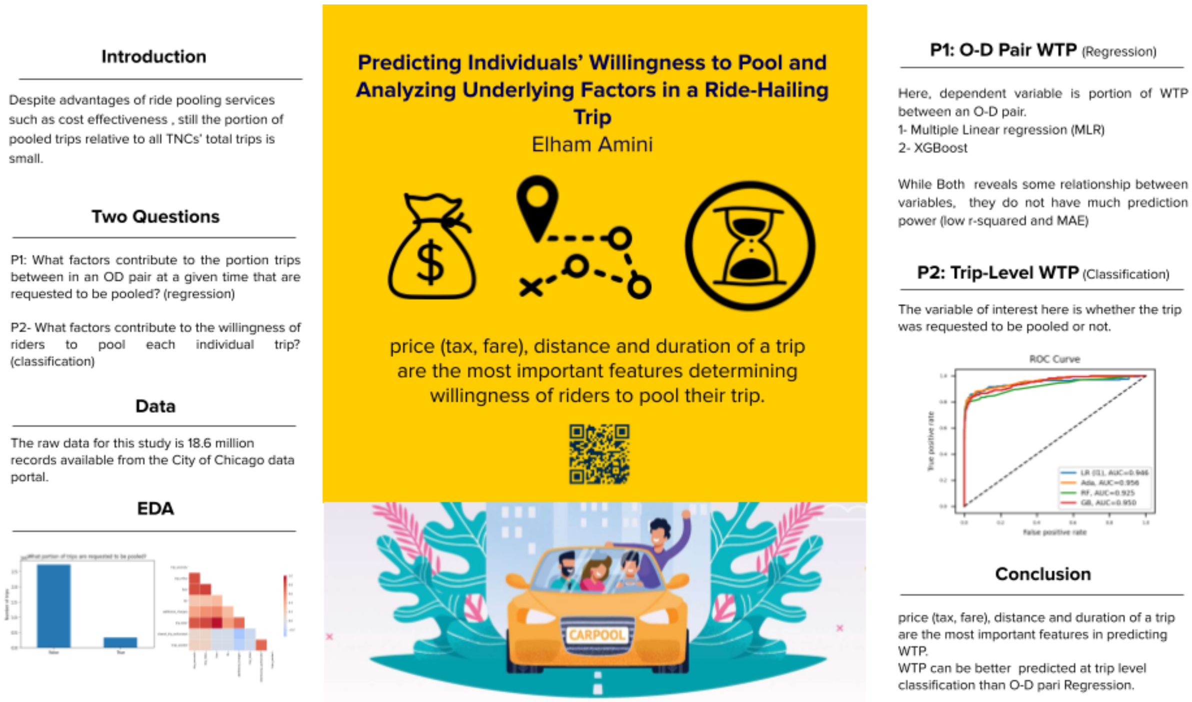
Predicting Individuals' Willingness to Pool and Analyzing Underlying Factors in a Ride-Hailing Trip



Elham Amini

Fall 2020

Introduction:

Transportation Network companies (TNCs) or ride-hailing services offer flexible, efficient, and convenient mobility which are promised to be a remedy for car dependency, traffic congestion, high parking costs, and environmental pollution. However, studies have shown that unintended consequences of ride-hailing services may outweigh some benefits by undermining public transportation, increasing vehicle-miles-traveled (VMT) and large levels of deadheading. [1,2] TNCs offer ride pooling services, in which a group of people share their ride for all or portion of their trips. Pooling the ride has the potential to mitigate negative impacts of ridehailing by consolidating VMT from several spatiotemporally matching trips. Despite advantages of ride pooling including cost effectiveness and congestion mitigation, the portion of pooled trips relative to all ridehailing trips is still small. Survey studies revealed that the underlying factors might be discomfort with sharing a ride with strangers or increase of travel time. [3]
In this project, I use a large novel dataset from ridehailing trips in Chicago, IL, to understand how riders choose to pool their ride. More specifically, I use ML and DL methods to understand and predict willingness to pool (WTP) for ridehailing trips. Note that requesting a pooled trip does not necessarily lead to matching it with another rider and the trip may complete solo. In the dataset, there are direct variables to show which trips are requested to be pooled and whether the pooling was successful or not.

The existing literature on this topic has been predominantly focused on the theoretical aspect of WTP as well as using survey data combined with the neoclassical econometric models to understand the determinants of WTP. This novel dataset provides an unprecedented opportunity to understand the pooling behavior in ridehailing trips using real world empirical data. A literature review of the subject reveals that several studies attempted to use ML and DL methods to estimate the *demand* for ridehailing trip. Yan et al. (2020) showed that a Random Forest model has a superior predictive power compared to traditional multiplicative demand estimation models. [4] Ghaffar et al. (2020) used a traditional econometric model (random-effects negative binomial regression) also for demand forecasting of ridehailing trips and finding the factors impacting travel demand. [5] Modeling of WTP and pooling behavior has received less attention while it has substantial impact on sustainability and the congestion effect of ridehailing services. In particular, in a recent study from NREL researchers, Hou et al. (2020) used a multiple linear regression and XGBoost model to identify the determinants of willingness to pool between OD pairs in Chicago. [6] I replicate this study here and susbatabally improve its performance. I also show a dual problem to this, as a direct classification of WTP for each trip leads to a much better prediction. I show that while ML models can reveal the most important factors (features) that influence WTP, the regression-based predictive model suffers from large error, showing large bias in presence of outliers (OD-pairs with low trip frequency).

Data Sources:

The raw data for this study is available from the City of Chicago data portal. [7] This spatiotemporal data is at the trip-leveland contains trip id, trip start/end timestamps, trip duration, miles, pickup/dropoff census tracts, fare, additional charges, tip, shared trip authorized (T/F), trips pooled, and a few other variables (See **Appendix A** for more about how the privacy riders is practiced). For this study, I focus on data from November 6, 2019 to January 5 2020. The raw dataset for 2-month worth of data consists of 18.6 million trip records.

Since the pick-up and drop-off is reported at the census tracts, we can augment many explanatory variables at the census tract level as well as some other variables embedded in the dataset. These control variables represent fixed effects and panel variation can be used as control variables in models. I append the data with socioeconomic/demographics variables (from American Community Survey), built environment (from ACS, Chicago Metropolitan Agency for Planning (CMAP), Longitudinal Employer-Household Dynamics (LEHD)), transit supply (GTFS), and points of interest (Google Map API) as explanatory variables. More detailed description, descriptive statistics of auxiliary data, which is augmented to the main dataset, as well as some exploratory data analyses are provided in the **Appendix A**.

Data Cleaning:

The data requires significant cleaning effort as well as dealing with a large number of missing values. I first remove all incorrect observations which can be characterized as inconceivable trips:

- Trips with total trip duration less than 1 minutes and longer than 5 hours
- Trips total distance traveled less than 0.25 miles and greater than 300 miles
- Trips with total fare equal to zero (fares are already rounded)
- Removing trips with extreme speeds (below 0.2 mph and above 80 mph - an auxiliary variable from trip distance and trip duration).

Nearly 16.5% of trips suffer from missing pick-up and/or drop-off census tracts. The data provider cites privacy concerns in masking these values. In order to deal with this issue, I first attempt to infer the pickup and dropoff tracts from the pickup and drop off community code. The city of Chicago has 77 community areas and within each community area there are multiple census tracts. I group the dataset by community area, and impute missing census tracts within each community area by trip-density weighted ranking of the non-missing census tracts in that group. This imputation reduces missing values to 5.8%, but may induce a modest bias in the estimation. The data includes some trips outside the boundaries of the City of Chicago (Cook County). Since the objective of this project is to analyze trips within the city of Chicago, I also remove tracts outside of the City of Chicago. For this purpose, I use tracts from the 2010 census boundaries (801 census tracts). At the end of the data cleaning process, about 30% of trips were removed, bringing down the total number of observations to 12.9 million records for two-month worth of data.

Exploratory Data Analysis:

After univariate analysis, I found that nearly 11% of all trips have been requested to be pooled. So, I have an imbalance target label (**Figure 1**). In the multivariate analysis, I found that "Additional Charges[1]" is the variable that has conspicuous correlation with "shared_trip_authorized" (**Figure 2**). So, this variable is expected to be one of the important features in model prediction. This likely because having an economical trip is one of the incentives for a person to request a shared trip. I

[1] The taxes, fees, and any other charges for the trip.

also remove total_trip since it is strongly correlated with fare and is already represented as a summation of fare, additional charges and tip. Moreover, trips_pooled should be removed from features in the prediction part, since it carries information about the target variable and may lead to data leakage. To avoid data leakage, I also normalize features after splitting the training/testing sets in all steps.

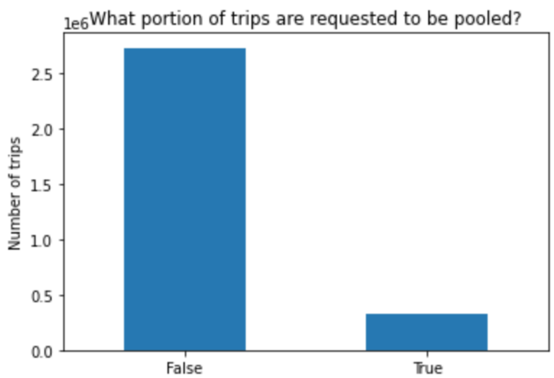


Figure1:Distribution of target label (imbalanced).

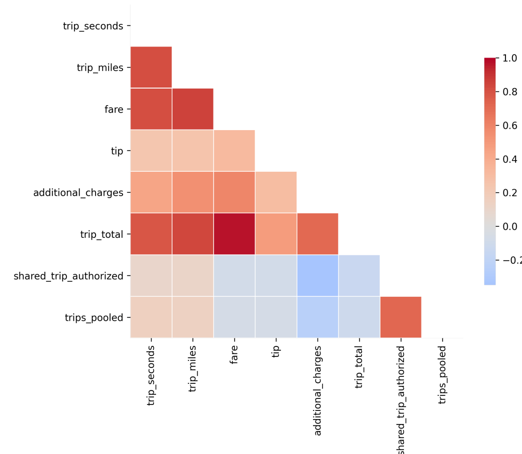


Figure 2: Heatmap of pearson correlations among variable pairs in the main dataset (multivariate analysis).

Problem Formulation and Methodology

The objective of this study is answering two questions:

1. What factors contribute to the portion trips between in an Origin_Destination (OD) pair at a given time that are requested to be pooled ? A regression problem aiming to predict the portion of WTP between an OD pair.
2. What factors contribute to the willingness of riders to pool each individual trip? I formulate this as a binary classification (share_trip_authorized = 0, 1)

For the regression problem between OD-pairs, I use a multiple linear model and compare the performance with the ensembles methods. For the second problem since the binary classifications here entails highly imbalanced classes, I use a dummy classifier as the baseline. Then compare the performance with 1. Logistic Regression with *l1* regularization (LR (l1)) 2. Random Forest (RF) 3. AdaBoosting (Ada) 4. Gradient Boosting (GB). The first classifier is a common classifier in the econometric studies which has the capability of feature selection among many explanatory variables. The rest of classifiers belong to the class of ensemble methods, which benefit from a pool of weak classifiers and entail automatic feature selection. I specifically chose these methods after an initial performance testing with several known ML methods.

Analysis and Results

P1: O-D Pair WTP (Regression)

In this section, I consider the dependent variable as the portion of willingness to pool in an oringin_destination (OD) pair. My aim is to analyze the effects of various independent variables such as trip duration on the variable of interest.

I first group data with pickup and dropoff census tracts, time of day, weekend or not, and airport pickup/drop off or not. Then, I supplement data with the census level socioeconomic/demographics, transit supply, and built environment variables at the pickup and dropoff census level.[2] Consistent with Hou et al. (2020) I used two approaches to analyze the effect of independent variables on the response variable. First, a multiple regression model is chosen for a better explanation of the relationship between various independent variables and the response with following equation:

$$WTP = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_n X_n$$

Second, a machine learning algorithm called XGBoost is chosen to predict the portion of WTP between an OD pair. For the evaluation, I use *mean absolute error* ($MAE = \frac{1}{N}\sum_{i=1}^{N}\left|d_i - \widehat{d}_i\right|$) as the performance metric for both methods.

The results from linear regression are presented in table 1. From the **figure 3**, I can see that WTP is positively correlated with trip_seconds and trip_miles. From **table 1**, I know these correlations are statistically significant. I can conclude that WTP increases as the average duration and distance of a trip increases. Moreover, I can see that airport dropoff/pickups are positively correlated with trip's duration and distance, meaning trips with dropoff/pickup at airport are on average longer than other trips. These results corroborate findings in the previous studies. I can also see that trips pickup on weekends is positively correlated with midnight and POI areas, meaning that on the weekend most ride hailing trips pickup are at midnight and from restaurants, cinemas, etc. Moreover, on morning peaks and evening peaks trips are longer. Also, pickup and dropoff from airports are negatively correlated with WTP.

[2] A complete list of variables used in this section is explained in table x in the supplementary material.

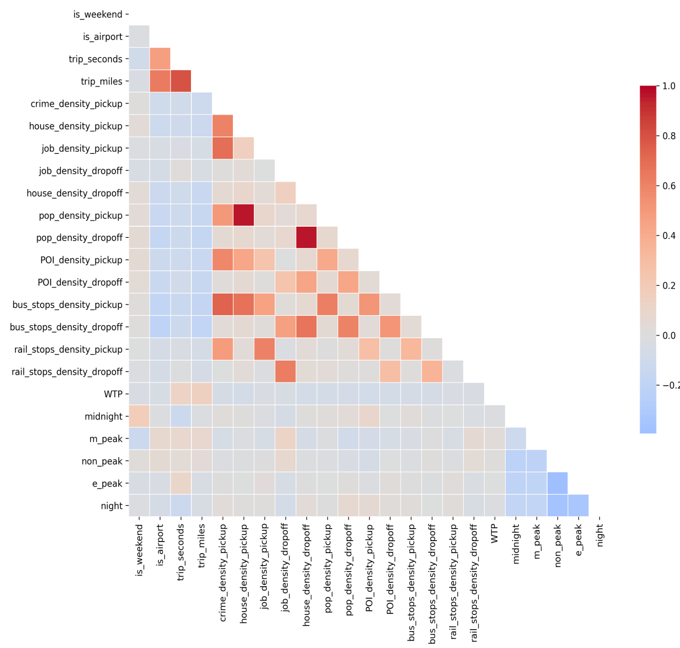


Figure 3. Correlation between variables

While by applying linear regression I could make a better understanding of the relationship between variables, the resulting R-squared is low and MAE is about 0.2, indicating the model is not robust and the variance in response is not explained by the variables well. To fill this gap, and as it is done in the previous studies, I use XGBoost to predict the portion of WTP in an OD pair. Feature importance from XGBoost method confirms that duration, distance of a trip and airport pickup/dropoff areas are among identifying factors in WTP. However, R-squared and MAE did not significantly improve from the linear model.

I can conclude that while formulating WTP as a regression problem between origin destination pairs can reveal the relationship between independent variables and their impact on the target label, resulting models do not have strong predictive power. A good response to this issue is formulating WTP as a binary classification model for each individual trip, which will be discussed in the next section.

Table 1. Linear Regression Coefficient

variable	Estimate (std err)
Is_weekend	-0.0169 (0.002)***
Is_airport	-0.3136 (0.006)***
trip_seconds	-0.1169 (0.050)***
trip_miles	1.5224 (0.035)***
crime_density_pickup	0.1107 (0.012)***
house_density_pickup	-04337 (0.066)***
job_density_pickup	-0.0842 (0.011)***
job_density_dropoff	-0.0511 (0.008)***
house_density_dropoff	-0.4865 (0.057)***
pop_density_pickup	0.1396 (0.01)**
pop_density_dropoff	0.3505 (0.065)***
poi_density_pickup	-0.1447 (0.061)***
poi_density_dropoff	-0.0927(0.0011)***
bus_stops_density_pickup	0.0193 (0.065)**
bus_stops_density_dropoff	0.0353 (0.061)***
rail_stops_density_pickup	-0.0489 (0.011)***
rail_stops_density_dropoff	-0.0166 (0.011)**
midnight	0.0079 (0.003)
m_peak	0.0302 (0.003)***
non_peak	0.02461 (0.02)***
e_peak	0.0201 (0.002)***
night	0.0059 (0.002)

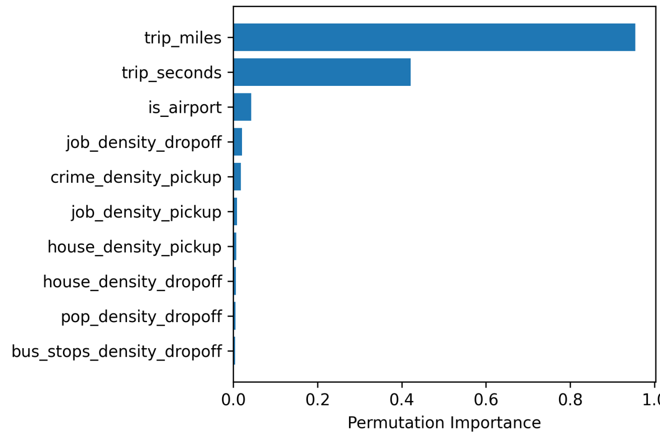


Figure 4. Permutation feature Importance for XGBoost model.

P2: Trip-Level WTP (Classification)

The variable of interest here is whether the trip was requested to be pooled or not. This variable is directly reported for each trip as true/false, thus a binary classification is suitable to identify the factors that determine the willingness of the rider to pool the ride or go solo and also predict whether a given trip is likely to be authorized to be pooled or remain solo. In this exercise, I ignore the time-series nature of the data, and consider each trip as an individual observation.

At the initial step, I compare the performance of the baseline (dummy classifier), which resembles a random guess with the most frequent label, to other candidate ML models with default hyperparameters. Thus, all available features are included in the models. I take a random sample of 100,000 observations and use 20% of data as the test set. The results show that the baseline dummy classifier has an accuracy of 89% and recall of 0%. All of candidate models significantly outperform the baseline, as shown in **Table 2**, but the recall and precision are lower than expectations. Note that while the accuracy for models is high, given the imbalanced nature of labels, precision, recall, and F1-Score are more important as the performance metrics.

Table 2:: Initial performance assessment of baseline model compared to other candidate models. The candidate models are trained with all features with default hyperparameters.

Models	***Accuracy***	***Precision***	***Recall***	***F1-Score***
Dummy Classifier (Baseline)	0.89	0.0	0.0	0.0
RF	0.96	0.91	0.72	0.80
LR (l1)	0.96	0.92	0.71	0.80
Ada	0.96	0.92	0.72	0.89
GB	0.96	0.91	0.72	0.80

Feature Selection:

After initial assessment of the models, I perform a feature selection procedure by focusing on the features that have highest predictive power. A permutation feature importance test reveals that additonal_charges, fare, trip_seconds, trip_miles, and tip are the top five features with highest importance in the model. **Figure 5** shows the top ten features from an untuned RF classification algorithm. I also use `selectkbest` library to find the most important features with Chi-Squared metric, which corroborates the prior finding. As I expected from the EDA, additional charges are one of the most important features in explaining the WTP at the trip level. I further conclude that training the models with only top five most important features slightly improves the performance measures (recall and precision), while significantly reducing the computational load. Thus, I use the aforementioned top five features for the rest of classification analysis.

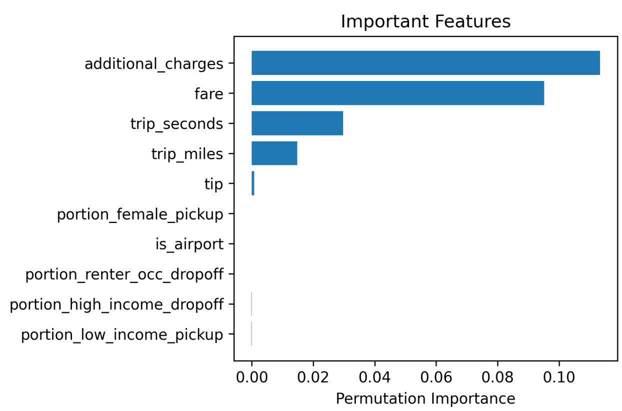


Figure 5: Permutation feature Importance for RF model.

Learning Curve:

In order to assess the robustness of models to the training size, I take bootstrap samples with different sizes from the whole dataset (12.9M records) and plot the learning curves. **Figure 6** shows the learning curves for candidate models. The convergence in the learning curve indicates that the model has been saturated with data with a training sample of nearly 5000 (0.04% of the whole dataset). The initial steep decline in the training score indicates inherent bias in the model without sufficient data, but it declines with appropriate numbers of samples.

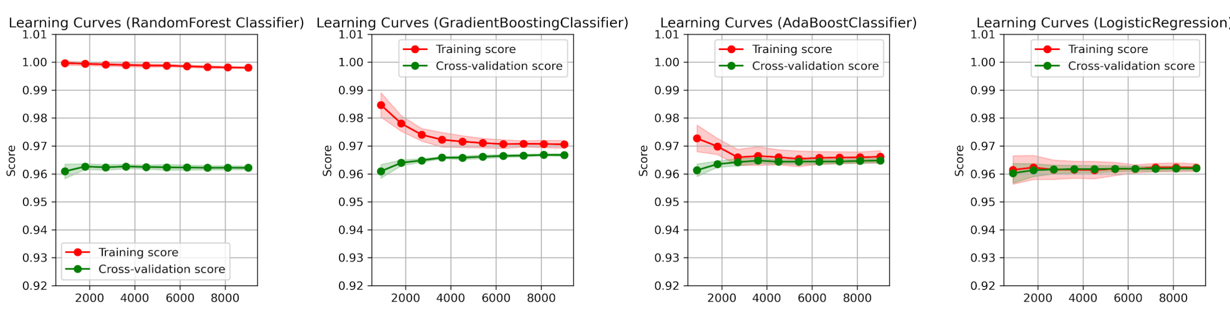


Figure 6: The learning curve for four candidate models using 5-fold cross validation over 100 iterations.

Optimal Model Selection and Hyperparameter Tuning

I use grid search with five fold cross validation to tune hyperparameters of the candidate models. The hyperparameter tuning is not global but attempts to test the parameters within a reasonable space. The (nearly) optimized models are retrieved by multiple scoring criteria based on recall, precision, and F1-score on the training set. **Table 3** shows optimal parameters for each model. In particular, the RF model was strongly overfitting before the hyperparameter tuning. The optimal model for each candidate model does not appear to overfit or underfit.

Table 3: Hyperparameter tuning results for candidate models

Candidate Models	***Number of Trees (n_estimators)***	***Max Depth***	***Number of Features***	***Learning Rate***	***Criterion***
RF	5	5	4	-	gini
LR (l1)	-	-	-	1	-
Ada	100	-	-	1	-
GB	100	7	5	0.1	friedman_mse

Performance Measurement

I evaluate performance of trained models on the validation set. As shown in **Table 4,** and based on the ROC curve (**Figure 7**), AdaBooting Classifier provides the best scores among other candidates. I also observe that the recall in the RF is slightly decreased and it is likely an artifact of hyperparameter tuning, where the overfitting is eliminated.

Table 4: Performance on the validation set. The performance improvement compared to the untuned model is shown in parentheses.

Models	***Accuracy***	***Precision***	***Recall***	***F1-Score***
RF	0.96	0.93 (+.02)	0.68 (-0.04)	0.95 (+0.15)
LR (l1)	0.96	0.92 (0.00)	0.70 (-0.01)	0.96 (+0.16)
Ada	0.96	0.94 (0.00)	0.73 (+0.01)	0.89 (+0.07)
GB	0.96	0.96 (+0.05)	0.72 (+0.04)	0.80 (+0.16)

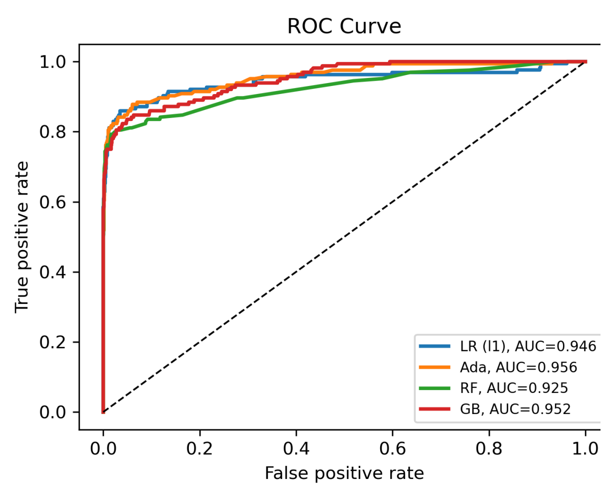


Figure 7: ROC Curve for classification of WTP in each trip

Bootstrap sampling

To ensure robustness of the model, I measure F1-score, precision, and recall for all candidate models for 100 bootstrapped samples size of 100,000 from the global dataset (12.9M). As **Figure 8** demonstrates, while there is significant overlap among the performances, AdaBoosting classifier, on average, outperforms other models in all metrics, which confirms the previous findings.

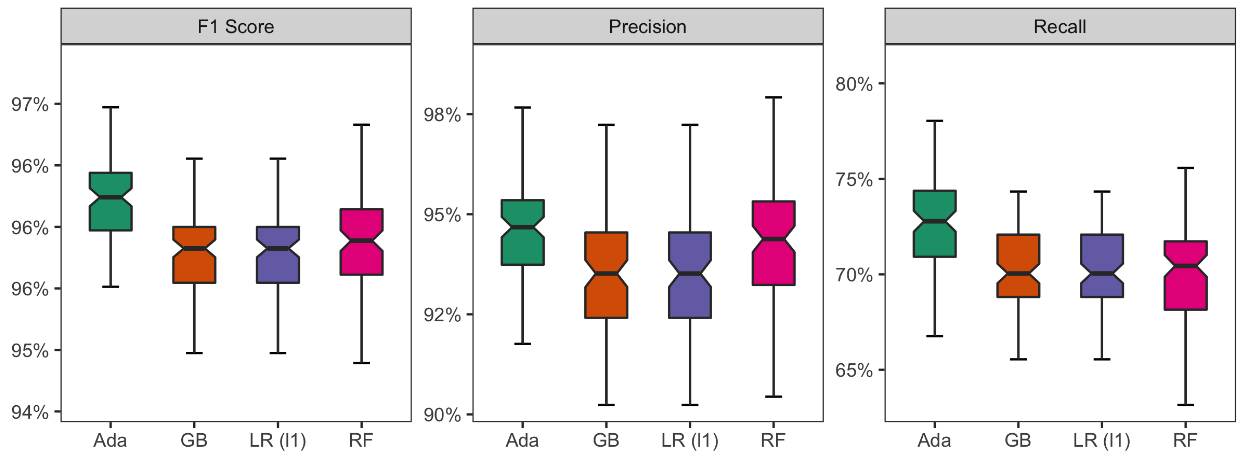


Figure 8: Distribution of model performance metrics over 100 bootstrapped samples from the entire dataset.

After feature selection, and tuning hyperparameters, AdaBoosting classifier was chosen based on its performance on the validation set. **Table 5** shows the performance of the final model on the held-out test set, which has been improved in comparison to the baseline.

Table 5: Performance of selected model (AdaBoosting) on the test set. The performance improvement compared to the untuned model is shown in parentheses.

Models	***Accuracy***	***Precision***	***Recall***	***F1-Score***
Ada	0.97 (+0.01)	0.94 (+0.02)	0.74 (+0.02)	0.96 (0.16)

Based on the findings in the feature importance part, I can conclude that travel impedance variables including trip_miles, trip_seconds, fare, additional_charges, and tip are the major contributing factors in willingness of riders to pool in each individual trip. While other auxiliary variables are correlated with WTP (such as socioeconomics, demographics, built environment, transit supply, POI), the predictive power of these variables is far smaller than main travel impedance variables.

Discussion and Conclusion

In this project I delved into understanding the pooling behavior from ridehailing trips in Chicago. Inspired by Hou et al. (2020), I first explored the WTP between OD-pairs as a regression problem.

While the regression problem reveals the determinants of portions of WTP, it falls short for high accuracy prediction of WTP. I trained a linear model and a XGBoost model, however, both show a poor performance and high error in prediction. I specifically noticed that including OD-pairs with low frequency of trips introduces high bias in the prediction.

The dual to this problem is predicting the probability of a trip being requested as pooled or not. This classification is far more efficient than the regression problem as described in Hou et al. (2020). With properly tuned ensembles models, I reach upwards of 96% accuracy. Cross validating the model the maximize F1-score also significantly improves the learner and increases the precision and recall. However even with the best model, the recall does not reach far beyond 75% while the accuracy consistently reaches above 96%. I observed that while a RF model can reach an accuracy slightly higher than other models, it is prone to overfitting. Note that because of imbalanced classes, precision, recall, and F1-score are more important than accuracy.

The other interesting observation was that I don't need the entire dataset (12.9M rows) for the ML process. A bootstrapped sample of data with proper size can represent the entire dataset and reduce compuatational load singifncalty. Using learning curves, I found that a bootstrapped sample with 5000 observations is sufficiently representative of the entire dataset for the learning purposes. I also observed that while many variables are correlated with the response, only a handful of them have the strongest predictive power. The socioeconomic, demographics, crime data, built environment, transit supply, POI density, start/end in airport or downtown, and time of day are all correlated with WTP for a trip to some degree. However, the ensemble methods can accurately predict the trip-level WTP with a handful of travel impedance variables including fare, additional charges, tip, trip miles and duration.

There are some limitations in this approach. First, because of imputing the missing pickup/dropoff tracts, the predictors are likely (slightly) biased. Second because the trips are anonymized, I cannot attribute the pattern to certain people. I am not observing WTP for each person on a number of rides. My findings cannot be generalized to human behavior. Finally, the pooling behavior after COVID19 pandemic has likely changed significantly. Since March 16 2020, no pooled ride was offered in Chicago and it will likely continue for a considerable amount of time. Thus, it is important to revisit the data once pooled rides are offered again.

References

1. Anair, Don, Jeremy Martin, Maria Cecilia Pinto de Moura, and Joshua Goldman. 2020. Ride-Hailing's Climate Risks: Steering a Growing Industry toward a Clean Transportation Future. Cambridge, MA: Union of Concerned Scientists. https://www.ucsusa.org/resources/ride-hailing-climate-risks
2. Erhardt, G. D. et al. Do transportation network companies decrease or increase congestion? Sci. Adv. 5, (2019)
3. Bansal P, Sinha A, Dua R, Daziano RA. Eliciting preferences of TNC users and drivers: Evidence from the United States. Travel Behaviour and Society. (2020)
4. Yan X, Liu X, Zhao X. Using machine learning for direct demand modeling of ridesourcing services in Chicago. Journal of Transport Geography. 2020
5. Ghaffar A, Mitra S, Hyland M. Modeling determinants of ridesourcing usage: A census tract-level analysis of Chicago. Transportation Research Part C: Emerging Technologies. 2020

6. Hou Y, Garikapati V, Weigl D, Henao A, Moniot M, Sperling J. Factors Influencing Willingness to Pool in Ride-Hailing Trips. Transportation Research Record. (2020)
7. City of Chicago Data Portal, 2020. Transportation Network Providers - Trips. Available at https://data.cityofchicago.org/Transportation/Transportation-Network-Providers-Trips/m6dm-c72p

Note

All data, codes, and Python scripts corresponding to this study can be retrieved from this link.

Appendix A

Privacy protection for riders: The data is entirely analyzed. The trips origin and destination are reported as the census tract to protect the privacy of riders. Time stamps are rounded to the nearest 15 minutes. Fares are rounded to the nearest $2.50 and tips are rounded to the nearest $1.00.

Auxiliary data sources: I supplement the census level socioeconomic/demographics variables from the American Community Survey (ACS) 5-year data (2014-2018). The other variables that can be reported at the census level are built environment variables (population, housing, employment densities), transit supply, crime rate, and points of interest (POIs). Prior research has shown that all these variables have statistically significant impact on travel demand and specifically demand for ridesourcing trips. Thus, they may explain the variance in WTP the ride as well. The other set of variables used here are called travel impedance variables. These variables reflect the resistance to the demand and include distance, duration and fare between the origin and destination. After an initial examination, I may remove some of these variables which exhibit a variance inflation factor (VIF) value over 10 to multicollinearity and its associated identification issues. **Table A1** provides a detailed description of supplemental variables.

Table A1: Auxiliary variables used in this study.

Block	Level	Variable/Feature	Source

Socioeconomic/Demographics Crime Rate	Pick up tract Drop off tract	• Percentage of female population • Percentage of White population • Percentage of Black population • Percentage of two or more population • Percentage of Hispanic population • Percentage of young adults (age 18-44) • Percentage of low income households (<$35000) • Percentage of high income households (>$100000) • Percentage of households without car • Percentage of renter occupied houses • Percentage of adult with at least bachelor's degree	ACS
Built Environment	Pick up tract Drop off tract	• Population density • Housing density • Employment density	ACS LEHD
Crime rate	Pick up tract Drop off tract	• Density of violent crime	CMAP
Transit Supply	Pick up tract Drop off tract	• Density of bus stops per square mile • Density of rail stops per square mile	GTFS/CMAP
POIs	Pick up tract Drop off tract	• Density of Restaurants, Museums, Governmental Buildings	Google Maps API

Table A2: Summary statistics of auxiliary variables augmented to the main dataset. (variables are normalized after split)

Variable	mean	std	min	25%	median	75%	max
trip_seconds	937.92	653.92	61.00	495.00	774.00	1182.00	8667.00
trip_miles	4.69	4.74	0.30	1.60	3.00	5.80	83.90
fare	9.96	6.94	0.00	5.00	7.50	12.50	137.50
tip	0.57	1.50	0.00	0.00	0.00	0.00	50.00
additional_charges	2.83	1.62	0.00	2.55	2.55	2.55	16.34
trip_total	13.37	8.49	0.72	7.55	10.35	15.05	166.60
pooled	0.08	0.26	0.00	0.00	0.00	0.00	1.00
portion_low_income_pickup	0.22	0.15	0.00	0.12	0.18	0.28	0.88
portion_high_income_pickup	0.43	0.19	0.00	0.30	0.48	0.58	0.79
portion_bachelor_up_pickup	0.67	0.26	0.00	0.57	0.76	0.85	0.95
portion_female_pickup	0.49	0.10	0.00	0.48	0.50	0.53	0.77
portion_no_car_pickup	0.32	0.15	0.00	0.19	0.32	0.43	0.78
portion_white_pickup	0.65	0.25	0.00	0.60	0.73	0.82	0.97
portion_black_pickup	0.13	0.22	0.00	0.03	0.04	0.12	1.00
portion_twoplus_pickup	0.03	0.02	0.00	0.02	0.03	0.04	0.16
portion_hispanic_pickup	0.14	0.17	0.00	0.05	0.08	0.12	0.99
portion_renter_occ_pickup	0.57	0.17	0.00	0.47	0.59	0.68	1.00
portion_young18_44_pickup	0.57	0.17	0.00	0.48	0.59	0.69	0.88

portion_low_income_dropoff	0.22	0.15	0.00	0.12	0.18	0.28	0.88
portion_high_income_dropoff	0.42	0.19	0.00	0.29	0.48	0.57	0.79
portion_bachelor_up_dropoff	0.66	0.27	0.00	0.55	0.76	0.85	0.95
portion_female_dropoff	0.48	0.11	0.00	0.47	0.50	0.53	0.69
portion_no_car_dropoff	0.32	0.16	0.00	0.19	0.32	0.43	0.78
portion_white_dropoff	0.64	0.25	0.00	0.58	0.73	0.82	0.97
portion_black_dropoff	0.13	0.22	0.00	0.03	0.04	0.12	1.00
portion_twoplus_dropoff	0.03	0.02	0.00	0.02	0.03	0.04	0.16
portion_hispanic_dropoff	0.13	0.17	0.00	0.05	0.08	0.12	0.99
portion_renter_occ_dropoff	0.57	0.18	0.00	0.47	0.59	0.68	1.00
portion_young18_44_dropoff	0.56	0.18	0.00	0.47	0.58	0.69	0.88
crime_density_pickup	10988.02	10980.26	4.32	3688.78	6267.20	12964.30	44714.11
crime_density_dropoff	11151.60	11145.46	4.32	3594.29	6460.05	13123.38	44714.11
job_density_pickup	81884.84	170450.15	0.00	4524.35	11889.38	116888.99	778628.22
job_density_dropoff	89147.22	180968.63	0.00	4524.35	11889.38	117635.91	778628.22
house_density_pickup	16951.99	15079.82	0.00	7137.74	11865.56	21249.21	193070.98
pop_density_pickup	26858.97	18465.35	0.00	15540.01	23262.79	36640.51	263992.62
house_density_dropoff	16746.45	14946.19	0.00	7137.74	11585.34	21249.21	193070.98
pop_density_dropoff	26489.95	18375.14	0.00	15042.92	22864.75	36640.51	263992.62
POI_density_pickup	312.87	220.83	2.93	160.94	258.19	397.20	1667.90
POI_density_dropoff	312.11	220.69	2.93	159.48	258.19	397.20	1667.90
bus_stops_density_pickup	88.23	47.65	3.04	53.52	79.19	116.72	281.45
rail_stops_density_pickup	16.72	14.66	0.26	7.39	11.16	19.27	74.13
bus_stops_density_dropoff	88.99	47.28	3.04	53.74	79.94	116.80	281.45
rail_stops_density_dropoff	16.93	15.03	0.26	7.25	11.90	19.27	74.13
is_airport	0.07	0.26	0.00	0.00	0.00	0.00	1.00

Table A3: Variables used in O-D Pair WTP (Regression)

Variable	**Explanation**
WTP	The portion of WTP=True trips for a given bin.
crim_density_pickup/dropoff	Crime density of the census tract from which trips started/ended.
house_density_pickup/dropoff	House density of the census tract per mile-squared from which trips started/ended.
job_density_pickup/dropoff	Job density of the census tract from which trips started/ended.
pop_density_pickup/dropoff	population density of the census tract from which trips started/ended.
bus_density_pickup/dropoff	bus density of the census tract from which trips started/ended.
rail_density_pickup/dropoff	railroad density of the census tract from which trips started/ended.
poi_density_pickup/dropoff	Point of interests (restaurants, government buildings, parks, and museums) density of the census tract from which trips started/ended.