Solution 1:

- (a) (i) Important parts:
 - Correct number of models for tuning
 - Correctly multiplying tuning models times the two learners that need tuning
 - Correctly adding $4 \cdot 10$ for learner comparison ()

$$\# \text{models} = \underbrace{4 \cdot 10}_{\# \text{ models outer resampling}} + 2 \cdot \underbrace{10 \cdot \underbrace{5 \cdot 200}_{\# \text{ models for one tuning}}}_{\# \text{ models for both tunings}} = 20040$$

- (ii) We would select the k-NN (k-Nearest Neighbors) learner since it achieves the best values for the AUC.
- (b) Less data for training leads to higher bias
 - More data for training and less data for evaluation lead to higher variance
- (c) Are the following statements true or not, explain your answer in one sentence.
 - (i) True, using 3-fold cross-validation leads to smaller train sets and therefore we are not able to learn as much as for, e.g., 10-fold cross-validation.
 - (ii) False, the outer loss doesn't has as much restrictions as the inner loss, e.g. the outer loss doesn't has to be differentiable.