

Olga Whelan  
Mike Roylance  
Ling 572  
**Homework 5**

q4)

p0	x0^2 score	Number of related features	Test accuracy
baseline	0	32846	0.72
0.001	13.816	2401	0.75
0.01	9.210	3895	0.75
0.025	7.378	5223	0.756666666667
0.05	5.991	7484	0.733333333333
0.1	4.605	8499	0.746666666667

q5)

p0	x0^2 score	Number of related features	Test accuracy
baseline	0	32846	0.816666666667
0.001	13.816	2401	0.856666666667
0.01	9.210	3895	0.85
0.025	7.378	5223	0.846666666667
0.05	5.991	7484	0.82
0.1	4.605	8499	0.853333333333

q6)

Conclusions
Binarizing the features has a positive impact on performance
Increasing or decreasing the p0 amount does not have a direct correlation on performance. For instance, performance decreases from 0.025 to 0.05 but increases from 0.01 to 0.025 for non-binarized features, but decreases from 0.01 to 0.025 for binarized features.

The baseline was the lowest performing among both binarized and non-binarized features.