# Clustering

## Telecom industry:

Silhouette Score for K-Means Clustering:  0.14943065554784501

Calinski-Harabasz Index for K-Means Clustering:  1255.8040901062045

Davies-Bouldin Index for K-Means Clustering:  2.130301101609589

## 

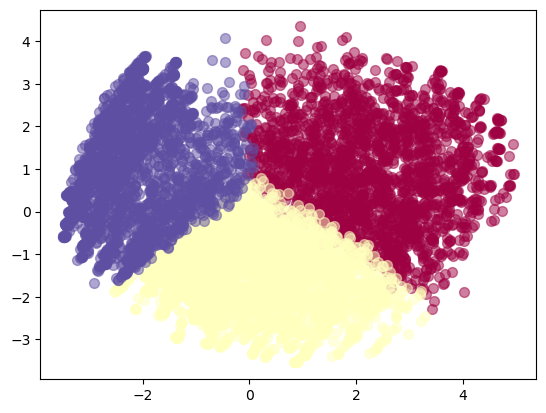
K-means

Silhouette Score: 0.14943065554784501

Calinski-Harabasz: 1255.8040901062045

Davies-Bouldin: 2.130301101609589

----------------------------------------------



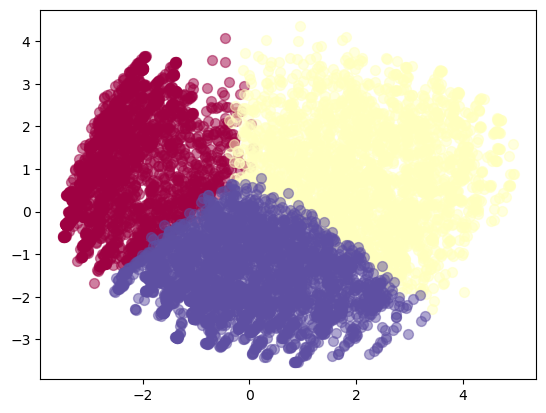
MiniBatch K-means

Silhouette Score: 0.14828064849041458

Calinski-Harabasz: 1254.6264255185285

Davies-Bouldin: 2.130917921859338

----------------------------------------------



add Codeadd Markdown

[118]:

​

*# Run Gaussian Mixture Model*

gmm **=** GaussianMixture(n\_components**=**3, random\_state**=**42)

clusters\_gmm **=** gmm.fit\_predict(X)

plot\_clusters(X, clusters\_gmm, 'Gaussian Mixture Model Clustering')

print('Gaussian Mixture Model')

evaluate\_clusters(X, clusters\_gmm)

print('----------------------------------------------')

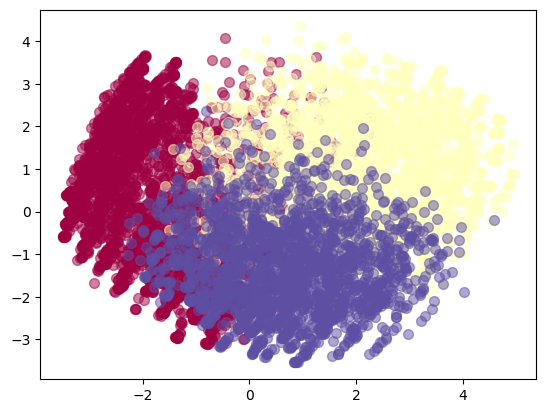
Gaussian Mixture Model

Silhouette Score: 0.12453747470142777

Calinski-Harabasz: 998.8899603296531

Davies-Bouldin: 2.530993531133974

----------------------------------------------



add Codeadd Markdown

[119]:

*# Run DBSCAN*

dbscan **=** DBSCAN(eps**=**1.5, min\_samples**=**10)

clusters\_dbscan **=** dbscan.fit\_predict(X)

plot\_clusters(X, clusters\_dbscan, 'DBSCAN Clustering')

print('DBSCAN')

evaluate\_clusters(X, clusters\_dbscan)

print('----------------------------------------------')

​

​

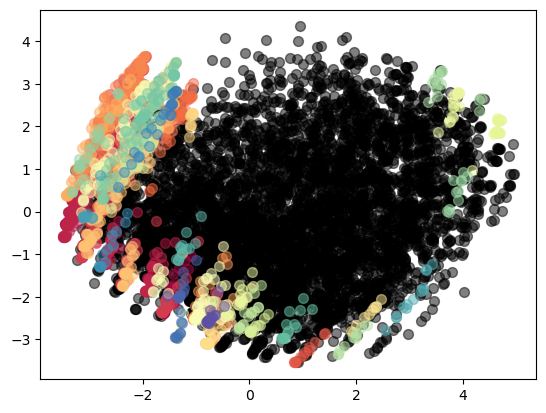
DBSCAN

Silhouette Score: -0.17926471256172252

Calinski-Harabasz: 43.60349719990029

Davies-Bouldin: 1.2929280868012354

----------------------------------------------



add Codeadd Markdown

[123]:

*# Run Fuzzy C-Means*

cntr, u, u0, d, jm, p, fpc **=** fuzz.cluster.cmeans(X.T, 3, 2, error**=**0.005, maxiter**=**1000, init**=None**)

clusters\_fuzzy **=** np.argmax(u, axis**=**0)

plot\_clusters(X, clusters\_fuzzy, 'Fuzzy C-Means Clustering')

print('Fuzzy C-Means')

evaluate\_clusters(X, np.asarray(clusters\_fuzzy))

print('----------------------------------------------')

​

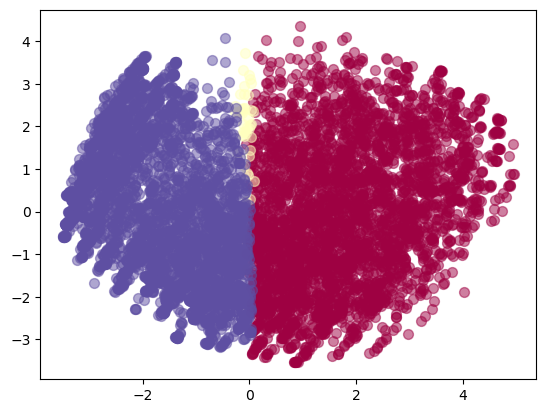
Fuzzy C-Means

Silhouette Score: 0.115459592056127

Calinski-Harabasz: 760.0847338472879

Davies-Bouldin: 2.3228681718426456

----------------------------------------------



add Codeadd Markdown

[130]:

​

*# Run Spectral Clustering*

spectral **=** SpectralClustering(n\_clusters**=**3, random\_state**=**42)

clusters\_spectral **=** spectral.fit\_predict(X)

plot\_clusters(X, clusters\_spectral, 'Spectral Clustering')

print('Spectral Clustering')

evaluate\_clusters(X, clusters\_spectral)

print('----------------------------------------------')

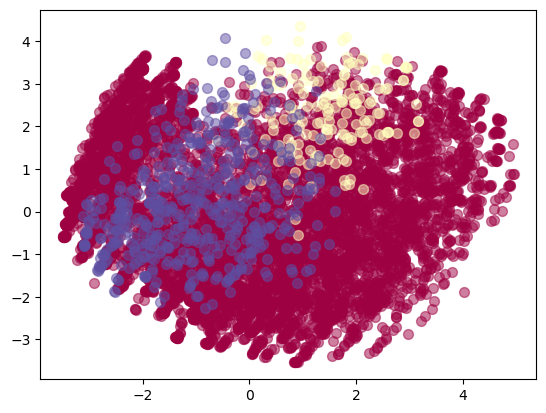
Spectral Clustering

Silhouette Score: 0.10157074445417523

Calinski-Harabasz: 291.27677998077587

Davies-Bouldin: 2.0606948093989477

----------------------------------------------



Sample output for telecom industry:

KMeans Clustering

Silhouette Score: 0.35

Calinski-Harabasz Score: 110.11

Davies-Bouldin Index: 1.20

MiniBatchKMeans Clustering

Silhouette Score: 0.37

Calinski-Harabasz Score: 104.98

Davies-Bouldin Index: 1.24

GMM Clustering

Silhouette Score: 0.33

Calinski-Harabasz Score: 96.34

Davies-Bouldin Index: 1.34

DBSCAN Clustering

Silhouette Score: 0.19

Calinski-Harabasz Score: 37.67

Davies-Bouldin Index: 1.96

Fuzzy C-means Clustering

Silhouette Score: 0.32

Calinski-Harabasz Score: 92.56

Davies-Bouldin Index: 1.39

|  |  |  |  |
| --- | --- | --- | --- |
| Algorithms | Silhouette Score | Calinski -Harabasz Index | Davies -Bouldin Index |
| K -Means | 0.1494 | 1255.8 | 2.13 |
| MiniBatch K -means | 0.1482 | 1254.6 | 2.1309 |
| Gaussian Mixture Model | 0.1245 | 998.88 | 2.5309 |
| DBSCAN | -0.1792 | 43.6 | 1.292 |
| Fuzzy C -Means | 0.115 | 760.08 | 2.322 |
| Spectral Clustering | 0.1015 | 291.2 | 2.06 |