

Assignment #5

머신러닝 이론과 실전

Due: October 11, 2021

1. One-vs-rest 전략을 이용하여 multiclass classification 을 수행하는 코드를 완성하시오. Classifier 는 statsmodels 패키지의 Logit 함수를 사용하시오. (절편포함)
2. Train 데이터로 위 1 번의 모델을 fitting 하시오. (veh.dat 사용)
3. Test 데이터의 각 관찰값별로, class 별 확률값과 최종예측된 class 를 출력하시오. (veh.test.dat 사용)
(아래 예시. Class 의 갯수는 데이터별로 다를수 있으니, 데이터에 맞게 자동 조정이 되어야 함)

```
ID, Actual class, Class 1, Class 2, Class 3, Final prediction
-----
1, 1, 0.8, 0.1, 0.1, 1
2, 2, 0.2, 0.7, 0.1, 2
3, 3, 0.1, 0.2, 0.8, 3
(skip: 처음 3 줄과 마지막 3 줄만 출력시킴)
198, 1, 0.8, 0.1, 0.1, 1
199, 2, 0.2, 0.7, 0.1, 2
200, 3, 0.1, 0.2, 0.8, 3
```

4. Test 데이터에 대한 Confusion Matrix 와 Accuracy 를 출력하시오. (아래 예시)

```
Confusion Matrix (Test)
-----
                Predicted Class
                1      2      3      4
Actual Class   1      239    14     6     8
                2      12    153     5    12
                3       2     4    98     2
                4       3     6     8   123

Model Summary (Test)
-----
Overall accuracy = .793
```