Predicting Restaurant Daily Customer Demand

Vasavi Kullanakoppal

Bellevue University

DSC 680: Applied Data Science

Prof. Matthew Metzger

July 12, 2025

Predicting Restaurant Daily Customer Demand

Business Problem

Restaurants usually experience unpredictable demand from customers, which appears as inefficiencies such as overstaffing, food wastage, or poor customer service during peak demand periods. The project aims to address the problem by building a predictive model that can accurately forecast restaurant demand on any day.

Background/History

Demand forecasting in the food service industry has traditionally had its basis in manager expertise and past averages. As booking systems, POS, and external contextual information such as holidays and weather become more readily available, there exists the potential for using advanced machine learning methods to yield better prediction accuracy. Data in the Recruit Restaurant Visitor Forecast dataset in Japan can provide a solid basis for model assessment.

Data Explanation

Data was obtained from Kaggle's Recruit Restaurant Visitor Forecasting dataset. Some of the key files include:

- air_visit_data.csv, which includes historical daily visitor counts for restaurants
- date_info.csv, which includes calendar facts such as day of week and holiday indicators
- air_store_info.csv, which houses restaurant location and restaurant type information
- air_reserve.csv, which contains reservation information like reservation and visits dates

Data preparation involved formatting date fields, aggregating reservation information into the daily granularity, combining multiple datasets into one based on date and store, creating lag and rolling average features, and converting categorical features such as restaurant genre and location.

The main columns employed are:

- visit_date = Visit Date
- visitors: Number of real customers
- holiday_flg: 1 if holiday, 0 otherwise
- day_of_week: Day of the Week
- num_reservations: Number of reservations
- Total_reserved_visitors: Total number of expected visitors through reservations
- lag_1, lag_7: One-day and seven-day lagged visitors
- reserve_ratio: Ratio of reserved to actual visitors

Visual content includes:

- A line graph of total daily visitors, displaying weekly seasonality and trend in growth
- Bar and box plots showing variation by day of the week and holiday effect
- A heatmap of weekday average visitors per store
- A scatter plot correlation between real visits and bookings
- Forecast overlays of actual, Prophet, and XGBoost forecasts

Methods

There are two models developed in order to address this issue. Prophet, which is a univariate time series model, served as the baseline and captured overall trend and seasonality. XGBoost Regressor, which is a machine learning model, was employed using all engineered features including lagged demand, reservations, and calendar features. Model performance was evaluated using MAE (Mean Absolute Error) and RMSE (Root Mean Squared Error).

Analysis

Model	MAE	RMSE

Prophet	8.59	10.04
XGBoost	0.51	0.82

Prophet achieved an MAE of 8.59 and RMSE of 10.04. It could replicate overall trends but did not react to short-term variability and outliers. XGBoost performed significantly better with an MAE of 0.51 and RMSE of 0.82 and produced accurate and responsive predictions based on multi-source features. Feature importance estimation identified reservation data, recent guest history, and calendar data as the most predictive.

Assumptions

The project assumes all reservation and visitor data are accurate and reflective of demand. It assumes there are always holiday impacts on visitor patterns and that weather has minimal influence in this dataset, which did not factor into the analysis.

Restrictions

Only Japanese restaurants were included in the sample, which may limit generalizability. Data at some stores was incomplete or irregular. Weather and potentially relevant local events were omitted.

Challenges

The hardest problems were handling missing values and outliers, aligning data between sources by date and store id, and avoiding data leakage when creating lag features.

Future Applications/Additional Uses

The model can be scaled across multiple restaurants and locations or areas. Future models can be extended with external influences such as weather, events happening in the area, and advertising campaigns. An API for real-time prediction can be integrated with the restaurant management system.

Recommendations

Use XGBoost or any similar machine learning algorithms for short-term demand forecasting.

Implement a retrain and monitoring pipeline for the same on a daily basis. Explore using time series trends as inputs along with machine learning to create better-accurate hybrid models.

Implementation Plan

In using this solution:

- Daily data imports for bookings, calendar, and guest logs
- Retrain the XGBoost model each week using new data
- Deploy the model through an API or an interactive dashboard for restaurant managers.
- Regularly track model performance and update and expand features as needed

Ethical Assessments

Fair practices include avoiding the use of discriminatory labor scheduling practices based solely upon anticipated demand. Customer and reservation information must be anonymized to ensure privacy. Care must be exercised in avoiding institutionalized discrimination, i.e., consistently understaffing on low demand days.

Conclusion

The project demonstrates that a machine learning algorithm like XGBoost can make accurate and reliable daily demand forecasts for restaurants. Because it is capable of managing multi-source data integration, it is highly flexible and realistic for real application in operation. While Prophet gave a decent baseline and covered long-term trends, it was not suitable for high-resolution, short-term forecasted applications.

This evidence gives validity to the integration of such a prediction system into restaurant management activities. It would maximize data-driven planning, minimize wastage of resources, and maximize overall efficiency in services.

References

Kaggle. (2018) *Recruit Restaurant Visitor Forecasting*.

https://www.kaggle.com/competitions/recruit-restaurant-visitor-forecasting