kmodes

November 15, 2022

20BRS1208 - P.Kailash

[ ]:
```
import pandas as pd
import numpy as np
from kmodes.kmodes import KModes
from sklearn.metrics import accuracy_score
```

[ ]:
```
df = pd.read_csv('C:\Coding\ML_python\machine-learning-lab-main\datasets\kmodes.
↪csv')
df.head()
```

[ ]:

	Person	hair color	eye color	skin color
0	p1	blonde	amber	fair
1	p2	brunette	gray	brown
2	p3	red	green	brown
3	p4	black	hazel	brown
4	p5	brunette	amber	fair

[ ]:
```
kmode = KModes(n_clusters=3, init = "random", n_init = 5, verbose=1)
clusters = kmode.fit_predict(df)
```

```
Init: initializing centroids
Init: initializing clusters
Starting iterations…
Run 1, iteration: 1/100, moves: 3, cost: 11.0
Run 1, iteration: 2/100, moves: 0, cost: 11.0
Init: initializing centroids
Init: initializing clusters
Starting iterations…
Run 2, iteration: 1/100, moves: 0, cost: 12.0
Init: initializing centroids
Init: initializing clusters
Starting iterations…
Run 3, iteration: 1/100, moves: 1, cost: 11.0
Run 3, iteration: 2/100, moves: 1, cost: 11.0
Init: initializing centroids
Init: initializing clusters
Starting iterations…
Run 4, iteration: 1/100, moves: 1, cost: 11.0
```

```
Init: initializing centroids
Init: initializing clusters
Starting iterations…
Run 5, iteration: 1/100, moves: 3, cost: 11.0
Run 5, iteration: 2/100, moves: 0, cost: 11.0
Best run was number 1
```

[ ]:
```
print(clusters)
```

```
[0 0 2 1 0 1 2 1]
```

[ ]:
```
df['cluster'] = clusters
df
```

[ ]:

	Person	hair color	eye color	skin color	cluster
0	p1	blonde	amber	fair	0
1	p2	brunette	gray	brown	0
2	p3	red	green	brown	2
3	p4	black	hazel	brown	1
4	p5	brunette	amber	fair	0
5	p6	black	gray	brown	1
6	p7	red	green	fair	2
7	p8	black	hazel	fair	1