

 A. One set for both selection and evaluation

B. Separate selection and evaluation sets (averaged over 100 reps)
C. Separate sets

weighted average (averaged over 100 reps)