**Assignment 6**

NAME – Aadit Agarwal

Roll No – 2018IMT 001

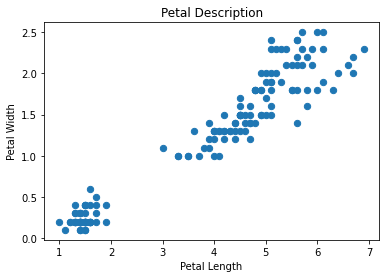
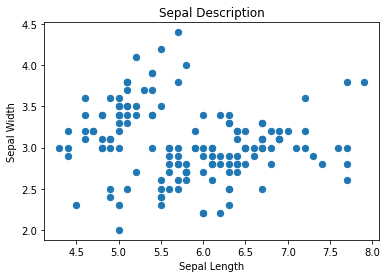
Course – Machine Learning (ITIT 4103)

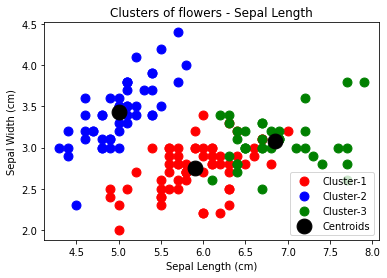
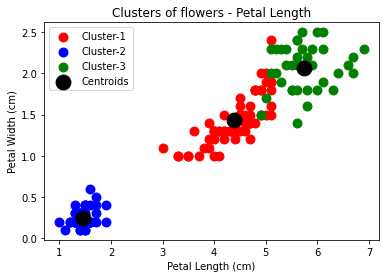
**Aim**

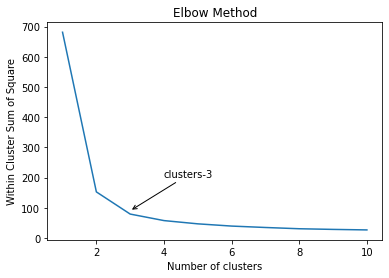
1. Considering the IRIS dataset discussed in previous assignment, apply EM algorithm to cluster the data (without considering the output labels) Use the same dataset for clustering using K-means algorithm. Compare the results of these two algorithms.

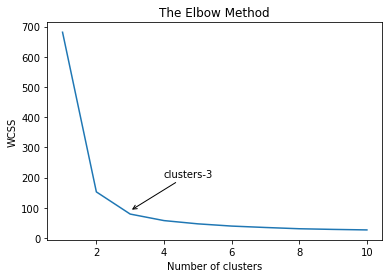
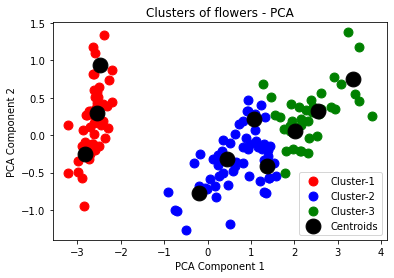
2. Apply PCA algorithm to obtain first two principal components and perform the clustering using both algorithms on the resultant data.  Compare the results of these two algorithms.

* Data Description



* K-Means



* K-Means PCA
* Results

Accuracy of K-means and EM models

1. K-Means model: 24 %

2. Elbow Method: 41.333333333333336% %

Accuracy of K-means and EM models with PCA

1. K-Means model with PCA: 88.66666666666667 %

2. Elbow Method with PCA: 64.0%

* Conclusions

It can be observed that in both, raw data and PCA data (dimensionally reduced data), EM algorithm seems to behave and perform better as compared to K-means model. EM Algorithm is a solid alternative to traditional k-means clustering on semi-supervised learning. It produces stable solutions by finding multivariate Gaussian distributions for each cluster.