Food Delivery Duration Prediction Business Summary
Techincal Details Excluded

Evie Xinqi Guo

Overview of the challenge

Food Delivery company connects consumers with their favorite resturants through machine-learning-backed service. Once the order arrives late, the convinence due to delivery now becomes frustration.

To ensure that the prediction of delivery time closely reflects the actual delivery time, information from the three sides involved - the store, the dashers' capacity, and the consumer's order - all needs to be taken into consideration.

Conceptually, the estimated delivery duration after the customer placing an order can be divided into these components (it's not necessarily a chain of events since some can be done in paralell):

- The time that a store takes to receive an order from Food Delivery company
- The time for a store to repare for an order
- Assigning the order to the right dasher
- The dasher travels to the store
- The dasher picks up order and goes to the destination
- The time of dashing trying to find the specific location of delivery after getting off delivery vehicle

Overview of the approach

- New features like "Weekday or Weekend" and "Day or Night" for are created for each delivery job based on exisiting date-and-time data to extract information useful for time prediction. These features makes the model more accurate since it captures the clustered effects like stores' kitchens being busier during weekends than weekdays, and delivery jobs with daylight being eaiser and faster than they are after dark.
- Other potential new features (e.g. average parking time at a specific lcoation) that could further improve the model are raised and discussed below.
- The potential models of time prediction are evaluated both by the mean of absolute errors with the actual delivery time and a categorical metric that better fulfills the business goal. Specifically, the categorical metric separates delivery predictions based on their over/underestimation of the actual duration. And there are 6 ordinal slots that ranks each delivery prediction from "Very late - over 10 min" to "Very early - over 10 min". When interpreting and evaluating the model performance, the lateness (underestimating the delivery) are penalized more than early estimations (overestimating the delivery time), and very late/early deliveries should be viewed as stronger signals of call for model improvement than being slightly late/early.

Models used

- A linear regression model was created first. The model does well in keeping the overall late delivary within 15%, and the very-late (over 5 minutes) deliveries is around 6.9%. But the RMSE score is high - 1672.92. Note: Overall lateness is defined as the proportion of late delivery, which ranges from very late (over 10 min) to about on time (late within 5 min).

Error Evaluation Linear Regression Model Early Late Category

Early Late Category	Frequency
Early more than 10 Min	0.677
Early 5-10 Min	0.102
Early within 5 Min	0.072
Late within 5 Min	0.050
Late 5-10 Min	0.031
Late Over 10 Min	0.069

- A mixed effect model with the store_id as random effects (meaning that we can later take these irrelevant clustering effect out to make our estimation of the meaningful variables more accurate). The mixed effect model outperformed the linear regression on both the RMSE (1627.13) and the lateness (very late: 6.3%, overall lateness: 14.1%).
- A XGBoost model also performs well with a low "very late" (6.8%) delivery and slightly higher overal lateness (14.9%) than the mixed effect model . Although the cross

validation RMSE of the XGBoost model (1478.721) is much lower.

Error Evaluation XGBoost Model Early Late Category

Early Late Category	Frequency
Early more than 10 Min	0.681
Early 5-10 Min	0.098
Early within 5 Min	0.072
Late within 5 Min	0.048
Late 5-10 Min	0.033
Late Over 10 Min	0.068

- As those the XGBoost and the mixed effect model are very different, averaging predictions is likely to improve the predictions. In the prediction for submission, I averaged the prediction between the XGBoost model and the mixed effect model, with higher weight given to the XGBoost model.

Introduction and potential NEW features to add

The dataset provides 16 important aspects of a typical delivery. But there are much more could influence the time cost of a duration:

Below I raise and discuss 5 potential new features that could improve the model performance, if collected

1. Community/Zipcode

The location of the destinations alters the delivery time. Destinations in suburb areas have easier access and parking than locations in urbdan enviroment, where tall buildings and

stricter security is more common. Further, during weekdays the parking at urban area might be harder than suburb community, and this flunctuation in parking time throughout the week is not captured by the existing estimated point-to-point duration.

2. Condo vs. Home

This feature provides a more detailed account of the environments of the destination than the Community feature above. Condos are in buildings (with stairs and/or security entrance), which means longer time is needed between getting off the delivery car and arriving at the customer's door.

3. Weather

Although the estimated point-to-point delivery by other models captures the weather of that day, the impact of weather on the first and the last mile of the delivery might still impacted by the weather. For instance, potential delay in delivery caused by heavy precipitation can be taken into calculation with this new feature.

4. Vehicle of delivery

Whether the vehicle of delivery has four wheels or two wheels significantly impact how big of an order a dasher can take. This would make the number of available dashers for a certain order more accurate, which enhances the capacity prediction.

5. Average parking time at a specific location

As more and more datapoints accumulate, the average time to park at a specific location would be more accurate. This feature could siginificantly improve the model fit.

Further reading: A relevant paper I found online

Order Fulfillment Cycle Time Estimation for On-Demand Food Delivery by Zhu et al.