Assignment - 29 - MACHINE LEARNING - 9

By Prakash Ghosh

Problem Statement: Compress racoon grey scale image into 5 clusters and visualize both raw and compressed image and look for quality difference

- The raw image is available in spicy.misc package with the name face.
- Hint:
  import numpy as np
  from sklearn import cluster, datasets
  from scipy import misc

In [1]:
```
# Import Libraries
import numpy as np
import pandas as pd
from scipy import misc

import scipy
from sklearn import cluster, datasets
from sklearn.cluster import KMeans
import matplotlib.pyplot as plt
```

Load Data

In [2]:
```
face=scipy.misc.face()
face.shape
```

Out[2]: (768, 1024, 3)

Analyze Data (RAW Image)

In [4]:
```
face = scipy.misc.face(gray=True)

plt.figure(figsize=(20, 4.6))
plt.imshow(face,cmap=plt.cm.gray)
plt.show()
```

Prepare Data (Compress Image)

In [5]:
```
# extract rows and columns
face = scipy.misc.face(gray=True)
rows = face.shape[0]
cols = face.shape[1]
image = face.reshape(rows*cols,1)
print(rows)
print(cols)
```

```
768
1024
```

K-Mean Analysis

In [9]:
```
kmeans = KMeans(n_clusters = 5)
kmeans.fit(image)
```

Out[9]:
```
KMeans(algorithm='auto', copy_x=True, init='k-means++', max_iter=300,
    n_clusters=5, n_init=10, n_jobs=1, precompute_distances='auto',
    random_state=None, tol=0.0001, verbose=0)
```

Visualize Data (Compress Image)

In [10]:
```
# Create compressed image using cluster labels
clusters = np.asarray(kmeans.cluster_centers_)
labels = np.asarray(kmeans.labels_)
labels = labels.reshape(rows,cols)

plt.imsave('Image1.png',labels)

# Visualize the compressed image
image = plt.imread('Image1.png')
plt.figure(figsize=(10, 3.6))
plt.imshow(image)
plt.show()
```