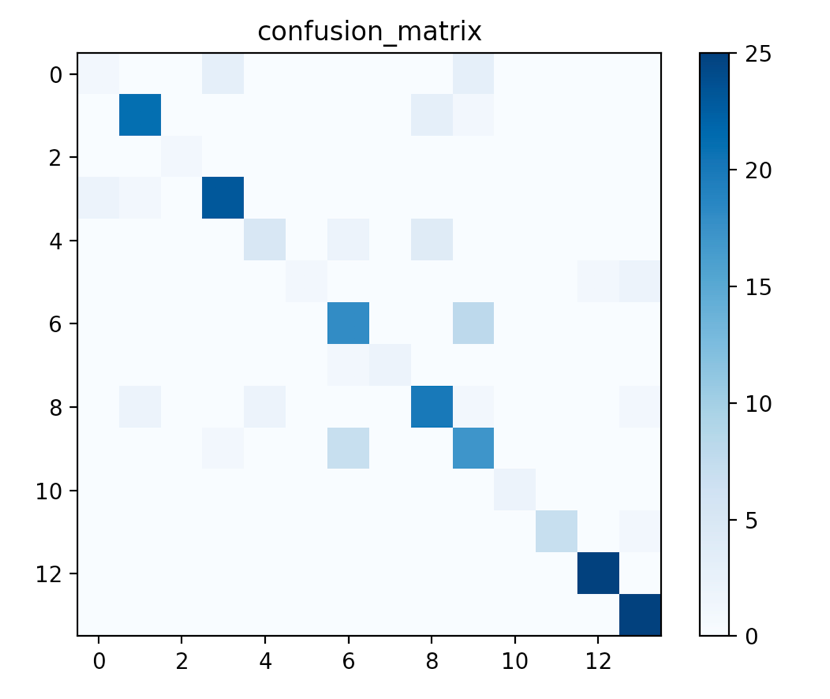
**Standard K means**

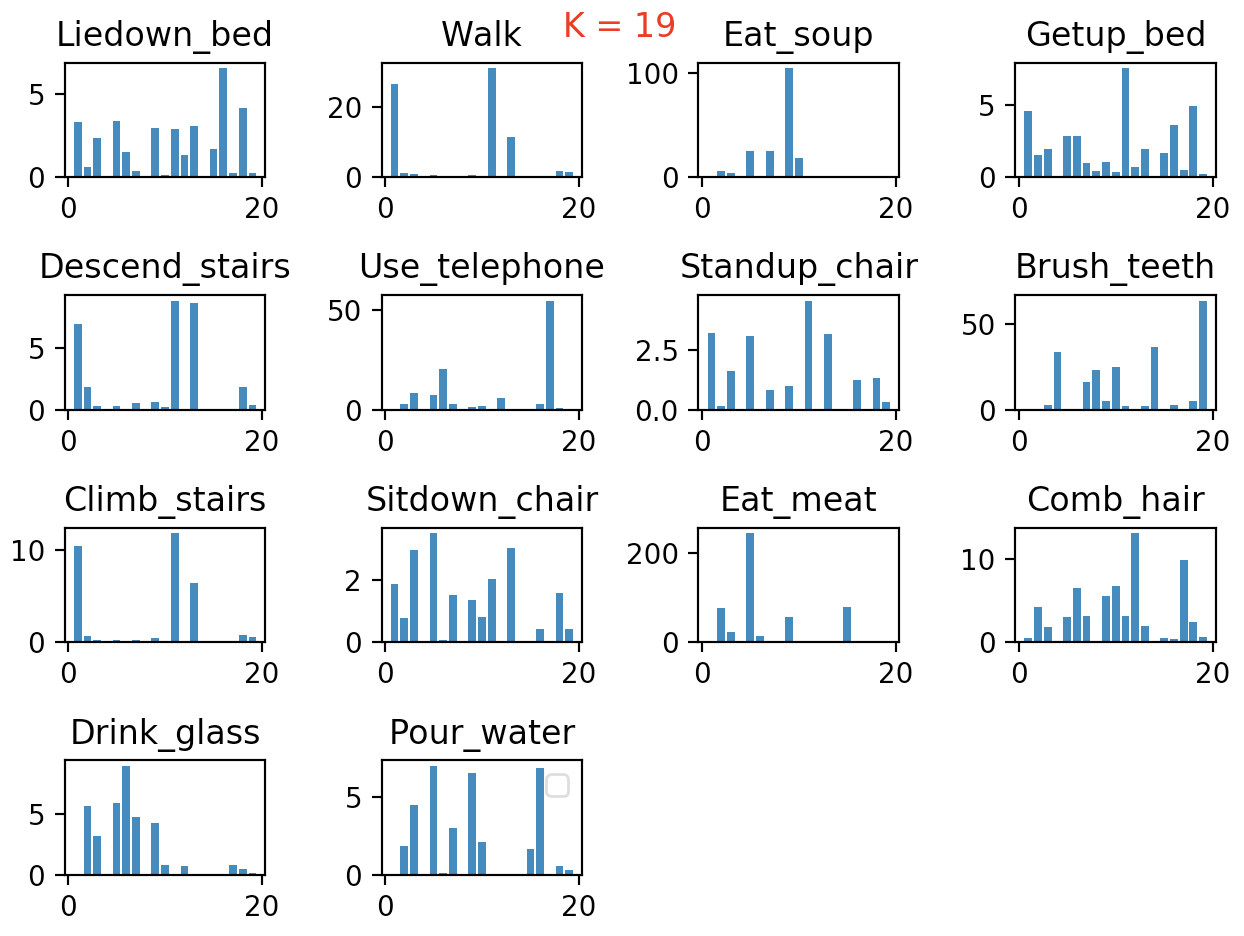
**Train test Split: 75%train, 25%test (The train error and test error does not have big difference, which means to avoid overfit problem)**

**Selection of k（test when length = 32）: Selection of fixed length（test when k = 19）:**

**K length**

****

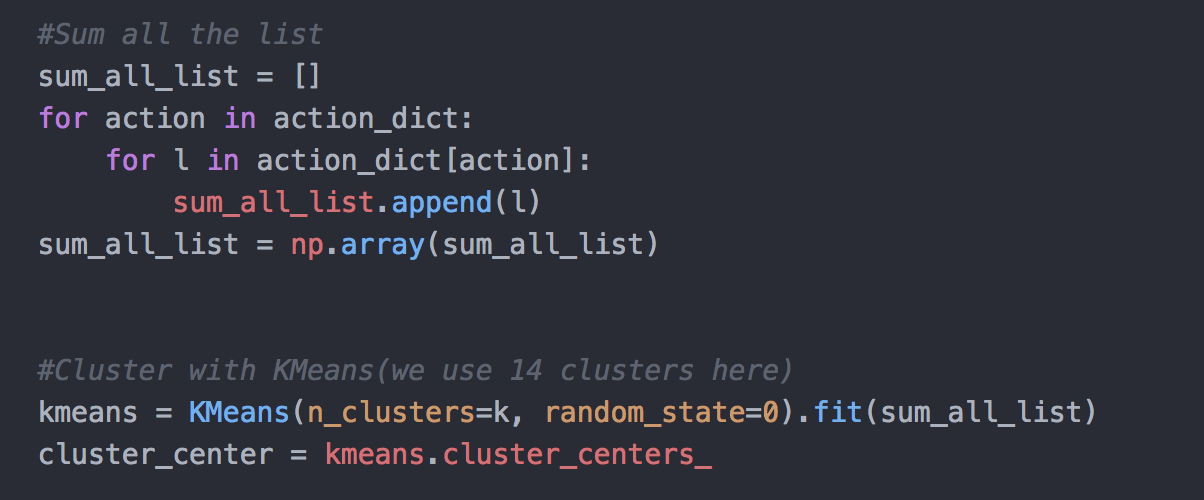
****

****

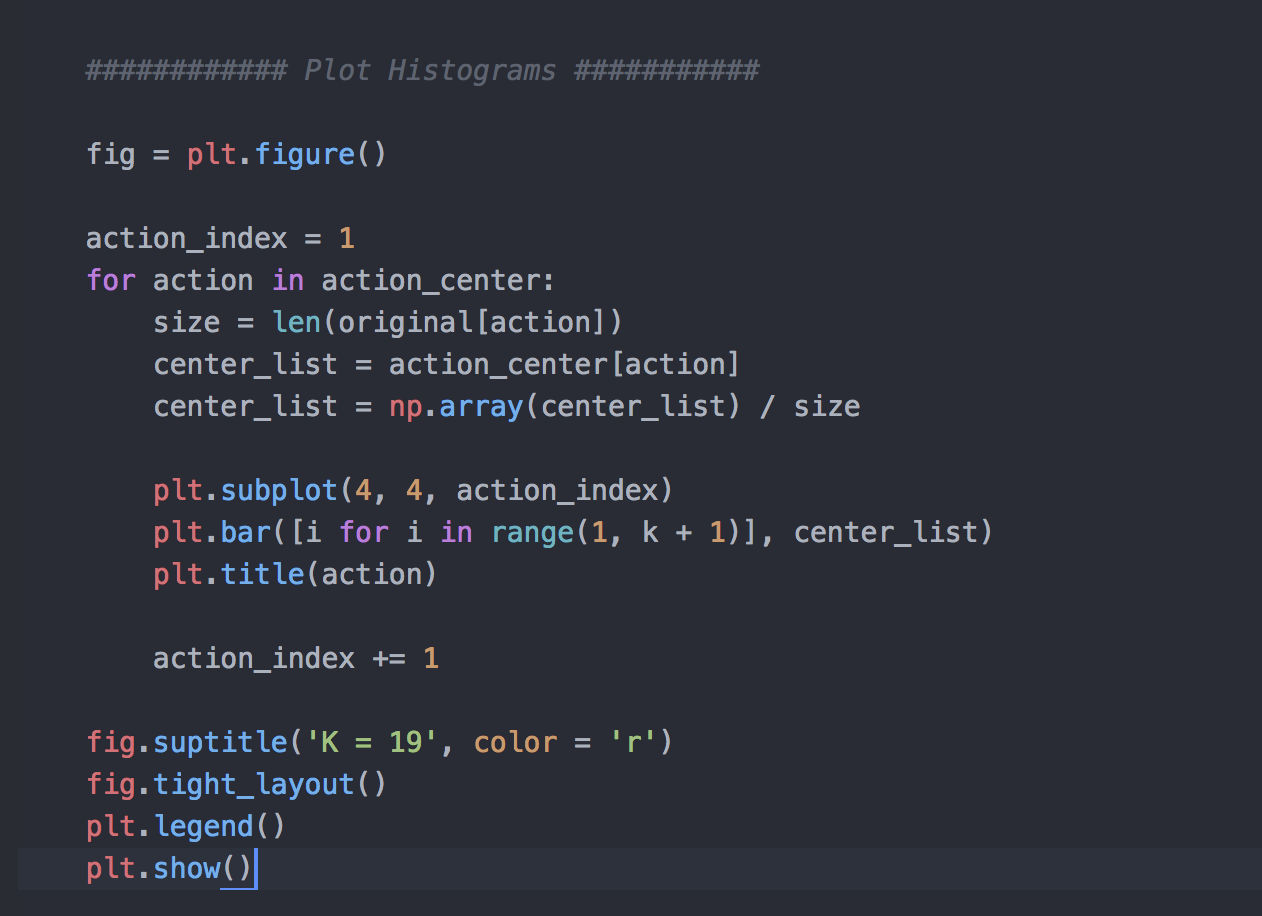
1. segmentation of the vector:

****

2. k-means

****

3. generating the histogram

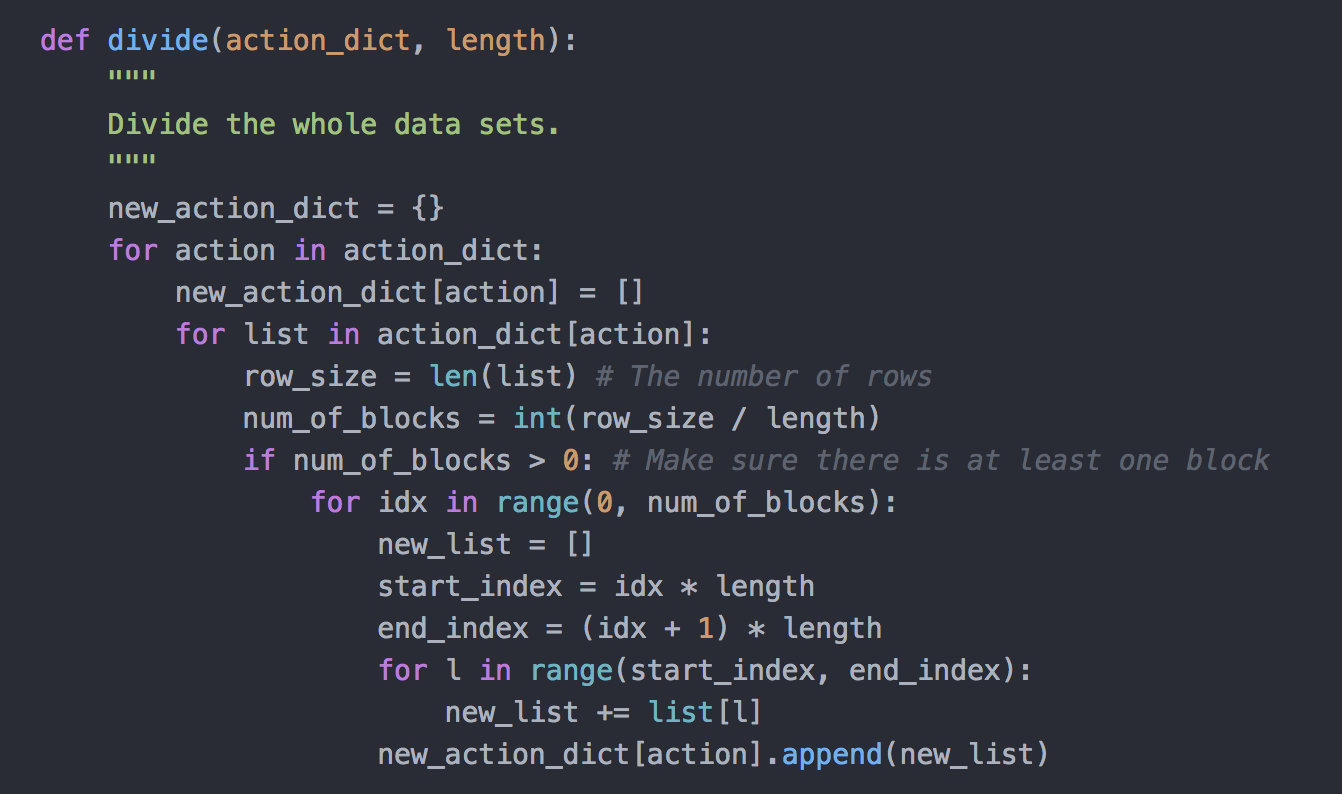
****

4. classification

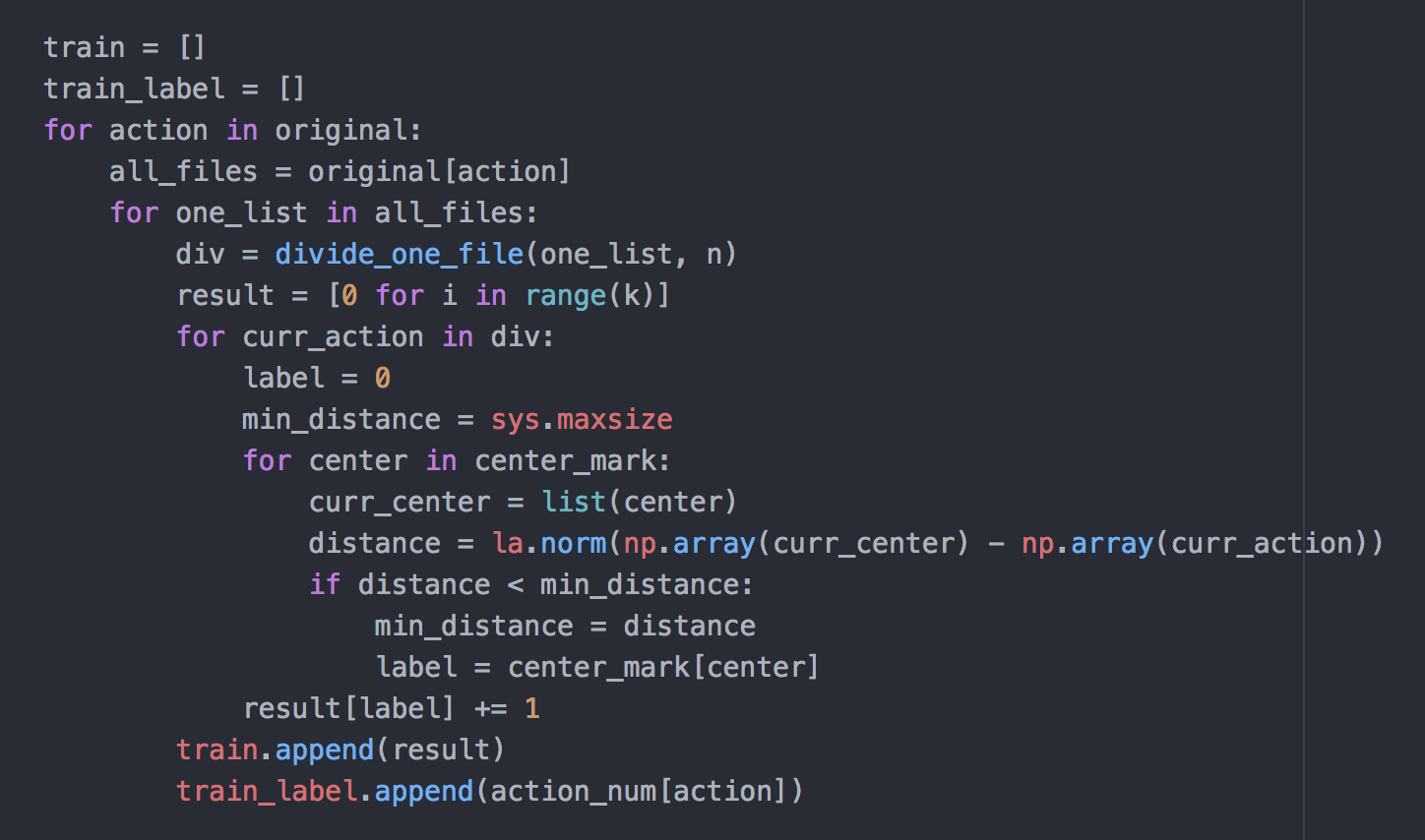
****

**Relevant Code:**

**Divide the file into segments**

****

**Generating training matrices and training labels.**

****