Python: 3.8.8 (default, Feb 24 2021, 15:54:32) [MSC v.1928 64 bit (AMD64)]

scipy: 1.7.1

numpy: 1.19.2

pandas: 1.1.3

sklearn: 0.24.2

Hello World!

\*\*\*\*\*\*\*\*\*\*\*\*\*

Part-1-k-Means Clustering

\*\*\*\*\*\*\*\*\*\*\*\*\*

C:\ProgramData\Anaconda2\envs\P37\lib\site-packages\sklearn\cluster\\_kmeans.py:881: UserWarning: KMeans is known to have a memory leak on Windows with MKL, when there are less chunks than available threads. You can avoid it by setting the environment variable OMP\_NUM\_THREADS=1.

warnings.warn(

Chart, line chart

Description automatically generated

\*\*\*\*\*\*\*\*\*\*\*\*\*

Appropriate K from elbow curve: 3

\*\*\*\*\*\*\*\*\*\*\*\*\*

Confusion matrix for k-means clustering using elbow\_k clusters:

[[50 0 0]

[ 0 48 2]

[ 0 14 36]]

Accuracy score for k-means with elbow\_k clusters: 0.8933333333333333

Confusion matrix for k-means clustering using k=3 clusters:

[[50 0 0]

[ 0 48 2]

[ 0 14 36]]

Accuracy score for k-means with k=3 clusters: 0.8933333333333333

\*\*\*\*\*\*\*\*\*\*\*\*\*

Part-2-GMM Clustering

\*\*\*\*\*\*\*\*\*\*\*\*\*

C:\ProgramData\Anaconda2\envs\P37\lib\site-packages\sklearn\cluster\\_kmeans.py:881: UserWarning: KMeans is known to have a memory leak on Windows with MKL, when there are less chunks than available threads. You can avoid it by setting the environment variable OMP\_NUM\_THREADS=1.

warnings.warn(

Chart, line chart

Description automatically generated

\*\*\*\*\*\*\*\*\*\*\*\*\*

Appropriate K from elbow curve: 3

\*\*\*\*\*\*\*\*\*\*\*\*\*

Confusion matrix for gmm clustering using k=aic\_elbow\_k clusters:

[[50 0 0]

[ 0 50 0]

[ 0 14 36]]

Accuracy score for gmm with k=aic\_elbow\_k clusters: 0.9066666666666666

Confusion matrix for gmm clustering using k=bic\_elbow\_k clusters:

[[50 0 0]

[ 0 50 0]

[ 0 14 36]]

Chart, line chart

Description automatically generated

Accuracy score for gmm with k=bic\_elbow\_k clusters: 0.9066666666666666

Confusion matrix for gmm clustering using k=3 clusters:

[[50 0 0]

[ 0 50 0]

[ 0 14 36]]

Accuracy score for gmm with k=3 clusters: 0.9066666666666666