

Build two models to predict the variable `cnt` from the `bike` dataset found on my website. We'll compare their ability to predict future data using k-fold cross validation.

Specifically, use any combination of numerical and categorical variables you want to create two multiple regression models each of which predicts `cnt`. We will use k-fold cross validation to determine which model is likely (but not guaranteed) to predict better on future data.

1. Load the library `caret`.
2. Read in the `bike` dataset from my website. Read the help file associated with this dataset.
3. Write yourself a Mean Squared Error function. This function should have signature `MSE(y, yhat)` and should return a single number: $MSE = N^{-1} \sum_{n=1}^N (y_n - \hat{y}_n)^2$.
4. Call the function `caret::createFolds` on the vector `cnt` and store the output in a variable named `folds`.
5. Recall how to extract a vector from a list that contains multiple vectors.
6. Create two vectors `mse_mr01` and `mse_mr02` using the function `rep`. These vector should be filled with K , the number of folds, `NA`s.
7. Write a for loop around each fold in your variable `folds`. Within each iteration of the for loop you should
 - (a) Create two datasets: `training` and `testing`. The dataset `training` should contain all data but the current fold. The dataset `testing` should contain only the data from the current fold.
 - (b) **For multiple regression model i**
 - i. Train your first multiple regression model to the variable `cnt` based on the `training` dataset using the variables of your choice for the first model.
 - ii. Predict the values of `cnt` for the observations of your testing dataset, using the model trained in the previous step. Let's refer to these predicted values as `yhat_mr0i`.
 - iii. Call your function `MSE` on `cnt` from the `testing` dataset, `y`, and your first model's predicted values `yhat_mr0i`. Store this value into a vector `mse_mr0i`.
8. After the for loop, find the mean of the vectors `mse_mr01` and `mse_mr02`.
9. Which mean MSE is smaller?
10. Are these MSEs reasonable quantifications of your model's ability to predict future (not your current) data? Why or why not?
11. Why is multiple regression model appears better at predicting future data?