Module 9: Model Selection and Regularization

Glossary

Cross-Validation

A family of methods to validate a model using only the existing data from the model; these techniques include K-fold cross-validation, leave-one-out cross-validation, and others

Hyperparameter

Any parameter used to control the learning that is set before training

K-Fold Cross-Validation

A specific type of cross-validation where the number of ‘folds’ is selected, and one ‘fold’ becomes the test set, while all other ‘folds’ become the training set

LASSO

A regression analysis method that performs both variable selection and regularization in order to enhance the prediction accuracy and interpretability of a model

Regularization

Techniques used to optimize machine learning models by minimizing the adjusted loss function and preventing overfitting or underfitting

Ridge Regression

A method of estimating the coefficients of multiple-regression models in scenarios where independent variables are highly correlated; this method performs L2 regularization

Sequential Feature Selection

A family of algorithms that are used to reduce the number of features in a model

Standardization

A way to transform data to make features in the data approximately the same scale