

Homework # 8

Biostatistics 210

DUE December 7, 2021 by 2pm

0. BACKGROUND

This assignment will ask you to discuss your understanding of predictive models. A useful resource will be the lecture and the Steyerberg text which can be accessed [here](#)

1. Define the terms optimism and overfitting in the context of prediction models.
2. List some of the effects that overfitting has on a model.
3. What problems does optimism cause for model fitting?
4. Steyerberg discusses various approaches to pre-screening variables to be including in a prediction model. What are the advantages of this approach? How does this strategy mitigate overfitting?
5. What are the advantages of methods which “shrink” coefficients (e.g., LASSO, ridge regression)?