3.)

(a) (i) (Partition) Divide the dataset into k equally-sized folds

(ii) Take turns leaving out 1 fold and training model on other (k-1) folds combined

(iii) Test the model predictions on the 1 left-out fold.

(iv) Aggregate the k test prediction results to form consensus prediction error for each candidate model.

(b) (i) Adv of k: • Better approximates true test error than validation set
• Less likely to give misleading/erroneous indication.

Dis: More computationally and time intensive than validation set approach.

(ii) LOOCV:

Adv of k-fold: • Less computationally and time intensive than LOOCV
• Less variance of test error than LOOCV

Dis of k-fold: • More bias of test error than LOOCV