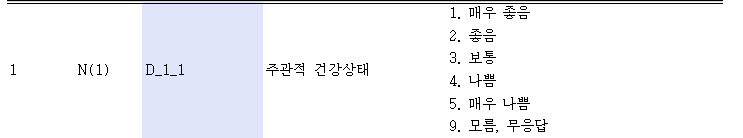
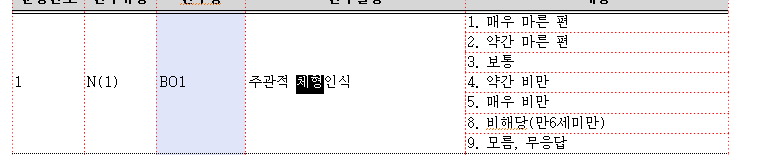
국민건강영양조사 2018 데이터 남자 모델

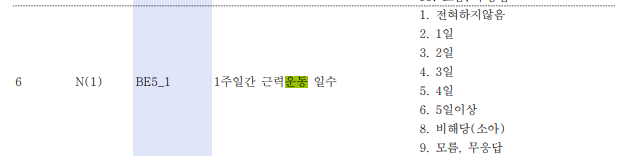
1. 남자 모델

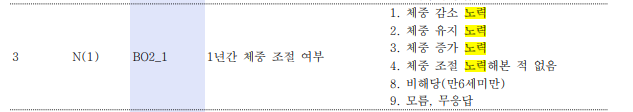
(1)변수 정규화 데이터의 규모를 줄임 (data reduction)

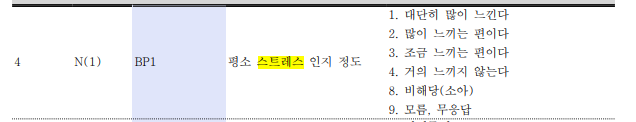
(1,2,3)(4,5)(9)

(1,2)(3)(4,5)(8,9)

 25이상이면 1, 아니면 0

(1)(2,3)(4,5)(8,9)

(1)(2)(3)(4)(89)

(1,2)(3)(4)(8,9)

이 3개를 ‘suicide’라는 변수로 하나라도 1이면 1이게끔 만들어주고 저 3개는 지움.

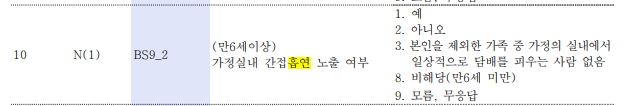


1 -> 1

2~19 -> 2

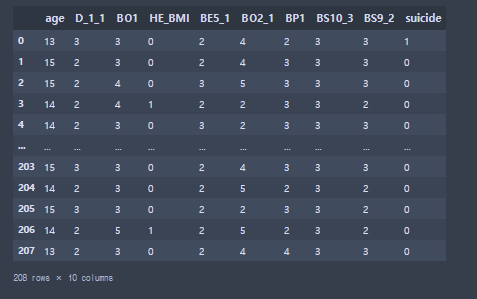
20이상 -> 3

888,999 -> 4

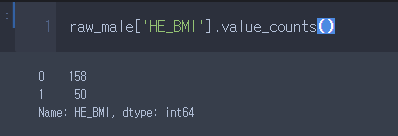
(1)(2)(3)(8,9)



🡺 타겟 변수 포함 10개, 타겟 빼면 9개



(2) 비만 target 비율



(3) Modeling

1. X, y 를 data, target 이라는 변수로 나눠 줌

2. 범주형 변수로 바꾸기 위해 팩터( astype(‘category’) )화 해 줌

3 . 정규화

4. Grid Search(RBF) + K-fold 5

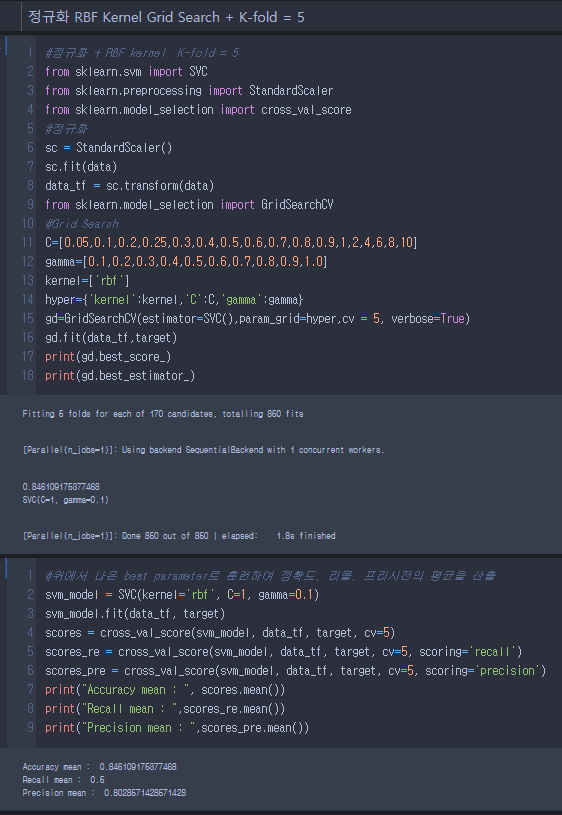
5. Grid Search(RBF) + K-fold 10

6. Grid Search(Linear) + K-fold 5

7. Grid Search(Linear) + K-fold 10

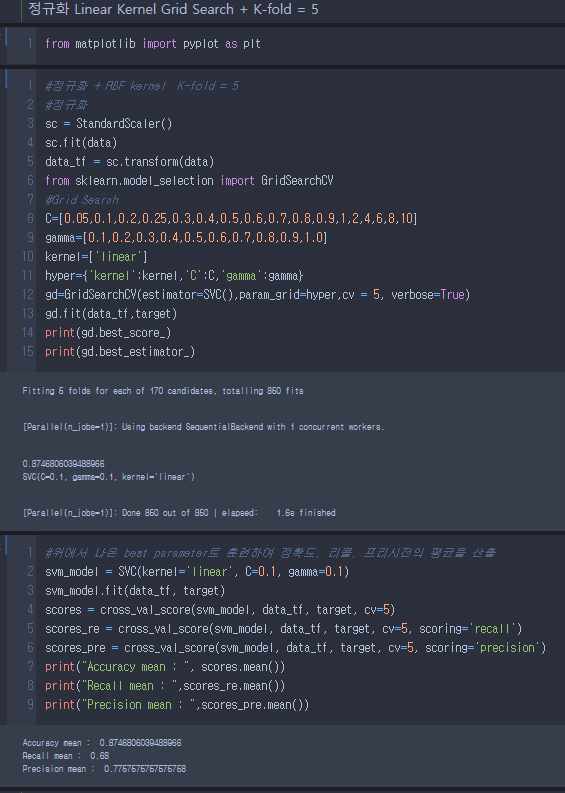
(4) 결과

-rbf





-Linear





(5) 변수 중요도 파악

