Swiggy Delivery Challenge - Avoiding Rejects

Swiggy has a bold vision to elevate the quality of life of urban consumers by offering unparalleled convenience. At the heart of this vision is solving the hyperlocal delivery challenge of

" Assigning the right delivery partners to the right set of orders at the right time"

This is one of those problems that is positioned at the cusp of marrying the state of the art methods in Artificial Intelligence and Operations Research in large-scale near-real-time applications.

Our delivery partners have the option to reject an order, if he/she wishes to. However, rejection of an order increases the delivery time for the customer, hence we want to avoid rejects.

Here are two sample datasets:
Assignment: Each row corresponds to an order assignment to the delivery partner.

Column name	**Description**
ORDER_ID	Unique Identifier for the ORDER
DE_ID	Unique Identifier for the DE
ASSIGNMENT_START_TIME	Start Time of the Assignment
ASSIGNMENT_END_TIME	End Time of the Assignment
reject_ind	Whether this assignment was rejected
reject_type	Reject Type of this Assignment
PLACED_TIME	Order Placed Time
DELIVERED_TIME	Order Delivered Time
LASTMILE_DISTANCE	Distance to travel in Last Mile (from Restaurant to Customer)
FIRSTMILE_DISTANCE	Distance to travel in First Mile (from DE Assignment Location to Restaurant)
LAST_MILE_TIME_PREDICTED	Time prediction for last mile
PAYOUT_MADE_TO_DE	Actual payout made to DE for this order
NUM_PING_COUNT_LAST10MIN	# of pings received from DE device in last 10 minutes
LAST_PING_TIME_LAST10MIN	time of last ping received from DE device (within last 10 minutes)
CUSTOMER_ZONE	Zone ID for the customer
CUSTOMER_LAT	Coordinates of the customer
CUSTOMER_LAT	Coordinates of the customer

Delivery Partners: Each row corresponds to a delivery partner.

Column name	Description
DE_ID	Unique Identifier for the DE
SHIFT_END_TIME	Shift end time for DE (in HH:MM)
DE_HOME_LAT	Home Location coordinate for the DE
DE_HOME_LNG	Home Location coordinate for the DE
DE_JOINING_DATE	Joining date of the DE
DE_ZONE_ID	Zone ID for the DE

These two datasets are attached in the email.

Additional details:

- Every 2 minutes, all the non-assigned orders are input to the assignment module. For each order, a set of DEs are evaluated, and the order is finally assigned one of these DEs.
- Delivery partners are allowed one reject per day, beyond which they are penalized.
- No payout is made to DE, if he rejects an order
- Every instance of a DE reject, is stored in the production tables, as a unique entry.

Delivering best customer experience is of supreme importance to Swiggy.

As a data scientist, solve the following challenges.

1. The first part of any data science problem is understanding the size, impact and cause of the problem (DE rejection in this case). With that context, can you do an exploratory data analysis?
2. Can you apply a machine learning algorithm on this dataset, to solve the problem? (Hint - The first run of model training might give you weird results)
3. Can you think about where and how the model should be used? Do you see any deployment challenges with the same? How do you overcome these challenges?

Requirements :

- Please submit a notebook of working code with results that you would be walking us through during the Technical depth round.
- You are free to use slides to present the results as well along with code walkthrough.