Technical Description of Forecast Models for Toronto Island Ferry Ticket Counts

Toronto Island ferry ticket counts data contain redemption count and sales count from May 2015 through June 2025. This project contains two sections, one is for redemption count forecast and the other sales count forecast.

Section 1: redemption count forecast

Part 1) Basic check of redemption count data

I use timestamp and redemption count. I check the data type of columns, duplicates, missing values, and outliers. The Dickey-Fuller test verifies the data is stationary. I generate the autocorrelation plots and do trend and seasonality decomposition.

Part 2) Model selection

Notice that the residuals center at 0 with a seasonal pattern, so I plan to try 10 statistical methods. Because the data contains a great number of outliers, MAE is the best metric to select models. With the help of the train set and test set, I conclude the Holt Winters model is the best candidate.

Part 3) Compare Holt Winters model to given baseline

I compare Holt Winters model to baseline based on MAE, RMSE, MAPE respectively. Both the metrics and the plots confirm that Holt Winters gives more accurate forecasts.

Section 2: sales count forecast

Part 1) Basic check of sales count data

I use timestamp and sales count. The basic check is similar to the description in Section 1.

Part 2) Model selection

I consider three groups of models: statistical models (StatsForecast from NIXTLA), conventional machine learning models (scikit-learn), and deep learning models (NeuralForecast from NIXTLA).

Assume redemption count is not given in Section 2 for the sake of comparison fairness because redemption count and sales count are strongly correlated (corr=0.98).

I use MAE as the metric to select the promising model for the same reason aforementioned in Section 1.

a) I compare the MAEs from the 10 models of StatsForecast from NIXTLA, and find Holt Winters is the best among them. I use Holt Winters as my baseline.
b) I compare the MAEs of the 14 models from scikit-learn. Extra Tree outperforms the rest.
c) I compare the MAEs from the five models of NeuralForecast. NBEATS beats the other four.
d) Extra Tree (ranks 1st) is superior to both Holt Winters (ranks 2nd) and NBEATS (ranks 3rd), thus I fine tune Extra Tree.

Part 3) Fine Tune Extra Tree

a) During the feature engineering procedure, I introduce new features by temporal embedding and time delay embedding.
b) I split the full training into a smaller train set and a validation set.
c) I create pipelines to transform the data using Min-Max scaler.
d) I shrink hyperparameters space using Bayesian Optimization, and then find the best set of hyperparameters using Grid Search. Learning curves indicate the tuned model does well.
e) The tuned model is evaluated by the test set.
f) To show the performance difference between Extra Tree and Holt Winters, I plotting two plots respectively:
    i) Observed sales count and Extra Tree predictions
    ii) Observed sales count and Holt Winters predictions

Conclusion: the chosen model is promising and ready to forecast the future counts.