**Report**

**Part 2 ‐ Experiment and metrics design**

Two methods to assess the effectiveness of the campaign. The first involves using smartphone apps to monitor drivers' locations and determining the percentage of time they spend in each city during working hours. The second method involves calculating the number of drivers who receive toll reimbursements for their working hours. However, this method may not be very informative as it is highly likely that some drivers will get reimbursement, making the results statistically significant but not necessarily meaningful. Considering that drivers' working hours may vary depending on the city they predominantly drive in, we should take into account the different activity patterns in each city. To address this, we propose tracking driver locations during weekends, as this time is typically more active, and assuming we can track their locations accurately. Offering discounts or cashback rewards to encourage higher app usage among riders on weekends compared to weekdays.

**Part 3 ‐ Predictive modeling**

**1. Perform any cleaning, exploratory analysis, and/or visualizations to use the provided**

**data for this analysis (a few sentences/plots describing your approach will suffice). What**

**fraction of the observed users was retained?**

1697 users out of 10000 users were retained. Based on the EDA, weekday\_pct and avg\_rating\_by\_driver had high impact on the user retention rate.

**2. Build a predictive model to help Ultimate determine whether or not a user will be active in**

**their 6th month on the system. Discuss why you chose your approach, what alternatives**

**you considered, and any concerns you have. How valid is your model? Include any key**

**indicators of model performance.**

I have implemented CatBoost classifier because of its ability to handle categorical features with the need of encoding the data using LabelEncoder or OneHotEncoder. This model had 80% accuracy given that no hyper-parameter tuning was performed. Accuracy can be improved by tuning the model. Furthermore, the model predicted that 17% users out of 10,000 users will be retained in their 6th month on the system.

**3. Briefly discuss how Ultimate might leverage the insights gained from the model to**

**improve its long-term rider retention (again, a few sentences will suffice).**

Focusing on the weekdays’ rate and offering incentives for both the riders and customers could increase the probability of booking more rides. The model showed that rating by the drivers also play vital role so drivers should be encouraged to provide genuine ratings.