How DoorDash leverages AI in its on-demand Logistics Engine

Raghav Ramesh
Machine Learning Engineer, DoorDash
May 2018
Outline

DoorDash Overview

AI in Logistics

Delivery Time Predictions

Batching Algorithms
Last mile, on-demand logistics

Three-sided marketplace

Restaurant Delivery

1600 cities by end of 2018
DoorDash Overview

Logistics Engine

Time Predictions

Batching

100,000+ Restaurants

300,000+ Dashers

10,000,000s of Deliveries
Marketplace

Merchants

Dashers

Consumers
DoorDash
Overview
Logistics
Engine
Time
Predictions
Batching

Marketplace

Merchants
Reach
Revenue

Flexibility
Earnings

Dashers
Selection
Convenience

Consumers
AI @ DoorDash

Core Dispatch
- Batching algorithms
- Hotspots

Recommendations / Personalization
- Search ranking
- Demand distribution

Supply/Demand
- Dynamic Pricing
- Delivery Time

Merchants

Dashers

Consumers

DoorDash Overview
Logistics Engine
Time Predictions
Batching
AI @ DoorDash

**Merchants**
- Food prep time
- Selection intelligence
- Parking prediction

**Dashers**
- Core Dispatch
- Batching algorithms
- Hotspots

**Consumers**
- Pay calculation
- Supply forecasting
- Incentives

**Delivery Time Predictions**
- Supply/Demand Management
- Dynamic Pricing

**Recommendations / Personalization**
- Search ranking
- Demand distribution

**Lifetime value**
- Acquisition
- Promotions

**Logistics Engine**
- Time Predictions
- Batching
Logistics Engine

Core Dispatch
- Batching algorithms
- Hotspots

Merchants

Recommendations / Personalization
- Search ranking
- Demand distribution

Dashers
- Delivery Time
- Dynamic Pricing

Consumers
Logistics Engine

Merchants

Core Dispatch
Batching algorithms
Hotspots

Dashers

Supply/Demand
Delivery Time
Dynamic Pricing

Recommendations / Personalization
Search ranking
Demand distribution

Consumers
Logistics Engine

The AI system that powers DoorDash deliveries
Logistics Engine

**Fast and efficient** deliveries

- On-time delivery to Consumer
- Increase marketplace efficiency
Logistics Engine

Balance Supply/Demand

Plan Routes

Dasher/Delivery matching
Logistics Engine

- **Balance Supply/Demand**
- **Plan Routes**
- **Dasher/Delivery matching**
Optimal Matching

In plain English
● Pick the best Dasher for a Delivery

In canonical Operations Research
● Vehicle Routing Problem

DoorDash specific considerations
● **Real-time** fulfillment
● Optimize supply for future demand
● Decide *when to assign*
Challenges

Complexity
- Combinatorial options
- Delivery constraints

Time constraint
- Asap delivery
- Real-time Consumer demand
- Dasher flexibility

Variance
- Merchant operations
- Traffic
- Weather

Combinatorial options:
- A_1
- A_2
- A_3
- A_4
- A_5

Delivery constraints:
- J_1
- J_2
- J_3
- J_4
- J_5

DoorDash Overview
Logistics Engine
Time Predictions
Batching
ML meets OR

Operations Research (OR)
- Given deliveries, Dashers, and business goals, determine cost function to use and identify the optimal matching

Machine Learning (ML)
- Set up the marketplace
  - Forecast supply and demand
  - Balance
- Calculate inputs to the cost function
  - Travel times, Food preparation times
- Calculate constraints
  - Variances, Batching estimates
- Auxiliary
  - Help with decision on when to assign
Logistics Engine: Summary

**Data**
- Delivery data
  - Location
  - Type
  - Size
- Dasher data
  - Location
  - Capacity
- Constraints
- Equipments

**Predictions**
- Total delivery duration
- Dasher travel duration
- Food prep duration
- Batching quality prediction
- Can two orders be batched?

**Optimizations**
- Cost functions
- Optimal matching
- Mixed integer programming

**Actions**
- Dasher assignment
Let's talk predictions

**Data**
- Delivery data
  - Location
  - Type
  - Size
- Dasher data
  - Location
  - Capacity
- Constraints
- Equipments

**Predictions**
- Total delivery duration
- Dasher travel duration
- Food prep duration
- Batching quality prediction
- Can two orders be batched?

**Optimizations**
- Cost functions
- Optimal matching
- Mixed integer programming

**Actions**
- Dasher assignment
Delivery Time Predictions

When every second is worth a million dollars
Delivery time predictions

- 10+ time point predictions for every delivery
- Capture each part of the delivery process
- Critical component of Logistics system
Lifecycle of a delivery
Lifecycle of a delivery

Merchant receives the order

Merchant starts preparing the food

Ready for pickup
Lifecycle of a delivery

- Merchant receives the order
- Merchant starts preparing the food
- Dasher picks up the food
- Dasher arrives at Merchant
- Dasher reaches location
- Dasher starts driving
- Dasher receives Delivery
- Waiting for food
- Parking

DoorDash Overview
Logistics Engine
Time Predictions
Batching
Lifecycle of a delivery

Merchant receives the order

Merchant starts preparing the food

Dasher picks up the food

Dasher reaches location

Dasher goes to Merchant

Waiting for food

Parking

Dasher starts driving

Dasher receives Delivery

Driving

Parking

Consumer receives food
Lifecycle of a delivery

- Merchant receives the order
- Merchant starts preparing the food
- Dasher receives Delivery
- Dasher starts driving
- Dasher goes to Merchant
  - Dasher reaches location
- Dasher starts driving
  - Parking
- Waiting for food
- Dasher picks up the food
- Total delivery time
- Driving
- Parking
- Consumer receives food
Total delivery time

- Merchant receives the order
- Merchant starts preparing the food
- Dasher picks up the food
- Consumer receives food

Total delivery time:
- Dasher reaches location
- Dasher starts driving
- Dasher goes to Merchant
- Waiting for food
- Parking
- Driving
- Parking

Parking
Driving
Overview

Goal

- Provide ETA to Consumer
- Set constraints for the optimizer

Inputs

- Order details
- Market conditions

Constraints

- Fast, Scalable
- Works for new markets

Directly impacts Consumer conversion and retention
Initial Model

Model

Gradient Boosted Decision Trees
(LightGBM)

Target: Total Delivery Time

Evaluation

Historical data for training
Validate on latest data

Evaluate on RMSE
## Features

<table>
<thead>
<tr>
<th>Order Features</th>
<th>Real time features</th>
<th>Historical aggregates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subtotal</td>
<td># of orders</td>
<td>Past X weeks average delivery times for</td>
</tr>
<tr>
<td>Cuisine</td>
<td># of total Dashers</td>
<td>Store</td>
</tr>
<tr>
<td>Type and price of items</td>
<td>Traffic, Travel estimates</td>
<td>City</td>
</tr>
<tr>
<td>Type of order - group, catering, etc.</td>
<td>Time of day - lunch / dinner</td>
<td>Market</td>
</tr>
<tr>
<td>Merchant details</td>
<td>Day of week - weekend / holiday</td>
<td>Time of day</td>
</tr>
</tbody>
</table>

Similarly:
- average parking times
- variance in historical times
Limitations

- Minimizing mean squared error makes sure we are right on average.
- But, it is more important to deliver by the estimated delivery time, not just be correct on average.
Limitations

- Minimizing mean squared error makes sure we are right on average.
- But, it is more important to deliver by the estimated delivery time, not just be correct on average.

What we have

\[ \text{Actual Delivery Time} = \text{Estimated Delivery Time} + \text{Error} \]

What we want

\[ \text{Prob} (\text{Actual Delivery Time} > \text{Estimated Delivery Time}) < X\% \]
Quantile Regression

- Modeling technique to generate **prediction intervals**
- Instead of Mean Squared Error, the **Loss function** used is:

\[ L(y, F(x)) = h_q(y - F(x)) \]

where
\[
h_q(z) = z \times (q - I_{z<0})
\]

- \( q \) is the percentile value to use
- \( I \) is the indicator function
Quantile Regression Model

- Choosing two $q$ values (say 0.05, 0.95) gives you a **prediction interval** where 90% predictions would lie
- Model evaluated using % **predictions within the interval** and average estimated value
- **Ideal value of $q$** identified through experiments to optimize retention
Extensions

- Geographical ensembling
  - Market specific model for larger markets

- Use of Embeddings
  - Store
  - Market
  - Time of day

- Near real time aggregates
Let's talk predictions

**Data**
- Delivery data
  - Location
  - Type
  - Size
- Dasher data
  - Location
  - Capacity
- Constraints
- Equipments

**Predictions**
- Total delivery duration
- Dasher travel duration
- Food prep duration
- Batching quality prediction
- Can two orders be batched?

**Optimizations**
- Cost functions
- Optimal matching
- Mixed integer programming

**Actions**
- Dasher assignment
Batching algorithms

Doing more with less
Context

What is batching?
- One Dasher working on more than one delivery simultaneously

Why is it important?
- Increases marketplace efficiency
- Help manage supply demand spikes

On-time delivery to Consumer
Increase marketplace efficiency
Challenges

Orders emerge on the fly

Batching deliveries is easier if orders are all pre-scheduled and have a delivery window

Food Quality

5 extra minutes due to batching could mean a burger is no longer hot

Scale

Optimization system can not exhaustively match all pairs of deliveries

Interdependence

Multiple deliveries being dependent on each other increases variance further
Batching Quality Model

Route of Normal Delivery
- \( \text{Pickup}_1, \text{Dropoff}_1 \)

Route of a Batched Delivery
- \( \text{Pickup}_1, \text{Pickup}_2, \text{Dropoff}_1, \text{Dropoff}_2 \)

Model Intuition
- Any delay in \( \text{Pickup}_2 \) affects \( \text{Dropoff}_1 \)
- We should not batch if the probability of delay is high
Model Details

Problem Type
● Classification

Target
● Boolean: Actual Pickup time > Estimated pickup time

Evaluation Metric
● PR-AUC

Use in optimization system
● Two approaches
  ○ Used as constraints
  ○ Use in cost function
Takeaways

- Machine learning helps **efficiently solve** traditional Operations Research problems

- Particularly important in **real-time, high variance** environments like DoorDash

- **Quantile Regression** is an effective technique to generate prediction intervals
The Future

- Automated menu processing
- Price optimization
- Self Driving
- Product placements
- DoorDash Drive
- Robotics
- Demand shaping
- Dispatch for new verticals
- Support experience
- Marketing
- Fraud prevention
We’re hiring!

www.doordash.com/careers

@raghavramesh | raghav@doordash.com

Thank you!

Promo Code: DOORDASHAI