PS9

Caden Finley

April 2023

1 Code

library(tidymodels) library(glmnet) library(magrittr) library(readr) set.seed(123456)

housing ¡- $read_table("http : //archive.ics.uci.edu/ml/machine - learning - databases/housing/housing.data", col_names = FALSE)names(housing) < -c("crim", "zn", "indus", "chas"$

$housing_split < -initial_split(housing, prop = 0.8)$

$housing_train < -training(housing_split)housing_test < -testing(housing_split)$

$housing_recipe < -recipe(medv ., data = housing)convertoutcomevariabletologsstep_log(all_outcomes())con$ $zn : indus : rm : age : rad : tax : ptratio : b : lstat : dis : nox)createsquaretermsofsomecontinuousvariabless$ 6)

run the recipe $housing_prep < -housing_recipehousing_train_prepped < -housing_prephousing_test_prepped < -housing_prep$

create x and y training and test data $housing_train_x < -housing_train_preppedhousing_test_x < -housing_test_preppedhousing_train_y < -housing_train_preppedhousing_test_y < -housing_test_prepped$

dimension of $housing_traindatais404rowsand14columnshave1morexvariablethantheoriginalhousingdata$

LASSO model log median house value

$tune_spec < -linear_reg(penalty = tune(), tuningparametermixture = 11 = lasso, 0 = ridge)set_engine("glmnet")set_mode("regression")defineagridoverwhichtotrydifferentvaluesofla$ $-grid_regular(penalty(), levels = 50)10 - foldcross - validationrec_folds < -vfold_cv(housing_train_prepped, v = 6)optimal_lambda_lasso < -tune_spec$bestTune$lambdacat("Optimalvalueo$ ", $optimal_lambda_lasso$, "")

ridge regression model library(caret) $ridge_model < -train(log_median_house_value ., data = train_data, method = "glmnet", trControl = trainControl(method = "cv", number = 6), tuneGrid = expand.grid(alpha = 0, lambda = 10^seq(-5, 5, by = 0.1)))$

Optimal value of lambda for Ridge $optimal_lambda_ridge < -ridge_model$bestTune$lambdacat("Optimalvalued$ ", $optimal_lambda_ridge$, "")