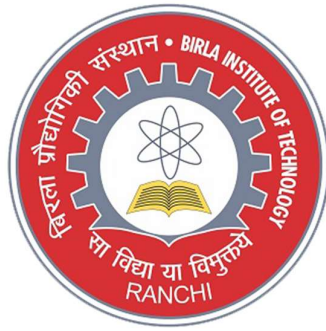


Birla Institute of Technology, Mesra,
Patna Campus



ML-Assignment

Name-Shubham Sourabh

Roll-Btech/15044/18

Sec-CSE 6th

#Assignment-8

P. Apply EM algorithm to cluster a set of data stored in a .CSV file. Use the same dataset for clustering using k-means algorithm. Compare the results of these two algorithms and comment on the quality of clustering. You can add Java/Python ML library classes/API in the program.

Code:-

```
#Imported libraries and dataset

from sklearn import datasets

from sklearn.cluster import KMeans

from sklearn.utils import shuffle

import numpy as np

import pandas as pd

#Loading iris dataset and defining the our target and data
iris=datasets.load_iris()
X=iris.data
```

```
Y=iris.target

#Shuffle of Data

X,Y = shuffle(X,Y)

#Defining model

model=KMeans(n_clusters=3,init='k-
means++',max_iter=10,n_init=1,verbose=0,random_state=3425)

#Training of the model

model.fit(X)

# This is what KMeans thought (Prediction)

Y_Pred=model.labels_

Y_Pred

#Accuracy of KMean model

from sklearn.metrics import confusion_matrix

cm=confusion_matrix(Y,Y_Pred)
print('\n\nThe Confusion matrix of K-Mean:\n',cm)
#print(cm)
```

```
print('\n')

from sklearn.metrics import accuracy_score

km=(accuracy_score(Y,Y_Pred))

print('The accuracy score of K-Mean: ',accuracy_score(Y,Y_Pred))
print('\n')

#loading data-set for EM algorithm

iris = datasets.load_iris()

X = pd.DataFrame(iris.data)

Y = pd.DataFrame(iris.target)


#Defining EM Model
from sklearn.mixture import GaussianMixture
model2=GaussianMixture(n_components=3,random_state=3425)

#Training of the model

model2.fit(X)
```

```
#Predicting classes for our data

uu= model2.predict(X)

#Accuracy of EM Model

from sklearn.metrics import confusion_matrix

cmem=confusion_matrix(Y,uu)
print('The Confusion matrix of EM-algo:\n',cmem)
#print(cm)
print('\n')
from sklearn.metrics import accuracy_score

em=(accuracy_score(Y,uu))

print('The accuracy score of EM-algo: ',accuracy_score(Y,uu),"\n")
```

Output:-

```
Windows PowerShell
Copyright (C) Microsoft Corporation. All rights reserved.

Try the new cross-platform PowerShell https://aka.ms/pscore6

PS C:\Users\vampirepapi> cd C:\Users\vampirepapi\Desktop\nowhere\6th-LABS\ML
PS C:\Users\vampirepapi\Desktop\nowhere\6th-LABS\ML> cd..
PS C:\Users\vampirepapi\Desktop\nowhere\6th-LABS> mlenv/Scripts/activate
(mlenv) PS C:\Users\vampirepapi\Desktop\nowhere\6th-LABS> cd C:\Users\vampirepapi\Desktop\nowhere\6th-LABS\ML
(mlenv) PS C:\Users\vampirepapi\Desktop\nowhere\6th-LABS\ML> python lab8.py

The Confusion matrixof K-Mean:
[[ 0  0 50]
 [47  3  0]
 [14 36  0]]

The accuracy score of K-Mean: 0.02

The Confusion matrixof EM-algo:
[[ 0  0 50]
 [45  5  0]
 [ 0 50  0]]

The accuracy score of EM-algo: 0.033333333333333333

(mlenv) PS C:\Users\vampirepapi\Desktop\nowhere\6th-LABS\ML> |
```