

Model Order Selection. In-class Exercise 3

EL-GY 6143 Intro Machine Learning. Prof. Sundeep Rangan

Question

A teacher tries to model the score of a child on a test as follows:

$$\text{Test score} \approx \beta_0 + \beta_1[\text{Hours studied}] + \beta_2[\text{Hours sleep before test}]$$

Below are three possible shortcomings of the model. For each shortcoming, indicate whether it would lead to (i) High irreducible error, (ii) High bias error or (iii) High variance error

- (a) The relation between test score and hours studied is not linear. After some amount of studying, the students test score will no longer improve.
- (b) The model should also include the age of the child.
- (c) The model was fit with data from one class with only 15 children.