

# Summary

This week, we learned about two important resampling methods: cross-validation and the bootstrap. These techniques help us evaluate the accuracy of models by estimating prediction error and assessing the uncertainty of parameter estimates. We also discussed how these methods can guide model selection and avoid overfitting.

## Concepts

- **Training vs. test error:** Training error is often lower but misleading; test error is what truly matters when evaluating model generalization.
- **Hold-out validation:** The data is split into a training set and a validation set. The model is trained on one and evaluated on the other to estimate performance.
- **K-fold cross-validation:** The data is divided into  $K$  equal parts. Each part takes a turn as the validation set, and the model is trained on the remaining  $K-1$  parts. This helps reduce variance in the error estimate.
- **LOOCV (Leave-One-Out Cross-Validation):** A special case of K-fold where  $K$  equals the number of observations. It uses nearly all the data for training but can have high variance.
- **Choosing K:**  $K = 5$  or  $10$  is typically a good balance between bias and variance in the error estimate.
- **Bootstrap:** A method for estimating the accuracy of parameter estimates by repeatedly sampling with replacement from the dataset and re-estimating the model.
- **Use of bootstrap:** It's especially helpful for estimating standard errors and confidence intervals, not ideal for estimating prediction error due to overlap in training and test sets.
- **Right vs. wrong cross-validation:** Cross-validation must include every step of the modeling process, including feature selection, to avoid biased performance estimates.

## Uncertainties

One thing I'm still unsure about is when it's appropriate to use the bootstrap instead of cross-validation, especially since bootstrap tends to underestimate prediction error. I also found it a bit confusing how to interpret the bias-variance trade-off in choosing  $K$  for cross-validation. Lastly, while we talked about using cross-validation for model selection, I'd like to see clearer guidance or examples on how to apply it when tuning hyperparameters or comparing models with different levels of complexity.