

Recruit Restaurants - Forecasting - Prediction Analysis - Optimal ARIMA models, and prediction on a week of days with optimal ARMA models

- With our 7-day differenced time-series of the daily average visitors of each genre, the next step is to fit it to optimal ARIMA models. Based on considering different AR and MA orders, and comparing the prediction performance with two metrics on a test week

Optimal Parameters for Different Genres - MAE

Genre Name	p	q	MAE
Okonomiyaki/...	5	3	2.66
Japanese food	6	3	1.37
Izakaya	5	3	1.65
Italian/French	7	3	2.58
Dining bar	6	1	0.64
Bar/Cocktail	7	1	0.96
Cafe/Sweets	4	2	1.79
Other	6	1	1.88
Asian	5	1	12.4
International cuisine	4	2	5.66
Yakiniku/Korean food	5	2	1.25
Creative cuisine	6	3	3.8
Western food	2	3	2.65
Karaoke/Party	6	3	9.86

Optimal Parameters for Different Genres - RMSE

Genre Name	p	q	RMSE
Okonomiyaki/...	7	2	3.46
Japanese food	7	3	1.91
Izakaya	6	2	2.0
Italian/French	7	3	3.0
Dining bar	6	2	0.86
Bar/Cocktail	7	3	1.2
Cafe/Sweets	7	2	2.65
Other	6	1	2.09
Asian	4	2	14.61
International cuisine	4	2	6.5
Yakiniku/Korean food	5	2	1.87
Creative cuisine	6	3	4.59
Western food	2	3	3.22
Karaoke/Party	6	3	11.45

ARIMA(p,d,q) Model - with statsmodels package (Python)

Based on utilizing $d=7$, i.e. a differencing of lag 7
To find optimal p and q, two metrics, MAE and RMSE, are evaluated on
on a test set consisting of a week, given an ARIMA model trained
on a training set.
For the search of optimal p, q, all ARIMA models for $p=1,2,\dots,7$
and $q=1,2,\dots,7$ have been considered

In above tables, the optimal choices of p and q for
each genre is presented.
MAE is the mean of the residuals, with respect to true and predicted
time series values.
RMSE is the square root of the mean of the residuals squared

- With our optimal models, based on MAE, the last step is to evaluate the models on a week of
days, the validation set, to really measure the performance of our models

