

問4 (10点)

数値を答える場合は、数値のみを解答欄に記入すること。単位などは不要。また、小数第N位まで、と書いてある場合は、小数第N+1位を四捨五入して答えること。

1.

クラスタAとBには次の2次元データが含まれる。

A: (0,0), (0,1)

B: (1,1), (1,2), (2,2)

クラスタA, B間の距離をWard法により求め、小数第2位まで答えよ。

3.77

2.

ai_mid4.csvにはワインの銘柄(class列)とその特徴(alcohol, malic_acid, ash列)が記録されている。alcohol, malic_acid, ash列を用いて階層型クラスターリングを実行せよ。クラスター間距離はWard法を用い、距離関数にはユークリッド距離を使うこと。必要であれば、ai_mid4_sample.ipynbファイルを用いて良い。

クラスタ数を3としたとき、class列とのクロス集計表を作り、銘柄Aのワインのクラスタ1, 2, 3,に属する個数をそれぞれ答えよ。

クラスタ1

6

クラスタ2

46

クラスタ3

7

ipynbファイルのアップロード:

選択されていません



htmlファイルのアップロード:

選択されていません

