

Alexandra Whiteford

12/1/2022

Project Proposal

Predicting Food Delivery Time

Problem Statement:

- Context:
 - As food delivery services have always been integral to city life, improving the accuracy of predicted delivery times is something that is always in need of reshaping and refining. I would like to build a prediction model for the food delivery service company GrubHub, based on the input of food delivery data and outputting the most feasible pick-up and delivery times to be shared with the driver and the customer.
- Criteria for Success:
 - Predicting accurate pick-up and drop-off times as delivery windows within 5 minutes or less of the predicted time. This is done so that deliveries are honest and data-driven, and prevent deliveries perceived as late, leading to an unsatisfied customer.
- Scope of Solution Space:
 - We have three variables to be used as an input to our model:
 - Orders: What are people ordering / From what supplier / How long does this supplier usually take to prepare the order / What time is the delivery set for
 - Customers: Who is ordering / Where are they located
 - Delivery Results: How did this delivery perform compared to past deliveries / How long did each step take / Was one step more challenging/longer than others
 - Data example:
 - Driver name, restaurant name, delivery company, customer name, order id, pick-up address, delivery address, delivery time, order size (\$), distance from restaurant
 - Drivers should be counted on to show up at the supplier at the specified pick-up time. Although in practice this isn't always the case, it can be minimized by having enough drivers.
 - If the food arrives too late, the potential customer will be irritated and dissatisfied.
 - Given the size of deliveries, it can be assumed that cars are needed. However, many delivery drivers in major cities use bicycles or scooters.
- Constraints:
 - Hard-to-read delivery address data as they may be presented in many different forms (apartment, floor, building, zip code).

- Difficult parking or access for pick-up, food not ready
- Unpredictability of traffic / route conditions
- Building security issues, parking difficulties, tricky-to-access customer location or high floor
- Supplier feedback: longer prep times on average, or drivers spending x3 more time at supplier location waiting for order to be ready
- Stakeholders:
 - Carey Tse: Director of Data Analytics
 - GrubHub data science team
 - Head of Driver partners
- Data Sources:
 - Food Delivery Dataset - Kaggle
 - <https://www.kaggle.com/datasets/gauravmalik26/food-delivery-dataset>
 - Google Maps API for traffic patterns (if I can access the API)

Additionally--

- Test/Train: We will use information from previous deliveries, such as the anticipated and actual delivery times, to train the model
- Performance metrics:
 - Early on-time (> 5 mins early)
 - Basically on-time (< 5 mins late)
 - A bit late (< 10 mins late)
 - How often will users be upset?
 - After 15-25 mins, doesn't matter HOW late delivery is, categorically it is late and the customer is not happy
- Main components of a delivery:
 - Pick-up from supplier (pick-up time)
 - Drive from supplier to customer (travel time)
 - Drop-off food to customer (drop-off time)
 - Can use these values to predict the time for each component and add them together to predict final delivery time
 - Hyperparameter (takes into account any random interferences and to act as a buffer; ex: setting pick-up time to 5 mins earlier so drivers are closer and look better when the pick-up seems quick)