Le domande di Olena

1. Che ne dici se in jupyter si fa prima letura dbs e vista delle caratteristiche e poi la creazione delle costanti?
2. Ma se provassimo senza la correlazione pure per vedere i risultati come cambiano? La lasso tipo serve pure per fare la selezione delle caratteristiche che sono inutili potremmo avere buoni risultati.

1. Se ho fatto bene il coso (ma non lo so poi ti chiedo e vediamo insieme) abbiamo:

se facciamo senza correlazione

-Cross-validated Risk estimate lasso

0.019359130007616666

-Cross-validated risk estimate ridge

0.02579195886586786

Però la domanda 5 mi fa sorgere un dubbio che porta a un'altra domanda, con che alpha l’ho fatto il coso?

1. Penso serva Il plot della lasso e ridge con la cross validation , non so come mettere insieme miky e main su questo
2. Serve però il plot quello ultimo della alpha e il squared loss

In main 41 42 e 43 , ma con lasso e ridge penso . I grafici che hai fatto con la lasso e cholensky sono già con il cross validate risk estimate?

1. La parte solo di ridge regression nel main è in più?
2. Dal 44 in poi è da togliere giusto?
3. Poi si fa la PCA per il miglioramento del cross.validate risk estimate.
4. Le correlazioni con la matrice e le caratteristiche cancellate poi le vediamo dove e quando ficcare

Non togliere le colonne

Vedere le colonne che hanno meno correlazione e poi toglierle eventualmente

Prima ci usciva alpha di 0.3 nello scoring, mo non più con l’ultima modifica

Prendere il grafico di alpha e farlo più zoommato

Ricerca test error come ridurre

Vedere formula per ottimizzazione lasso

Accuracy – r2 , mse

Provare plot 45.. miky test error e train