

Part I - Variable Subset Selection

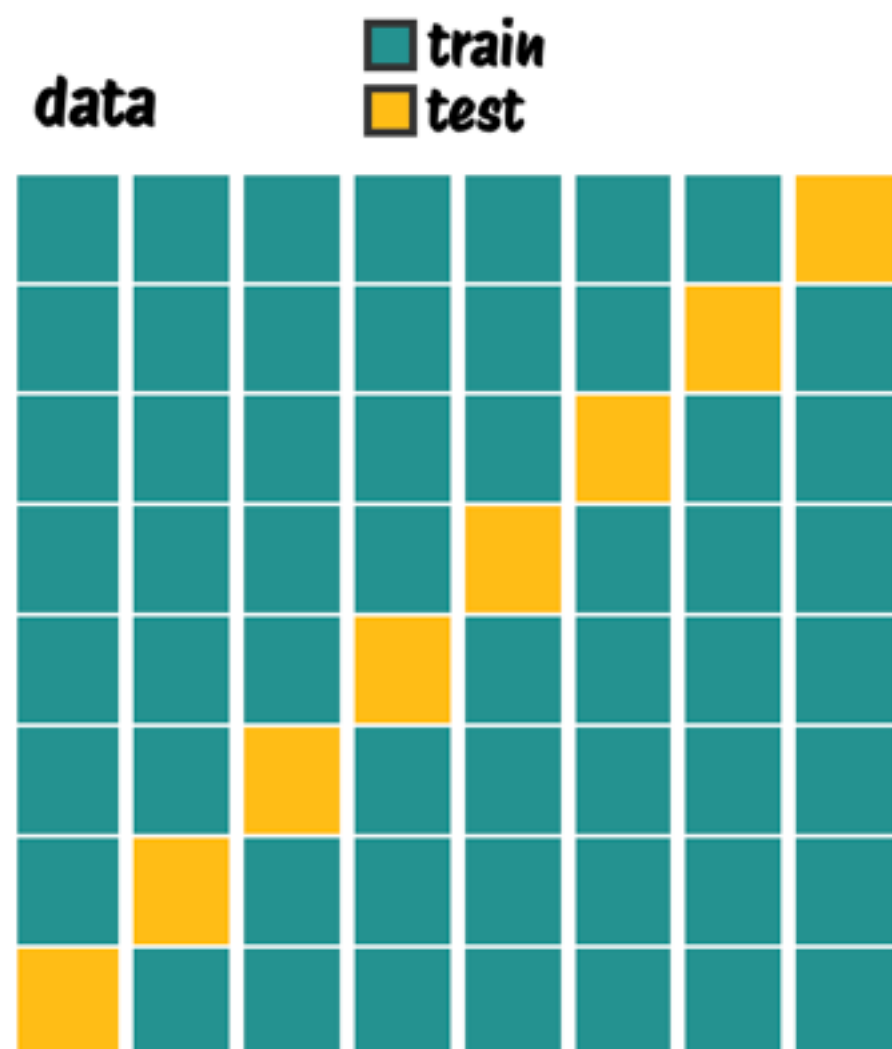
Model Selection Criteria: Validation by Prediction Error

Last week: how to use cross validation to choose a set of predictors by directly estimate prediction error using cross-validation techniques

$$\text{e.g. } \text{MSE} = \frac{\text{RSS}}{n}$$

$$\text{RMSE} = \sqrt{\frac{\text{RSS}}{n}}$$

$$R^2 = 1 - \frac{\text{RSS}}{\text{TSS}}$$



Now: indirectly estimating test performance using an approximation