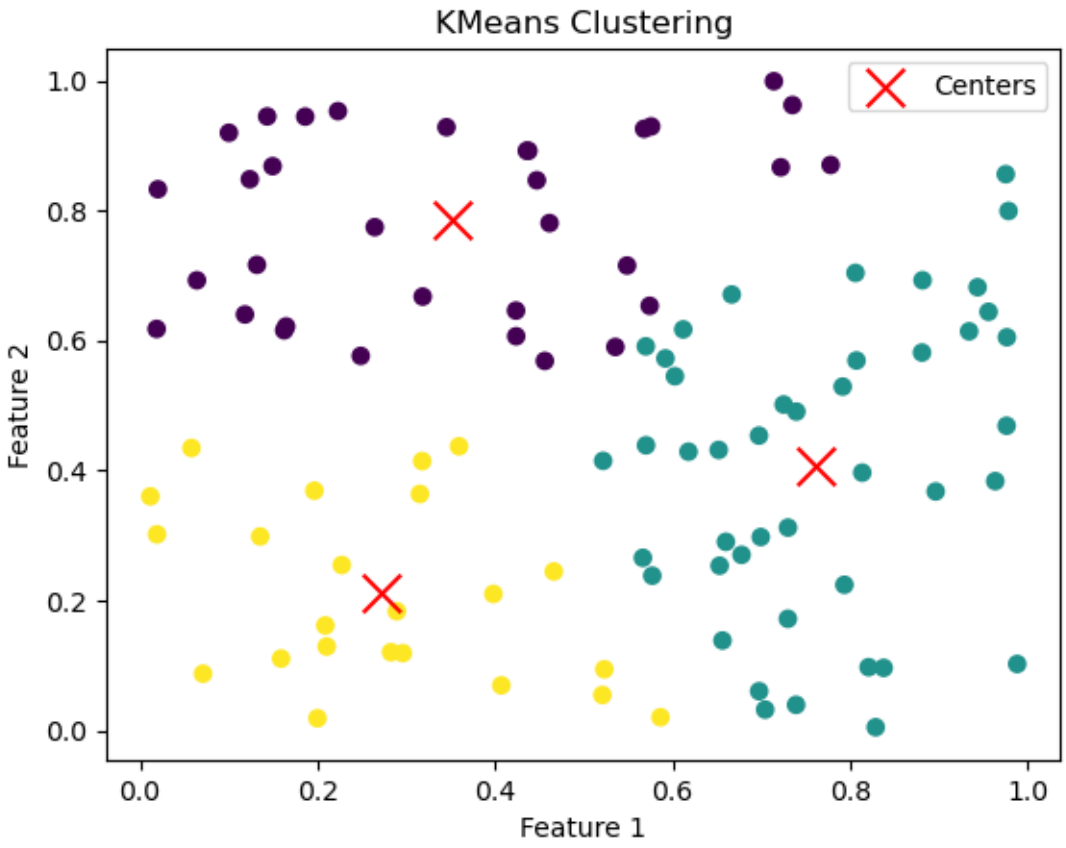
Kmeansclusters

March 6, 2025

[31]:
```
from sklearn.cluster import KMeans
import numpy as np
import matplotlib.pyplot as plt

# Generating some random data points
np.random.seed(0)
X = np.random.rand(100, 2)  # 100 data points with 2 features

# Creating KMeans instance
kmeans = KMeans(n_clusters=3, random_state=0)

# Fitting the model to the data
kmeans.fit(X)

# Getting the cluster centers and labels
centers = kmeans.cluster_centers_
labels = kmeans.labels_

# Visualizing the data and clusters
plt.scatter(X[:, 0], X[:, 1], c=labels, cmap='viridis')
plt.scatter(centers[:, 0], centers[:, 1], marker='x', c='red', s=200,␣
 ↪label='Centers')
plt.title('KMeans Clustering')
plt.xlabel('Feature 1')
plt.ylabel('Feature 2')
plt.legend()
plt.show()
```

```
C:\Users\Thinkpad\anaconda3\Lib\site-packages\sklearn\cluster\_kmeans.py:1446:
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(
```


KMeans Clustering
Centers
Feature 2
Feature 1
0.0
0.2
0.4
0.6
0.8
1.0

[ ]: