We have different dimensions for classifying a learners style. We want to find which questions are most influential for fitting the dataset.

('ILSF5AR2', 'ILSF21AR6', 'ILS25AR7K', 'ILSF37AR10K', 'ILSF41AR11')

neg\_mean\_squared\_error

cv=10

linear\_model.LinearRegression(positive=True)

Features: 1023/1023('ILSF5AR2', 'ILSF21AR6', 'ILS25AR7K', 'ILSF37AR10K', 'ILSF41AR11')

Best accuracy score: -1.18

Best subset (indices): (1, 5, 6, 9, 10)

Best subset (corresponding names):   
('ILSF5AR2', 'ILSF21AR6', 'ILS25AR7K', 'ILSF37AR10K', 'ILSF41AR11')

Logistic Regression:

Features: 1023/1023('ILSF5AR2', 'ILS25AR7K', 'ILSF33AR9K', 'ILSF37AR10K', 'ILSF41AR11')

Best accuracy score: -1.53

Best subset (indices): (1, 6, 8, 9, 10)

Best subset (corresponding names): ('ILSF5AR2', 'ILS25AR7K', 'ILSF33AR9K', 'ILSF37AR10K', 'ILSF41AR11')

1 subsets of each size up to 5

Selection Algorithm: 'sequential replacement'

ILSF1AR1K ILSF5AR2 ILSF9AR3 ILSF13AR4K ILSF17AR5 ILSF21AR6 ILS25AR7K

1 ( 1 ) " " " " " " " " " " " " " "

2 ( 1 ) " " " " " " " " " " "\*" " "

3 ( 1 ) " " " " " " " " " " "\*" "\*"

4 ( 1 ) " " "\*" " " " " " " " " "\*"

5 ( 1 ) "\*" "\*" "\*" "\*" "\*" " " " "

ILSF29AR8 ILSF33AR9K ILSF37AR10K ILSF41AR11

1 ( 1 ) " " " " "\*" " "

2 ( 1 ) " " " " "\*" " "

3 ( 1 ) " " " " "\*" " "

4 ( 1 ) " " " " "\*" "\*"

5 ( 1 ) " " " " " " " "

> coef(step.model$finalModel, 5)

(Intercept) ILSF1AR1K ILSF5AR2 ILSF9AR3 ILSF13AR4K ILSF17AR5

-0.7222217 1.9834737 1.6170949 1.3110348 1.4542281 1.1528008

|  |  |  |
| --- | --- | --- |
| Literature based: | F1;F13;F25;F33;F37 |  |