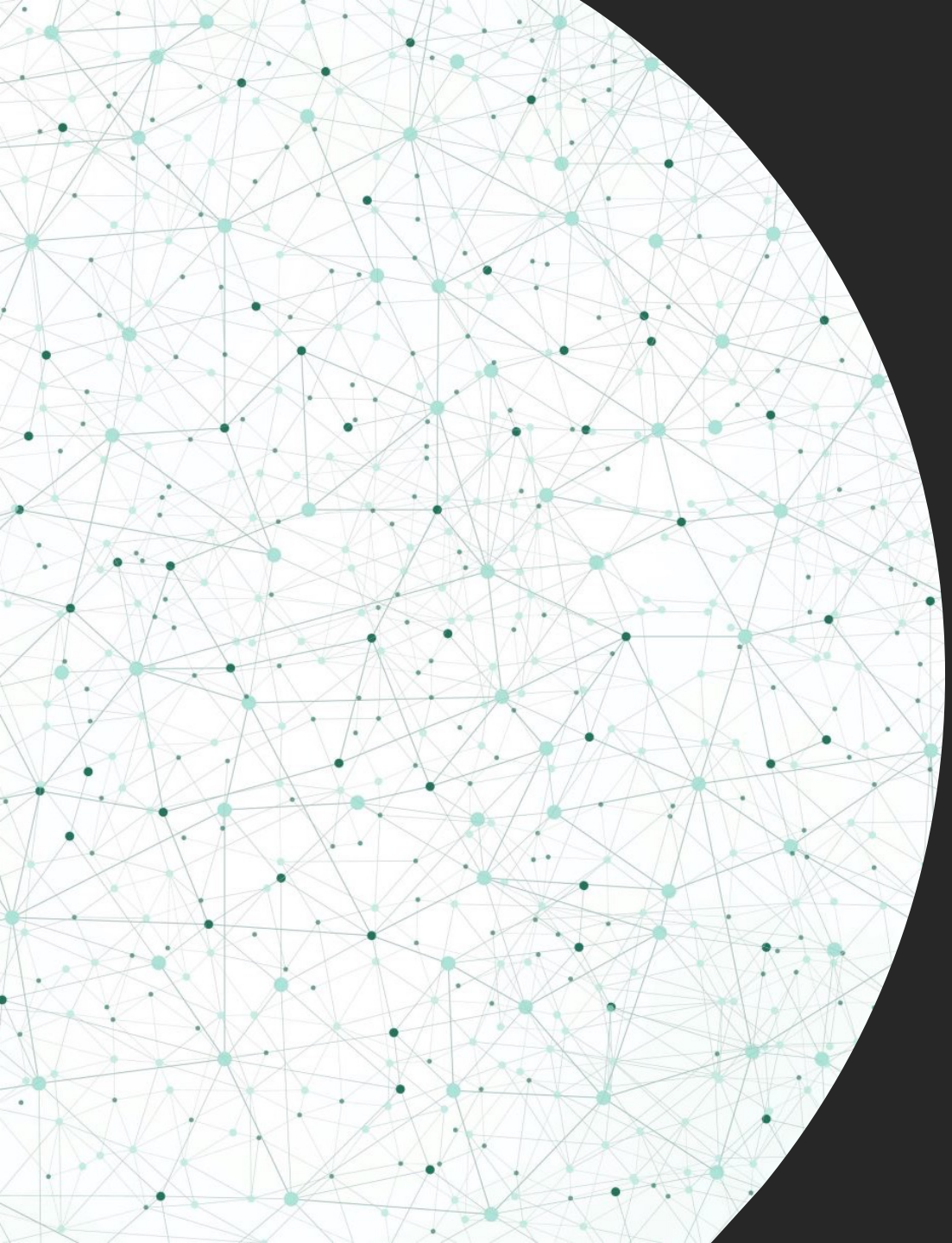


Logistic regression is all you need

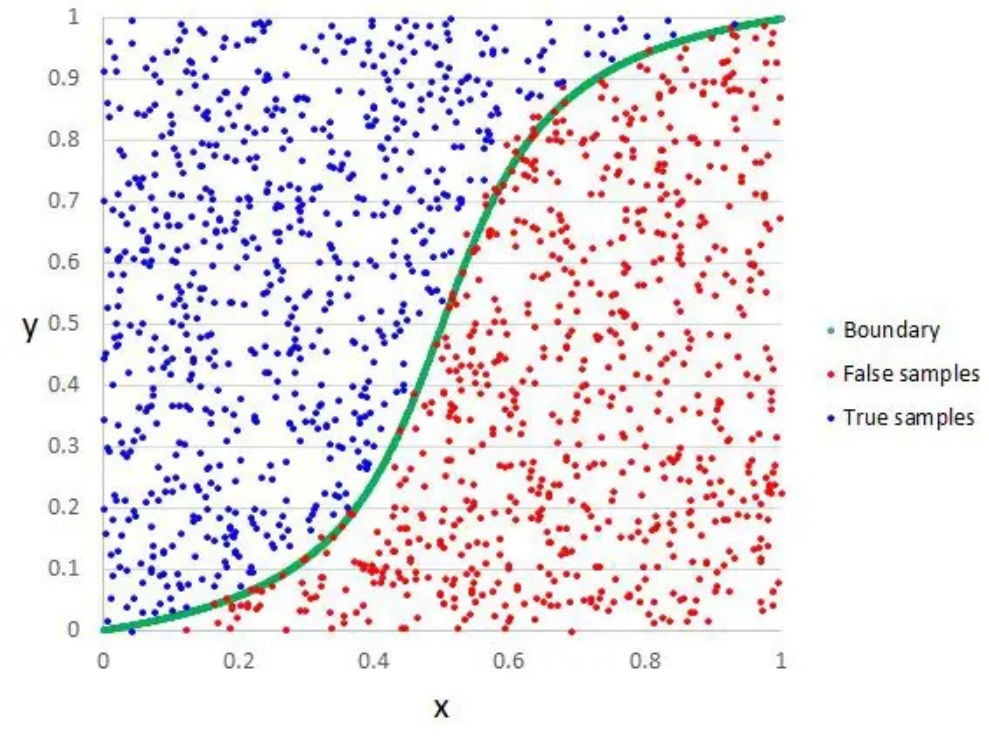
"And feature selection, but I didn't try that before the deadline…"

Thanks for the fruitful discussion

- Feifan, Linlan, Ruoqi and Mediha
- Logistic regression, feature selection, and not submitting more performant models

Method

- Bayesian search optimization in a cross-validation framework



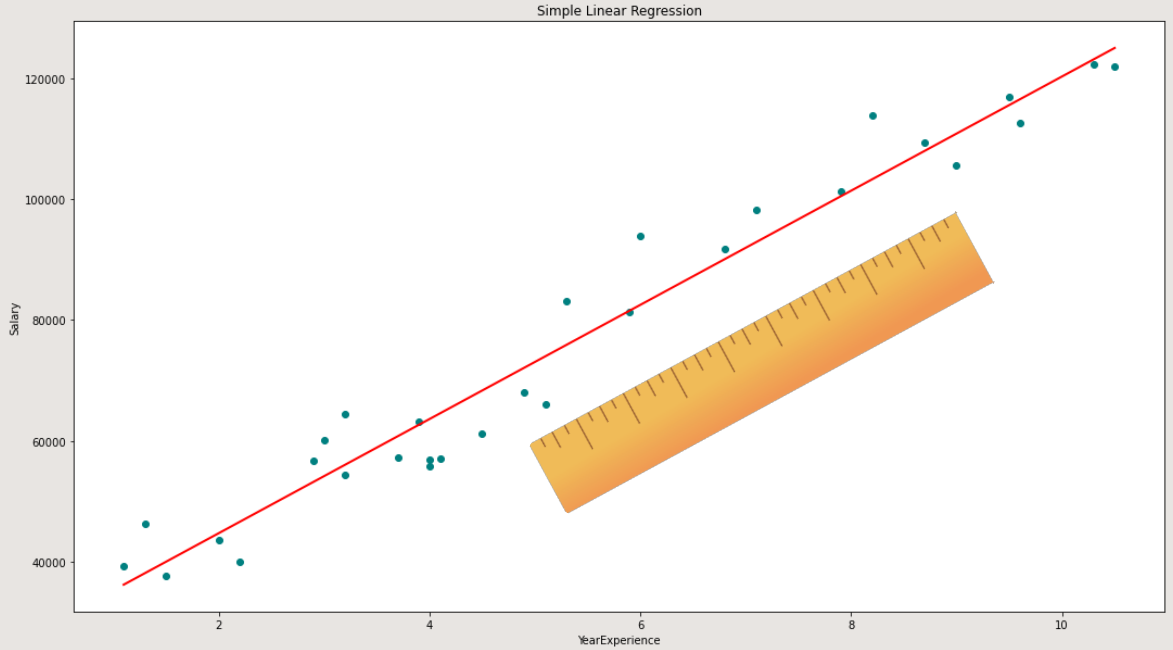
Fixed solver 'liblinear', max_iter = 10000

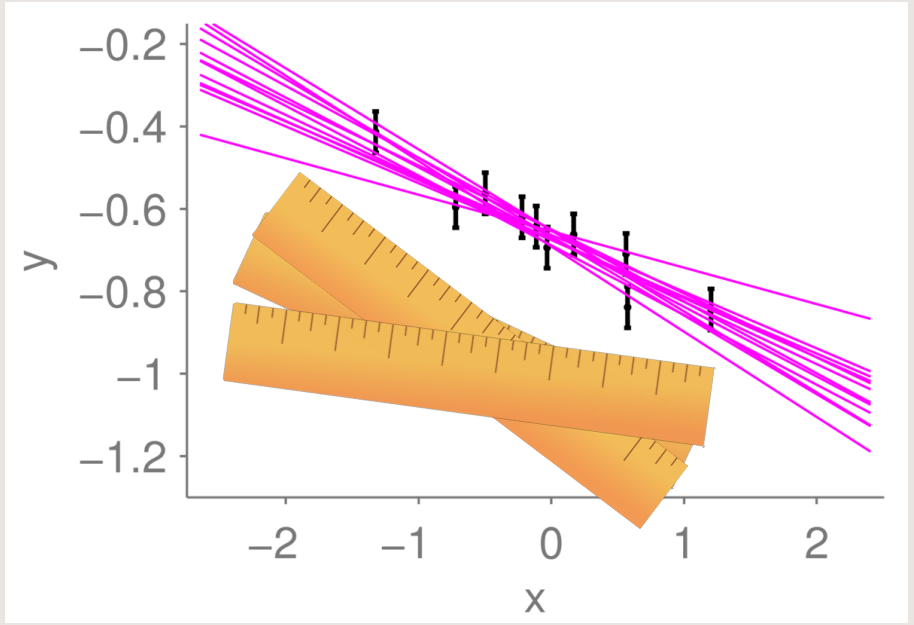
hparams:

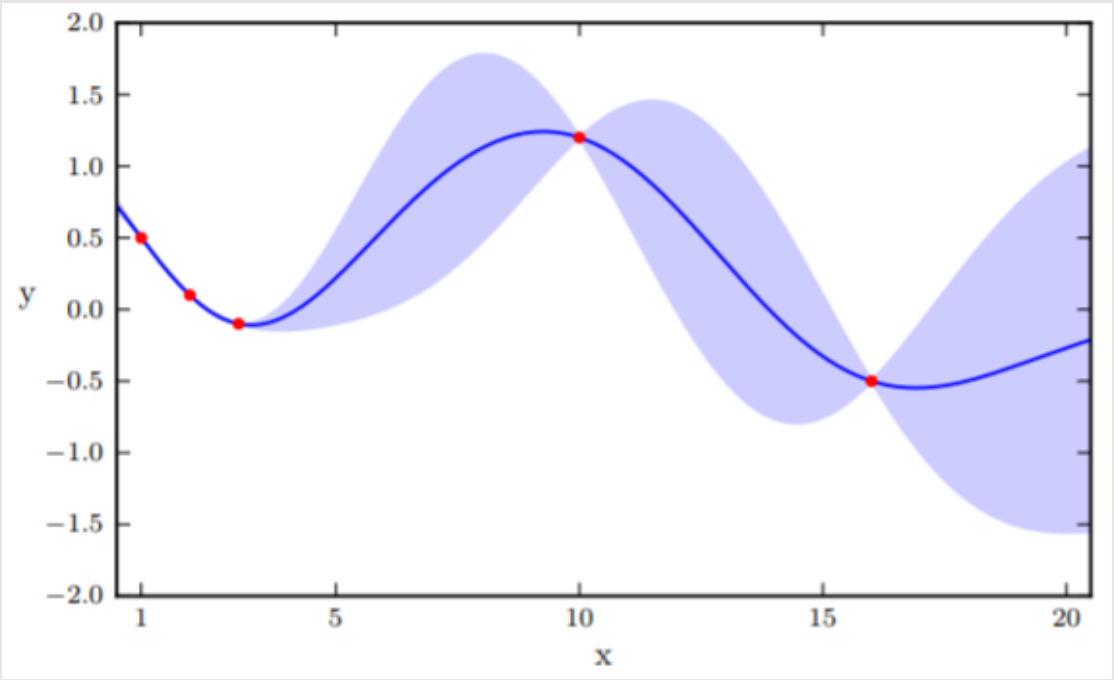
- {C: 0.037189794410786 95,
  penalty: 'l1'}

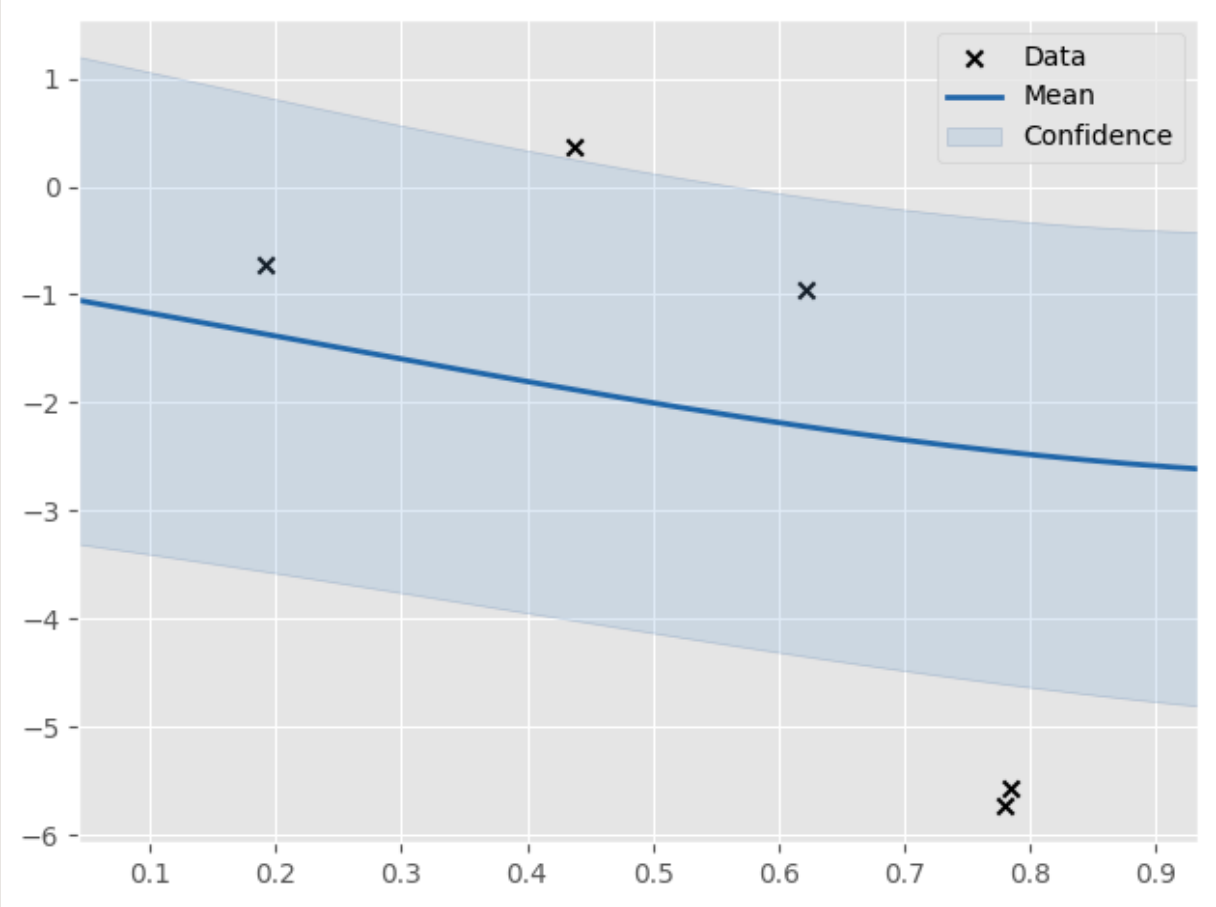
LogisticRegression (all dim)

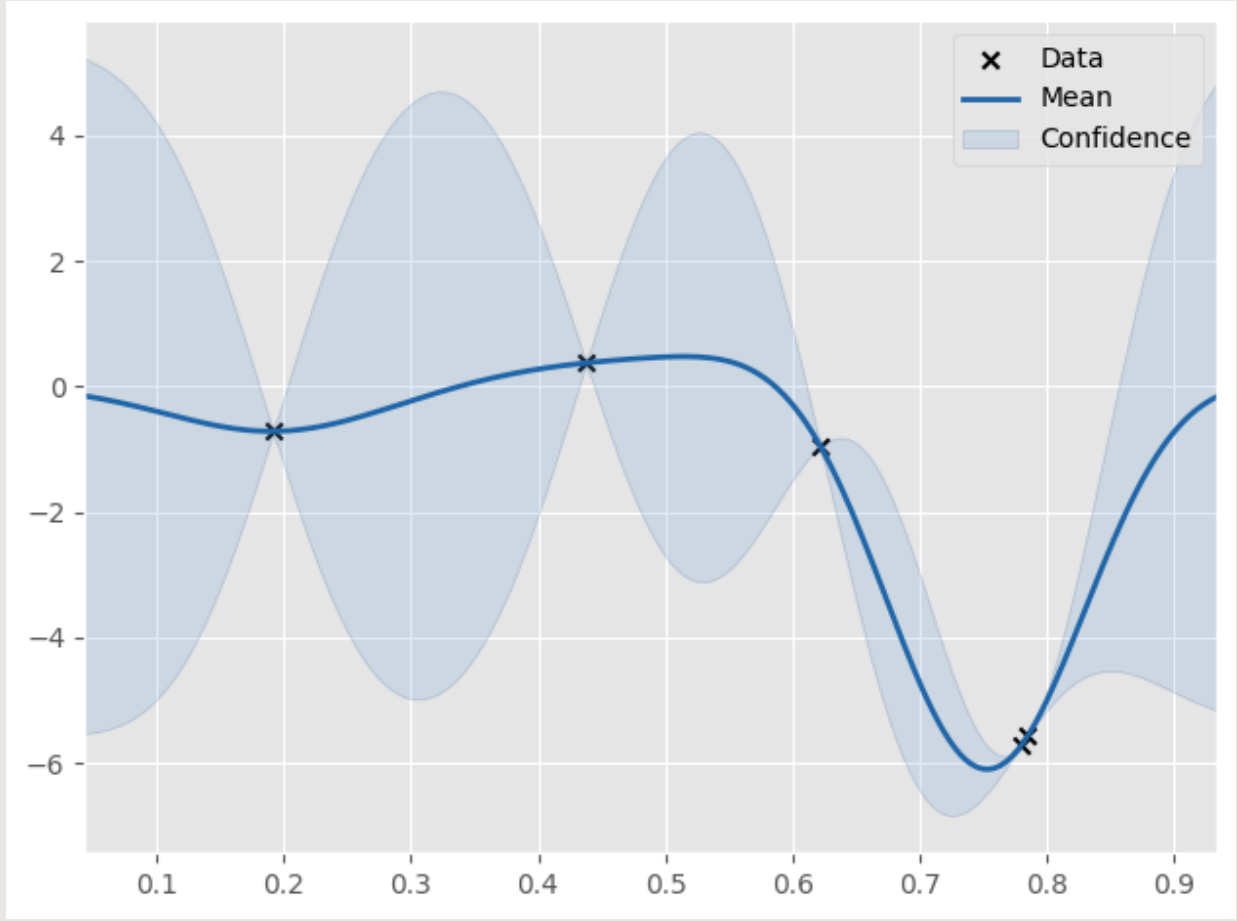
Best balanced accuracy score: 0.783687802635171


Simple Linear Regression
Salary
YearExperience
120000
100000
80000
60000
40000
2
4
6
8
10


−0.2
−0.4
−0.6
−0.8
−1
−1.2
y
−2
−1
0
1
2
x


2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
y
1
5
10
15
20
x





- Prior

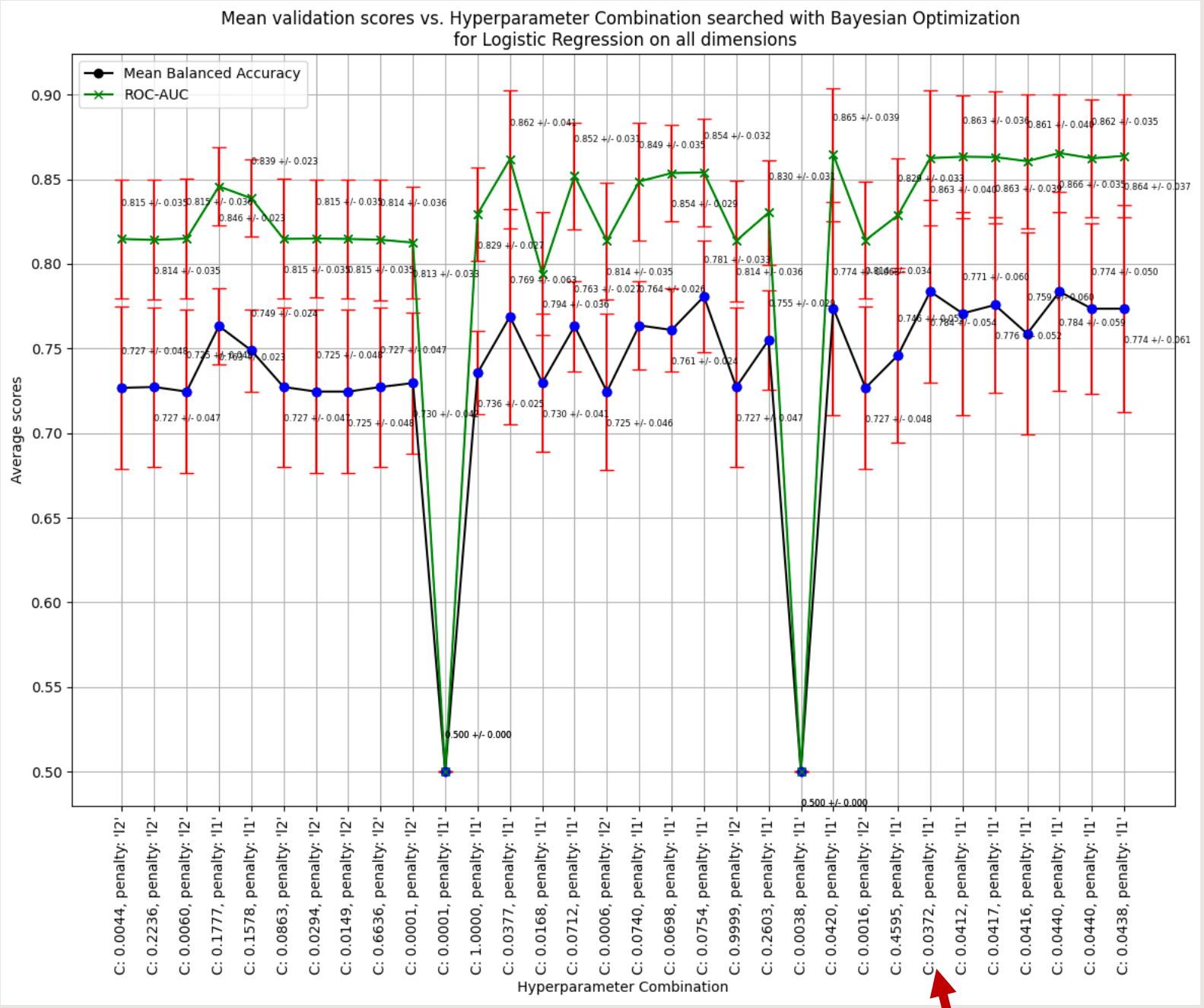
- X: hyperparameter
- Y: Metric
- Acquisition function proposes a new hparam

Why Bayes Optimization?

CUSTOMIZABLE, NON-RANDOM

MAYBE BETTER GENERALIZATION? (HIGHER TEST SCORE)

GOOD FOR RESOURCE-CONSTRAINED SETTINGS


Mean validation scores vs. Hyperparameter Combination searched with Bayesian Optimization
for Logistic Regression on all dimensions
Mean Balanced Accuracy
ROC-AUC
Average scores
Hyperparameter Combination
C: 0.0044, penalty: 'l2'
C: 0.2236, penalty: 'l2'
C: 0.0060, penalty: 'l2'
C: 0.1777, penalty: 'l1'
C: 0.1578, penalty: 'l1'
C: 0.0863, penalty: 'l2'
C: 0.0294, penalty: 'l2'
C: 0.0149, penalty: 'l2'
C: 0.6636, penalty: 'l2'
C: 0.0001, penalty: 'l2'
C: 0.0001, penalty: 'l1'
C: 1.0000, penalty: 'l1'
C: 0.0377, penalty: 'l1'
C: 0.0168, penalty: 'l1'
C: 0.0712, penalty: 'l1'
C: 0.0006, penalty: 'l2'
C: 0.0740, penalty: 'l1'
C: 0.0698, penalty: 'l1'
C: 0.0754, penalty: 'l1'
C: 0.9999, penalty: 'l2'
C: 0.2603, penalty: 'l1'
C: 0.0038, penalty: 'l1'
C: 0.0420, penalty: 'l1'
C: 0.0016, penalty: 'l2'
C: 0.4595, penalty: 'l1'
C: 0.0372, penalty: 'l1'
C: 0.0412, penalty: 'l1'
C: 0.0417, penalty: 'l1'
C: 0.0416, penalty: 'l1'
C: 0.0440, penalty: 'l1'
C: 0.0440, penalty: 'l1'
C: 0.0438, penalty: 'l1'