

2022년 2학기 경영데이터분석2 기말고사 답안지

전공: 문화관광외식학부

학 번 :

1912168

이름: 이호림

1.

(1) -/+ / + / + / +

(2) 1347

(3) b1은 유의한 음수로 b2,b4,b5은 유의한 양수로 추정하여 H1,H2,H4,H5는 채택하지만 b3는 유의한 음수로 추정하여(원래는 양수로 추정되야 함) H3은 기각한다.

(4) 0.164/0.246

(5) 444/0

(6) salarylow와 salarymedium의 2p-value가 알파인 0.05보다 작아 유의하고 estimate의 값이 양으로 유의하기 때문에 두 가설은 모두 채택한다.

(7) 0.769

(8) -1.231/ 안한다(음수 값이 나왔으니 0 즉, 퇴직 안한다.)

2.

\* According to the majority rule, the best number of clusters is 3

(1) 최적의 k개수는 3이다.

(2) 68/62/45(=2/3/1)

(3) 3

(4)

cluster	age	family	number	expense
1	0.5995973	1.2017996	-0.2776988	-1.05451729
2	0.4891084	-0.2743669	0.7750974	0.77617902
3	-0.9716331	-0.5713553	-0.6485512	-0.08591767

상대적으로 다른 군집에 비해 age가 음수의 값을 보이고 있고 family의 변수도 동일하게 다른 군집에 비해 제일 값이 작게 나왔다. number 변수 또한 다른 군집에 비해 값이 제일 작게 나온 특성을 지닌다. 즉 군집번호2(원래 cluster3)인 군집들의 age, family, number 변수는 상대적으로 다른 군집에 비해 제일 적은 값을 지닌다는 특성을 갖는다.

3.

```
k  Accuracy  Kappa
5  0.9473107  0.8499216
6  0.9426158  0.8382240
7  0.9385932  0.8279771
```

Accuracy was used to select the optimal model using the largest value.

The final value used for the model was k = 5.

(1) 최적의 k개수는 5이고 이에 따른 값은 각각 0.9473107/0.8499216 이다.

```
Importance
satisfaction 0.7150
years        0.6722
hours        0.5436
evaluation    0.5294
projects     0.5156
```

(2) 가장 중요한 변수는 satisfaction이고 중요도는 0.7150 이다.

```
[1] 1
Levels: 0 1
```

(3) 예측 값이 1이 나왔기에 퇴직여부는 퇴직한다 이다.

4.

- best parameters:

```
gamma cost
0.8 0.2
```

- best performance: 0.03180791

(1) A의 값은 0.9681921(1- 0.03180791)

Parameters:

SVM-Type: C-classification  
SVM-Kernel: radial  
cost: 0.2

Number of Support Vectors: 260

(2) 260개

1	2	3	4	5	6	7	8	9	10
0	1	0	0	0	0	0	0	0	0

(3)