Rolling update mechanism

Forecast Framework

Two types of forecasts are constructed: “base forecast” which only includes daily overall sale features and a more complete model called “advanced forecast” which includes both daily overall sale features and daily discount features. Every machine learning model is fitted for these two types of forecasts and compares their accuracy where the only difference between them is whether discount features are included.

Base forecast framework

Discount features are not included in the base forecast. In the base forecast, overall sale on day t is the function of the overall sales in the past five days[3].

(1)

represents the overall sale on day t.represents the different machine learning models chosen.

Advanced forecast framework

The advanced forecast includes discount features. In the advanced forecast, overall sale on day t is the function of the overall sale, and discount features in the past five days.

(2)

represents the overall sale on day t.represents the discount features on day t such as quantity discount, direct discount, and overall discount. represents the different machine learning models chosen.

Training, Cross-validation, and Out-of-sample evaluation

**In-sample training** data and out-of-sample testing data are split to evaluate out-of-sample accuracy. To be specific, training data and testing data are split in a rolling mechanism.

Assuming that dayt is the testing data, its corresponding training data is:

Train set = [dayt-1, dayt-2, dayt-3, …, dayt-m] (3)

In Equation (3), m=10.

Assuming that the index of rolling is i, then for rolling i, the testing data and training data are shown below:

Train set = [dayi, dayi+1, dayi+2… dayi+9] (4)

Test set= [dayi+10] (5)

**Cross-validation** is used to select hyperparameters in machine learning models. Ten-fold cross-validation with five repeats is used to evaluate the performance of each hyperparameter. The training set is randomly divided into ten subsets of the same size, nine sets are used to train for hyperparameter selection and 1set is used to test the performance of hyperparameters. Each subset will be treated as testing set at least once.

The overall performance of the model is the average of ten subsets. Next, retain the hyperparameters with the best performance and then estimate parameters with the entire training set. After doing this, the result is to get the best model for each training set.

**Out-of-sample evaluation**. The best hyperparameter and parameter selected are retained for each training set during the process of constructing the forecast for the out-of-sample testing set. When forecasting the overall sale on day t, pass data and models are used from dayt-10 to dayt-1 as input to the selected best model and then get the forecast result. To be more specific, in the next round of forecasting on dayt+1, pass data and models from dayt-9 to dayt are carried out. This kind of rolling update mechanism have relatively good performance in small-scale data with high dimension.