

Inteligentni industrijski sistemi
II semestralni ispit
21.01.2021.

Napomena:

U **train_test_split** kao zadnji argument dodati **random_state=x**, gdje je **x** slučajan broj dodijeljen svakom studentu/ici.

1. Za podatke dane u clustering.csv potrebno je provesti klastering primjenom K-Means algoritma. Analizu uraditi za train/test odnos 80/20%. Kao prediktore koristiti svojstva x_1 i x_2 .

Potrebno je:

- a. Odrediti optimalan broj klastera K .
- b. Prikazati grafik tačaka klasificiranih po klasterima.
- c. Prikazati koordinate centara klastera.
- d. Predvidjeti kojim klasterima pripadaju sljedeći parovi: (-5, -5), (0, 0) i (5, 5).

2. Za podatke dane u SVM.csv potrebno je provesti klasifikaciju korištenjem Support Vector Machines algoritma. Analizu uraditi za train/test odnos 70/30%. Koristiti linearni kernel, za atribut uzeti x_1 i x_2 .

Potrebno je:

- a. Prikazati grafik neklasificiranih podataka.
- b. Prikazati indekse tačaka koje predstavljaju support vectore.
- c. Prikazati broj support vectora za svaku klasu.
- d. Uraditi predikciju za parove (-10, -7) i (0, 7).
- e. Dati procjenu tačnosti modela.

Primjer odgovora na pitanje (kod + odgovor):

```
In [7]: 1 # Importing the dataset
2 import numpy as np
3 import pandas as pd
4 from sklearn.preprocessing import StandardScaler
5 from sklearn.cluster import KMeans
6 from sklearn.metrics import silhouette_score
7
8 # Load the dataset
9 data = pd.read_csv('clustering.csv')
10 X = data[['x1', 'x2']]
11
12 # Finding the optimal number of clusters
13 sil = []
14 for k in range(1, 10):
15     kmeans = KMeans(n_clusters=k, random_state=0)
16     kmeans.fit(X)
17     sil.append(silhouette_score(X, kmeans.labels_))
18
19 # Optimal number of clusters is 3
20 optimal_k = sil.index(max(sil)) + 1
21
22 # Clustering the data
23 kmeans = KMeans(n_clusters=optimal_k, random_state=0)
24 kmeans.fit(X)
25 labels = kmeans.labels_
26
27 # Predicting the cluster for the given points
28 points = np.array([[-5, -5], [0, 0], [5, 5]])
29 predicted_labels = kmeans.predict(points)
```

R-squared= 0.9565151137
MSE= 1