

Clustering application: medical image analysis

In [1]:

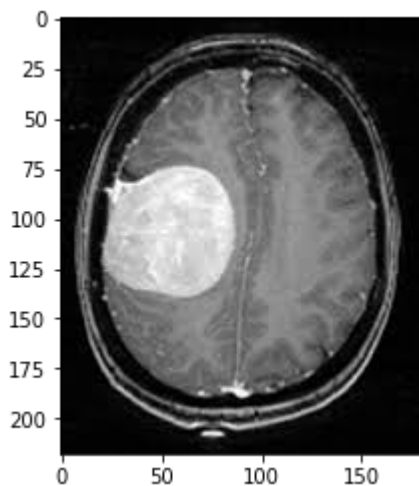
```
# Import packages
import matplotlib.pyplot as plt
import matplotlib.image as mpimg
from sklearn.cluster import KMeans
```

In [2]:

```
# Read image
img = mpimg.imread('tumor.jpg')
```

In [3]:

```
# Show image
plt.imshow(img);
```



In [4]:

```
# Show underlying data
img.shape
```

Out[4]:

```
(218, 180, 3)
```

In [5]:

```
# Reshape array such that every row of resulting matrix corresponds to a pixel
x = img.reshape((img.shape[0] * img.shape[1], 3))

# Apply K-Means algorithm
kmeans = KMeans(n_clusters=3, random_state=1234)
kmeans.fit(x)
```

Out[5]:

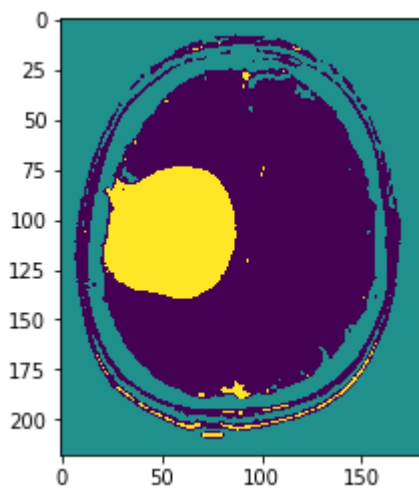
```
KMeans(algorithm='auto', copy_x=True, init='k-means++', max_iter=300,
       n_clusters=3, n_init=10, n_jobs=1, precompute_distances='auto',
       random_state=1234, tol=0.0001, verbose=0)
```

In [6]:

```
segmented_image = kmeans.labels_.reshape((img.shape[0], img.shape[1]))
```

In [7]:

```
# Show image
plt.imshow(segmented_image);
```



In [8]:

```
# Get size of tumor in numbers of pixels
(segmented_image == 2).sum()
```

Out[8]:

3789