Recruit Restaurant Visitor Forecasting (Kaggle)

Danijel Kopčinović

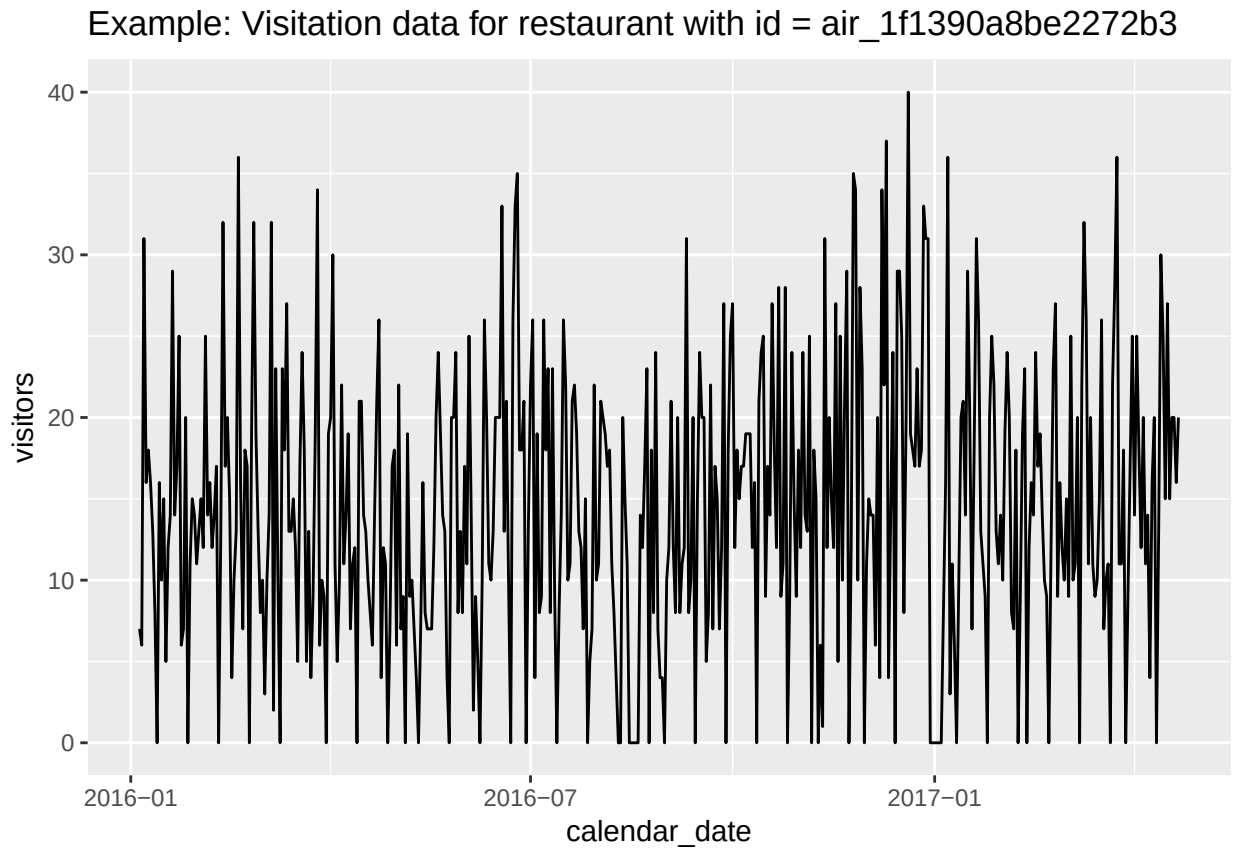
Description

This is a problem put by Recruit Holdings, owner of a line of restaurants. The challenge is to use reservation and visitation data to predict the total number of visitors to a restaurant for future dates. This information will help restaurants be much more efficient and allow them to focus on creating an enjoyable dining experience for their customers. Problem link: https://www.kaggle.com/c/recruit-restaurant-visitor-forecasting/.

Data

The data used to solve the problem consists of:

- air_reserve.csv: reservations from the air reservation system (92.378 entries).
- hpg_reserve.csv: reservations from the hpg reservation system (2.000.320 entries).
- air_store_info.csv: restaurant information from the air reservation system (829 entries).
- hpg_store_info.csv: restaurant information from the hpg reservation system (4.690 entries).
- store_id_relation.csv: relation matching restaurants between two reservation systems (150 entries).
- air_visit_data.csv: visitation data from the air reservation system (252.108 entries).
- date_info.csv: basic date information about the holidays and days of week (517 entries).

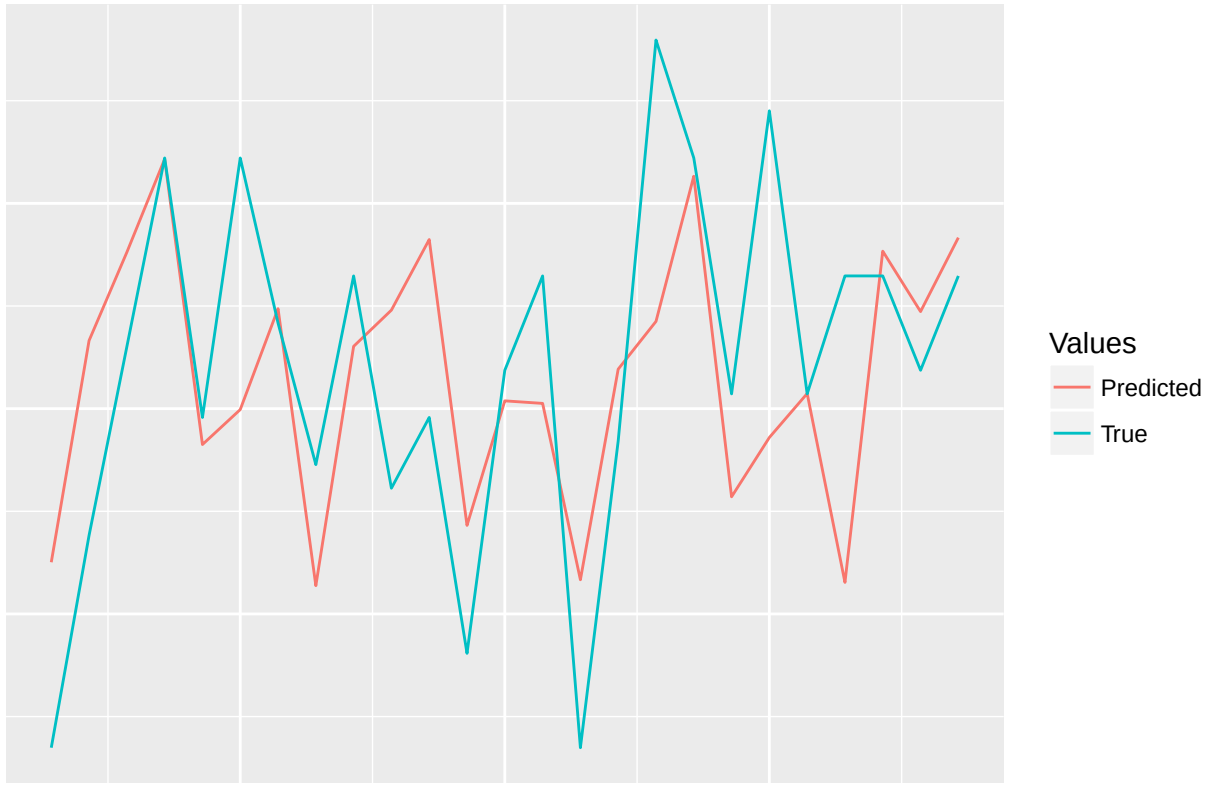


Modelling

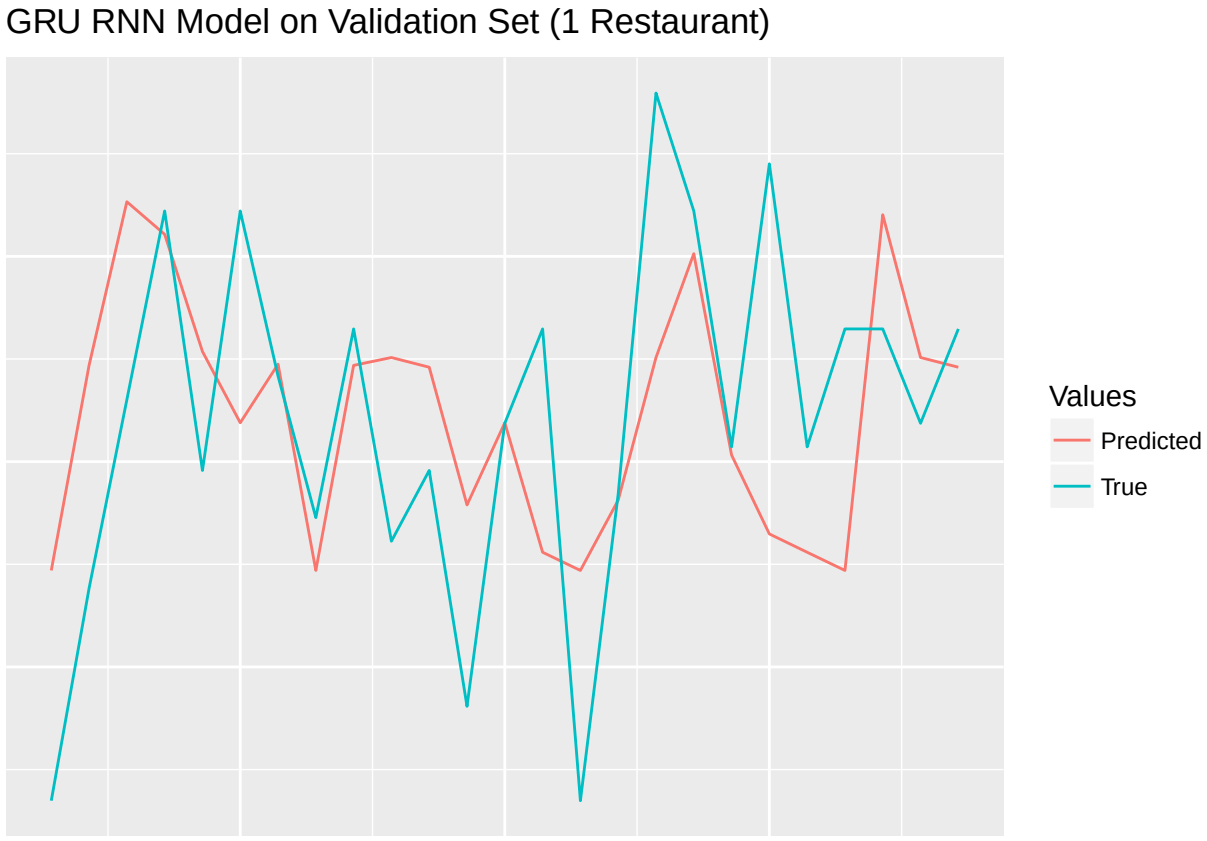
Since the goal is to predict (forecast) visitation for each restaurant, we will use time related data that we have: reservations from both systems, past visitation data and date information (to match days of the week and holidays with the effect on visitation).

We will use three approaches:

- ARIMA (autoregressive integrated moving average)
- neural network regression
- GRU-RNN (recursive neural network)

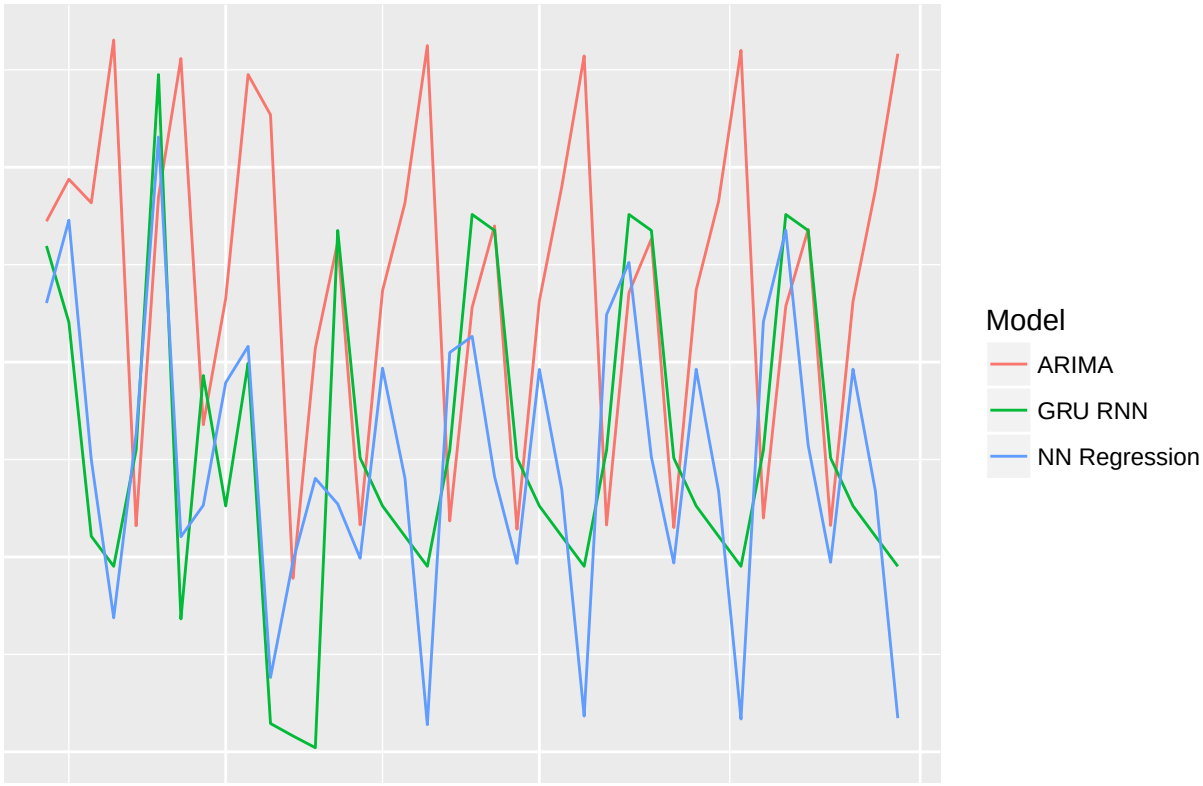

ARIMA Model on Validation Set (1 Restaurant)
Values
Predicted
True


NN Regression Model on Validation Set (1 Restaurant)
Values
Predicted
True



Forecasting

All 3 models seem to work quite well with the data so we will make forecasts with all models and check against new data (results unknown to us - used for evaluating the applications on Kaggle).



The final valuation result was 0.499 (root mean squared logarithmic error) where the best result was 0.465.

1382	▼179	mobaohua	0.499	3	2mo
1383	▼179	jithin	0.499	9	1mo
1384	▼34	Danijel	0.499	28	1h
1385	▼180	DY	0.499	37	7h
1386	new	Bhupender Singh	0.499	6	7h
1387	▼181	Qing Liu	0.499	24	2mo
1388	▼181	qkn	0.499	14	1mo

Contact and Info

Danijel Kopčinović, IT Market

Mail: danijel.kopcinovic@itmarket.hr

Tel: +385956472127