SDS 384 Final Project Report
E-scooter Demand Prediction in Austin

Ziyu Fan, Jeil Oh, Zizhe Jiang

1 Introduction and Literature review

Shared micromobility services, such as electric scooters, have gained popularity in recent years as a convenient and affordable transportation option in urban areas. However, managing the demand for these services can be challenging for operating companies, especially as they try to balance the availability of scooters with ridership demand. To address this issue, we conducted a study to identify the key factors that have a significant impact on the demand for shared electric scooters in Austin, Texas, using machine learning techniques to analyze data provided by the city of Austin.

The literature on shared micromobility services has grown rapidly in recent years, with many studies focusing on factors that affect ridership demand. For example, a study by Fishman et al. found that factors such as trip distance, topography, and weather conditions have a significant impact on the use of shared electric scooters[1]. Similarly, a study by Brakewood et al. identified the importance of trip purpose, land use, and public transit accessibility in predicting the demand for shared micromobility services[2].

Machine learning techniques have also been applied to analyze data on shared micromobility services. For example, a study by Nguyen et al. used machine learning models to predict the demand for shared electric scooters in San Francisco based on various factors, such as weather, events, and land use[3]. Another study by Lee et al. used machine learning to analyze data from the city of Chicago's Divvy bike-share program and identified the factors that contribute to the variability in ridership demand across different stations[4].

In this study, we aim to build on this existing literature by applying machine learning techniques to analyze data on shared electric scooter trips in Austin, TX. We will identify the key factors that have a significant impact on the demand for these services, which can help operating companies optimize the distribution and availability of scooters in response to ridership demand. By doing so, we hope to contribute to the development of more efficient and sustainable urban transportation systems.

2 Data

The data used in this study was obtained from the City of Austin's Shared Micromobility Vehicle Trips dataset, which includes information on shared electric scooter trips in Austin, Texas since April 2018. Till now, the dataset includes over 15 million observations of electric scooter trips. The data covers a wide range of locations in Austin, including downtown and surrounding neighborhoods, as well as various times of day and days of the week. This allows us to analyze the demand for shared electric scooters in different areas and at different times.To avoid the influence of the COVID-19 pandemic, we limited our analysis to the 2019 data only.

The 2019 data includes information on approximately 5.4 million scooter trips taken between January 2019 and December 2019. Each trip record includes information such as the start and end time of the trip, the trip duration, the distance traveled, and the starting and ending geolocation of the trip. In addition, we linked the trip data to demographic information from the United States Census Bureau's American Community Survey (ACS). Specifically, we used census tracts to link the scooter trip data to demographic variables such as income, age, and race/ethnicity of the areas where the trips were taken.

This dataset provides a rich source of information on the usage patterns of shared electric scooters in Austin during the pre-pandemic period, which can be used to identify factors that influence ridership demand. The large sample size, detailed trip information, and demographic variables linked through census tracts make it possible to conduct a robust analysis of the factors that affect the demand for shared electric scooters, and to investigate potential disparities in ridership across different demographic groups in Austin.

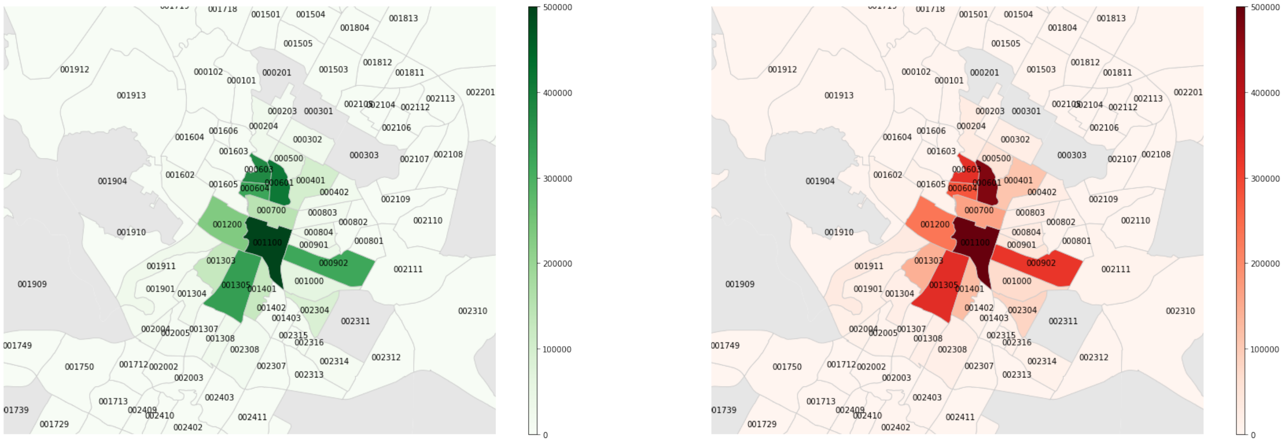
3 Exploratory Data Analysis and Hypothesis

3.1 Exploratory Data Analysis

We conducted several exploratory analyses to gain insights into the patterns and trends in the shared electric scooter data from Austin, TX. These exploratory analyses provided important insights into the patterns and trends in the data, which were used to inform the subsequent machine learning analyses. Some of the key analyses we performed include:

3.1.1 Census Tract Traffic

We created choropleth maps to visualize which areas had the highest scooter traffic. These maps revealed that downtown Austin had the highest concentration of scooter trips, followed by UT Austin main campus, upper west campus, and Bouldin Creek. Figure 1 shows the choropleth maps of the census tract starts and ends.



(a) Start

(b) End

Figure 1: Choropleth maps of census tract starts and ends

3.1.2 Day of the Week Traffic

We investigated the distribution of scooter trips across different days of the week to identify peak usage times. We found that the highest traffic occurred on Fridays and Saturdays, with fewer trips taken during weekdays. Furthermore, we examined how location affected day of the week traffic differently. We found that certain areas, such as downtown and Bouldin Creek, had higher traffic during weekends compared to other areas. On the other hand, we observed that the University of Texas area had higher traffic on weekdays, particularly from Monday to Thursday. In West Campus, we found that Thursday and Friday had the highest traffic. Figure 2 shows the distribution of scooter trips across different days of the week, while Figure 3 shows how location affected the day of the week traffic differently.

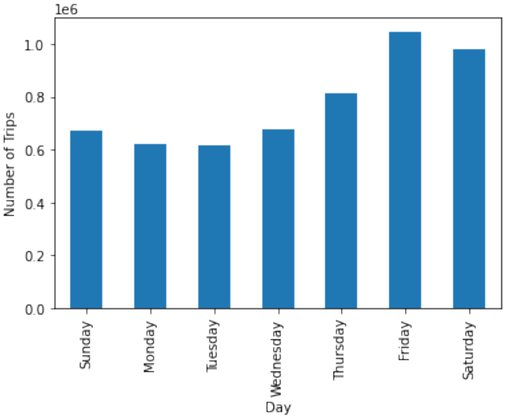


Figure 2: Total Trip Counts by Day of week

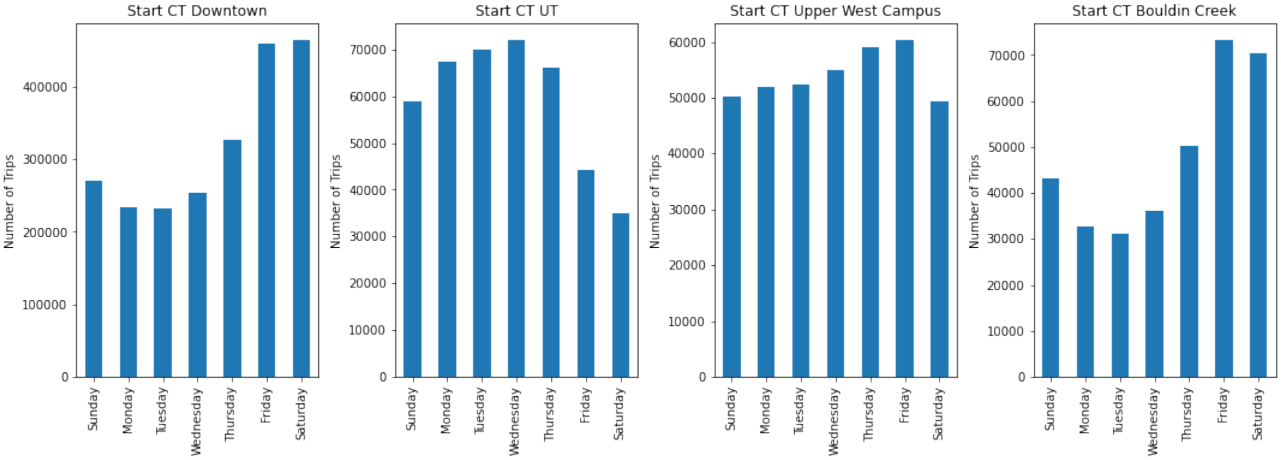


Figure 3: Daily traffic across the top 4 census tracts

3.1.3 Monthly Traffic

We also examined whether certain months were more popular for scooter use than others. Our analysis revealed a bimodal distribution in the number of trips taken, with the two highest peaks occurring in March and September of 2019. To investigate the possible reasons behind the unusual peak in March, we looked at the events happening in Austin during that time. We found that March is typically a busy month for Austin, with several major events taking place, including the South by Southwest (SXSW) festival and the Rodeo Austin festival. It is likely that the influx of visitors to the city for these events led to increased demand for micromobility options such as scooters. Similarly, we looked at the events happening in September to explain the peak in that month. We found that the Austin City Limits music festival and the Texas Tribune Festival were two major events taking place in September. As with March, the increased foot traffic in the city during these events likely contributed to the spike in scooter trips. Figure 4 shows the distribution of scooter trips across different months.

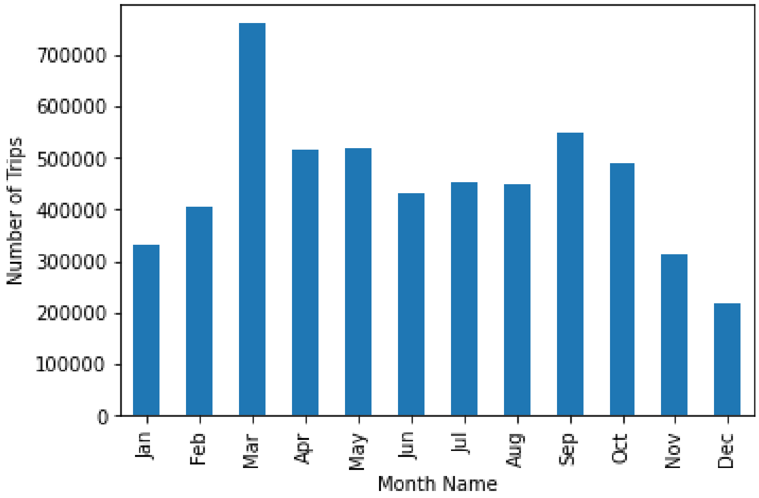


Figure 4: Total Trip Counts by Month

3.1.4 Time of Day Traffic

Finally, we explored how time of day affected ridership. Our analysis showed that peak usage times were during the afternoon and evening hours, with the highest number of trips taken between 3:00 PM and 8:00 PM. Figure 5 shows the distribution of scooter trips across different times of the day.

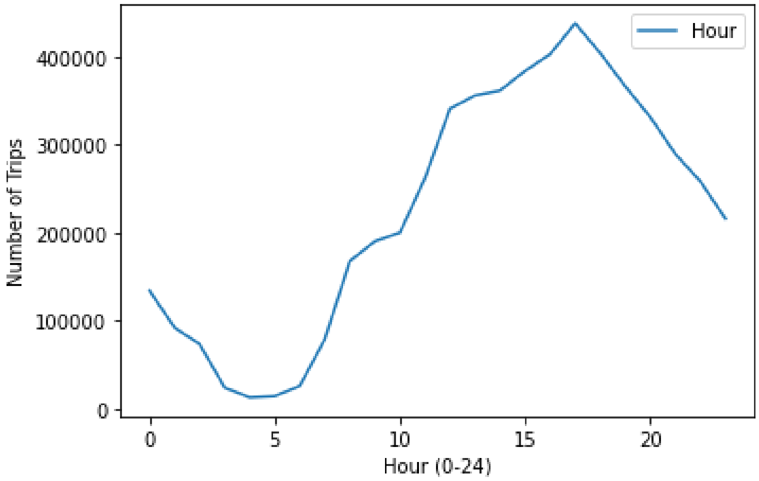


Figure 5: Total number of trips per hour

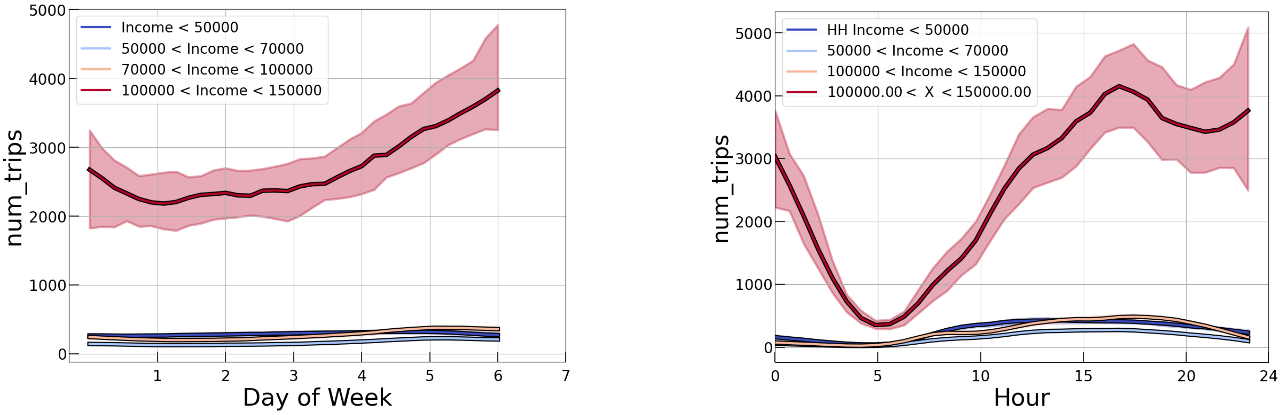
3.2 Exploratory Data Analysis by KLLR

In this section, we present an exploratory data analysis (EDA) of scooter trips using kernel density estimation (KDE) from the kernel density estimation library (kl). The aim of this analysis is to identify trends and patterns in the number of trips with respect to different groups such as income and race.

The EDA generated four key plots, each offering insights into the relationship between scooter trips and different demographic factors:

- Number of trips vs. day of the week grouped by income ranges: The plot reveals that people with the highest income are more likely to ride scooters across all days of the week. This suggests a possible correlation between income level and scooter usage.

- Number of trips vs. hour of the day grouped by income ranges: Similar to the previous plot, the highest income group shows a higher propensity to ride scooters throughout the day, indicating that income may play a role in determining scooter usage at different hours.



(a) KLLR for Hours

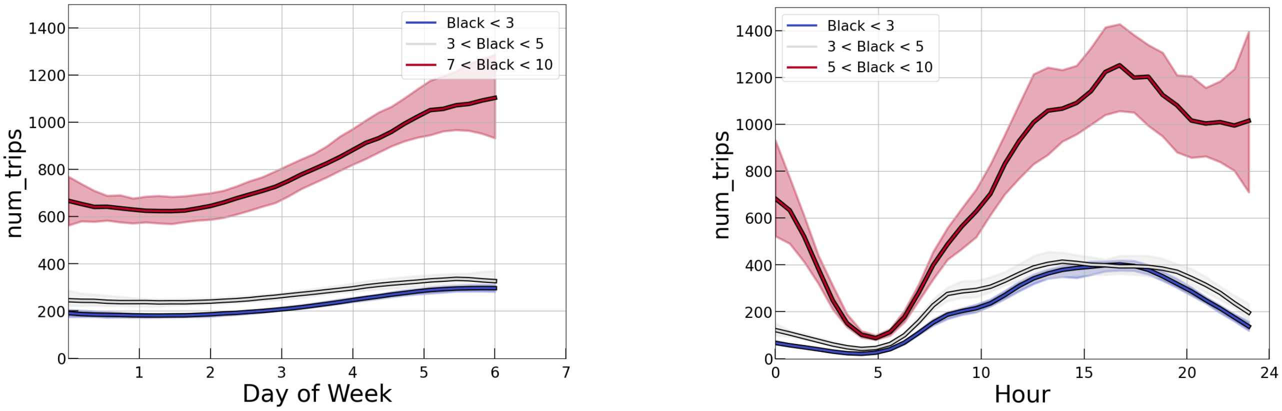
(b) KLLR for day of weeks

Figure 6: KLLR grouped by income range

- Number of trips vs. day of the week grouped by the percentage of the Black population: The plot shows that areas with a higher percentage of Black residents tend to have more scooter trips throughout the week. This finding indicates a potential relationship between the racial composition of an area and scooter usage patterns.

- Number of trips vs. hour of the day grouped by the percentage of the Black population: As with the previous plot, areas with a higher percentage of Black residents exhibit increased scooter usage across all hours of the day. This further strengthens the notion that the racial composition of an area might be correlated with scooter usage patterns.

The exploratory data analysis using kernel density estimation provides valuable insights into the trends and patterns of scooter trips based on income and race. The analysis reveals potential correlations between income, race, and the number of trips taken on specific days and times. By understanding these relationships, policymakers and city planners can make informed decisions to better serve the transportation needs of various demographic groups and optimize



(a) KLLR for Hours

(b) KLLR for day of weeks

Figure 7: KLLR grouped by Race

scooter distribution across different neighborhoods. Further research, including the examination of additional demographic factors, may yield more comprehensive insights into the factors influencing scooter usage patterns.

3.3 Objectives

The primary objectives of this work are the following:

- We construct a regression problem for hourly E-scooter demand prediction.
- We evaluate the predictive performance of five frequently used models (Linear Regression, Ridge regression, Gaussian Process Regression, Random Forest, XGBoost (eXtreme Gradient Boosting), Neural Network) based on the real-world dataset.
- We examine the connections between demographic data features and scooter demand prediction through model interpretation techniques to improve our understating of electric scooter demand.

4 Modeling and validation

Machine learning models boast exceptional data processing capabilities and can identify intricate relationships (including linear and non-linear) between features and prediction targets. As a result, using ML to predict electric scooter demand can enhance our understanding of urban transportation networks and create a model with robust generalization ability.

4.1 Linear Regression

In this experiment, we aimed to create a linear regression model to predict the number of trips taken in a certain area based on various input features. The dataset included information about county, day of the week, hour, demographics, and income. We used Ridge regression, a linear regression model that adds regularization to prevent overfitting and reduce the multicollinearity. The model was trained and tested, and performance was evaluated using Mean Squared Error (MSE) , Mean Absolute Error (MAE) , and R^2 score. The input features were selected: 'Day of Week', 'Hour', 'Men', 'Women', 'Hispanic', 'White', 'Black', 'Native', 'Asian', 'Pacific', and 'Income'. The output feature was 'number of trips'. During the training the dataset was split into training and testing sets, with an 80-20% ratio, using a random state of 42. A Ridge regression model with an alpha of 1.0 was created. Then we using the training data to fit the model, then it was used to predict the number of trips on the test set. Performance metrics were calculated, including Mean Squared Error (MSE) and the plot for Actual vs predicted results.

From the result we found that the Ridge regression model was able to predict the number of trips with a Mean Squared Error (MSE) of 3130676.5, a Mean Absolute Error (MAE) of 728.80, and an R^2 score of 0.1328. The relatively low R^2 score indicates that the model is not effectively capturing the relationship between the input features and the target variable. The plot of actual vs predicted results showed a plot where most of the data cannot fit the model and most of predicted value are higher than the actual value. One possible explanation for the suboptimal performance is the presence of multicollinearity among the input features, particularly within the demographic variables ('Men', 'Women', 'Hispanic', 'White', 'Black', 'Native', 'Asian', 'Pacific') and 'Income'. Multicollinearity occurs when two or more predictors in a regression model are highly correlated, leading to unstable estimates of the regression coefficients and a reduced ability to interpret the individual contributions of each predictor to the model.

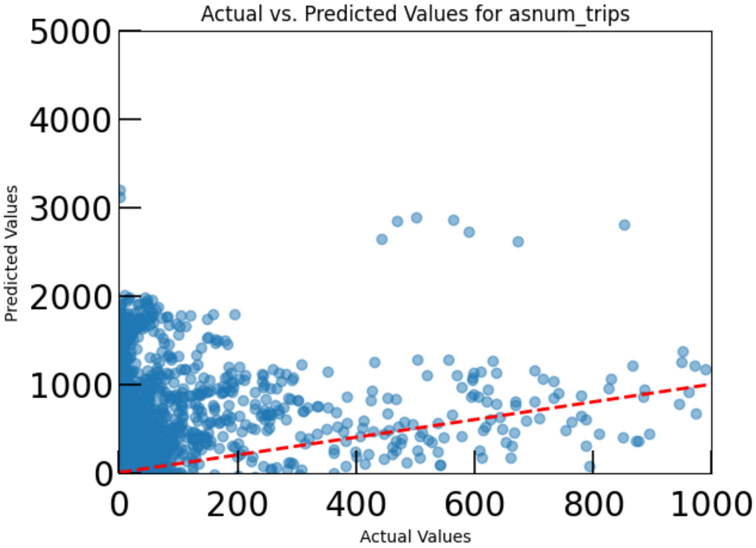


Figure 8: Actual vs Predicted Results for Linear Regression.

4.2 Ridge regression

In order to address the issue of multicollinearity present in the dataset, Ridge Regression was employed. Ridge Regression is a regularization technique that introduces a penalty term (the L2-norm) to the linear regression model's coefficients. This method helps mitigate multicollinearity by preventing the coefficients from becoming too large, which can lead to overfitting and instability in the model.

Using Ridge Regression on the dataset, the following results were obtained: a Mean Squared Error (MSE) of 3131517.81, a Mean Absolute Error (MAE) of 729.14, and an R^2 score of 0.1326. Compared to the initial linear regression model, the Ridge Regression model yielded slightly different performance metrics, with a marginally higher MSE and a slightly lower MAE and

R^2 score. These results indicate that Ridge Regression had a limited impact on improving the model's performance for this particular dataset. It should be note that the Ridge Regression model's performance could be further optimized by tuning the regularization parameter (alpha). A systematic approach such as cross-validation or grid search can be employed to identify the optimal value of alpha that yields the best model performance. But the R^2 score is so low, which means this dataset is not linearly Separable.

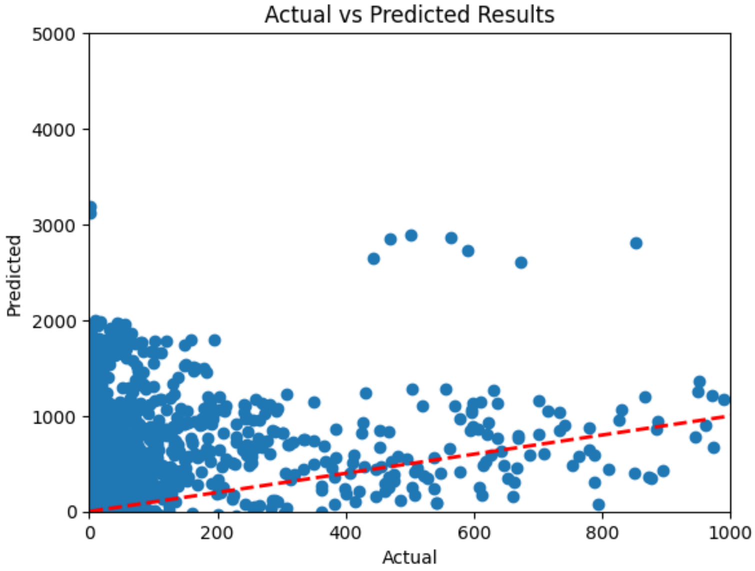


Figure 9: Actual vs Predicted Results for Ridge Regression.

In conclusion, while Ridge Regression was employed in an attempt to mitigate the effects of multicollinearity, the model's performance did not significantly improve. Linear model cannot fit this model well, we need to explore other methods, such as non-linear model.

4.3 Gaussian Process Regression

In this research, we aimed to create a Gaussian Process Regression (GPR) model to predict the number of trips taken in a certain area based on various input features. The dataset included information about county, day of the week, hour, demographics, and income. We used a combination of Radial Basis Function (RBF) kernel and White kernel to create the GPR model, which inherently captures non-linear relationships between input features and the target variable. The model was trained and tested, and its performance was evaluated using Mean Squared Error (MSE), Mean Absolute Error (MAE), and R^2 score. The dataset was split into training and testing sets, with an $80-20\%$ ratio, using a random state of 42. A GPR model with a kernel composed of an RBF kernel with a fixed length scale and a White kernel with noise level bounds was created. The GPR model was able to predict the number of trips with a Mean Squared Error (MSE) of 332974.18, a Mean Absolute Error (MAE) of 107.998, and an R^2 score of 0.9078. The GPR model's performance in predicting the number of trips was highly accurate, indicating that it captures the non-linear relationships between input features and the target variable well. Additional analysis could be performed to understand the contributions of each input feature and their relationships with the target variable. It might also be beneficial

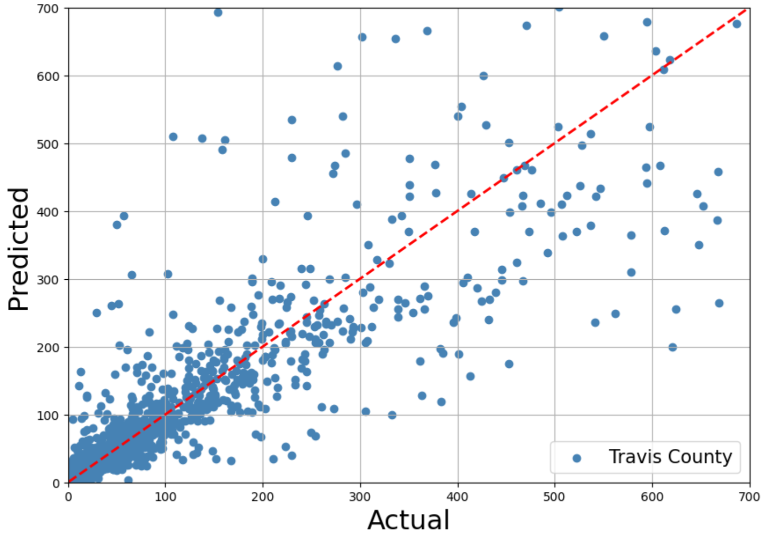


Figure 10: Actual vs Predicted Results for GPR.

to experiment with other kernels, tune the hyperparameters, or apply feature engineering to enhance the model's performance.

4.3.1 Compare Linear model and kernal model

Based on the results from both the linear regression and Gaussian Process Regression (GPR) models, it is evident that this dataset is not linearly separable. The linear regression model, with a Mean Squared Error (MSE) of 3131517.81, Mean Absolute Error (MAE) of 729.14, and an R^2 score of 0.1326, demonstrates relatively poor performance. In contrast, the GPR model achieved a considerably better performance, with an MSE of 332974.18, MAE of 107.99, and an R^2 score of 0.9077.

One possible reason for GPR's superior performance is its ability to capture non-linear relationships between input features and the target variable, thanks to its flexible kernel function. Linear regression models are based on the assumption of linearity between the predictors and the outcome, which is evidently not the case for this dataset. By incorporating non-linear relationships, GPR is able to better model the underlying data structure and thus provide more accurate predictions.

In terms of complexity, GPR can be more computationally expensive than linear regression, as it requires calculations and optimizations of kernel functions, which can increase with the number of data points. However, this additional complexity is worthwhile, as it provides better predictive capabilities for this particular dataset. It is essential to strike a balance between model complexity and performance, and in this case, the GPR model offers a superior trade-off.

4.4 Random Forest and XGBoost

In this study, we employ two popular decision tree based ensemble learning methods, Random Forest and XGBoost, to develop our electric scooter demand prediction models. Ensemble learning combines the strengths of multiple base models to achieve higher accuracy and better generalization performance. Both Random Forest and XGBoost are powerful and versatile techniques that have demonstrated success in a variety of tasks.

Random Forest is a collection of decision trees, where each tree is trained on a random subset of the dataset with replacement, known as bootstrapping. This process helps to reduce the model's variance and improve its overall performance. The final prediction is obtained by averaging the outputs of all individual trees.

XGBoost, short for eXtreme Gradient Boosting, is a more advanced ensemble method that builds decision trees sequentially, with each tree focusing on correcting the errors made by the previous one. The final prediction is a weighted sum of the outputs of all individual trees. XGBoost employs gradient boosting, an optimization technique that minimizes the loss function to improve the model's performance.

For the modeling and validation process, the data subsets were first randomized and divided into two parts: an 85% training set and a 15% test set. The training set was used for the development of machine learning models, while the test set was reserved for evaluating the best-performing model. Throughout the development and assessment phases, the mean squared error (MSE) was employed to gauge the model's predictive performance on both the training and test sets. A lower mean squared error (MSE) signifies better predictive performance.

To enhance the model's prediction and generalization abilities, a 5-fold cross-validation technique was applied to the training set during the development process. This method divides the training set into five subsets, using each subset for validation while training the model on the remaining data. Moreover, machine learning algorithm selection and hyperparameter tuning were carried out based on the average predictive performance across the five validation sets. For hyperparameter tuning, we focused specifically on the *max_depth* parameter. Max_depth represents the maximum depth of the decision trees, controlling the complexity of the model. By adjusting this parameter, we could find the optimal balance between bias and variance, preventing overfitting and underfitting. This comprehensive approach ensures the development of a robust and accurate model for electric scooter demand prediction.

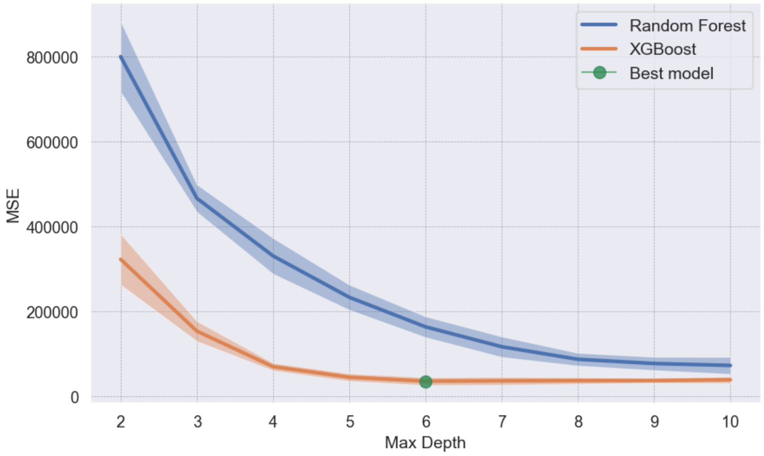


Figure 11: Hyper-parameter tuning based on the maximum depth of Random Forest and an eXtreme Gradient Boosting (XGBoost).

During the tuning process where the maximum depth parameter was varied from 2 to 10, XGBoost outperformed random forest across all depths. Notably, the combination of XGBoost and a depth of 6 yielded the lowest MSE, indicating the highest level of performance (Fig. 11).

4.5 Neural Network

We also employ a neural network model to predict electric scooter demand. Neural networks are powerful machine learning algorithms capable of approximating complex non-linear functions, making them suitable for this task. Our model comprises two fully connected hidden layers, each containing 100 neurons and utilizing the ReLU activation function, followed by an output layer with a single neuron for the regression task. The model is optimized using the Adam optimizer and is designed to minimize the mean squared error loss function, with mean absolute error as an additional evaluation metric.

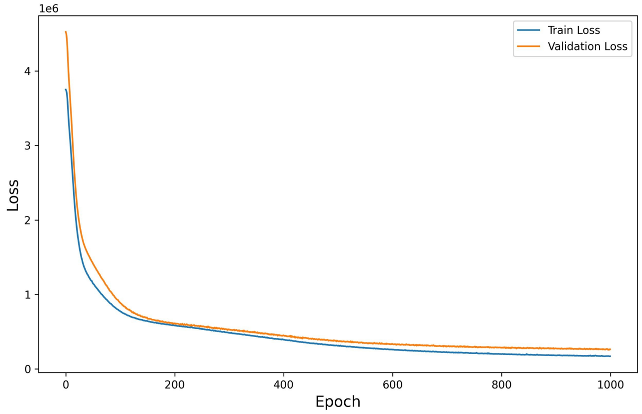


Figure 12: Training and Validation loss for the neural network model.

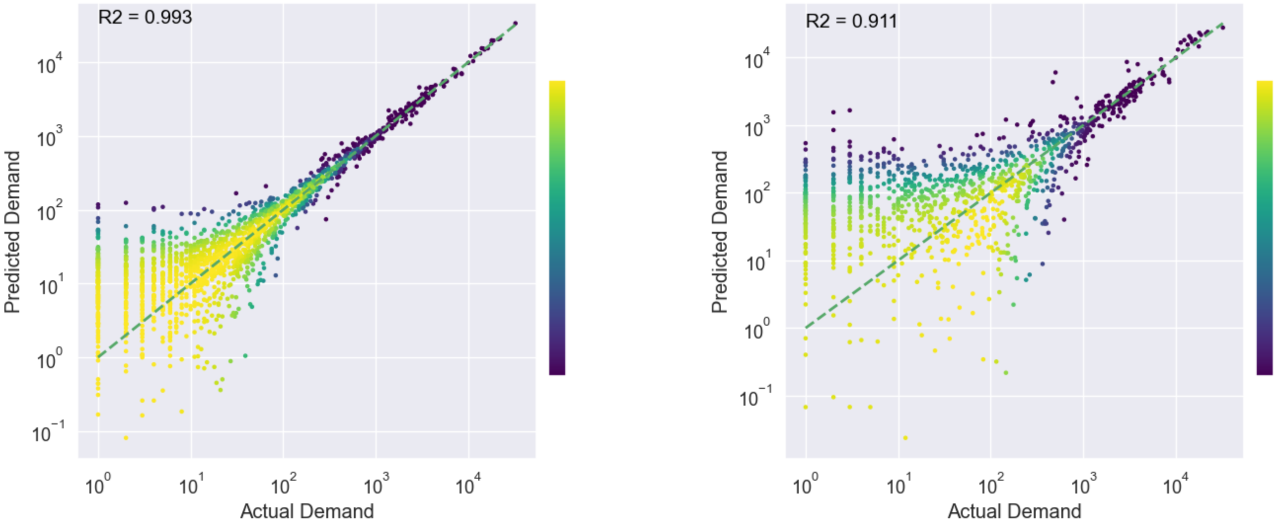
For the modeling and validation process, we partition the dataset into an 85% training/validation set and a 15% test set, ensuring that the data is randomly split as in the previous section. We then employ a 5-fold cross-validation technique on the training/validation set to improve the model's predictive performance and generalization capabilities. In this process, the training/validation set is divided into five equally-sized subsets, where the model is trained on four subsets and validated on the remaining one. This process is repeated five times, with each subset serving as the validation set once. Throughout the cross-validation process, we train and evaluate the neural network model on each fold, recording the validation loss and mean absolute error (Fig. 12).

5 Results

The XGBoost model achieved an average mean squared error (MSE) of 35139.13 ± 9499.18 across the five folds, while the neural network model achieved an average mean squared error (MSE) of 170643.69 and a mean absolute error (MAE) of 163.57 across the five folds. These metrics indicate that our model demonstrates a reasonable degree of accuracy in predicting electric scooter demand. It is important to note that the relatively low variance observed in the performance metrics across the folds suggests that the model is robust and less prone to overfitting. In addition to the cross-validation results, we also assess the performance of our trained model on the test set, which comprises 15% of the data. This evaluation offers insights into how well the model generalizes to new, unseen data. On the test set, the XGBoost model achieved an R^2 of 0.993, and the NN model achieved an R^2 of 0.911, further confirming its predictive accuracy and reliability in electric scooter demand prediction (Fig. 13).

Table 1: ML models performance comparison.

	MSE	MAE	R^2
Linear Regression	3130676.50	728.80	0.1328
Ridge Regression	3131517.81	729.14	0.1326
GPR	332974.18	107.99	0.908
XGBoost	18722.52	42.82	0.993
Neural Network	234112.27	178.40	0.911



(a) XGBoost model predictive performance for E-scooter demand

(b) Neural Network model predictive performance for E-scooter demand

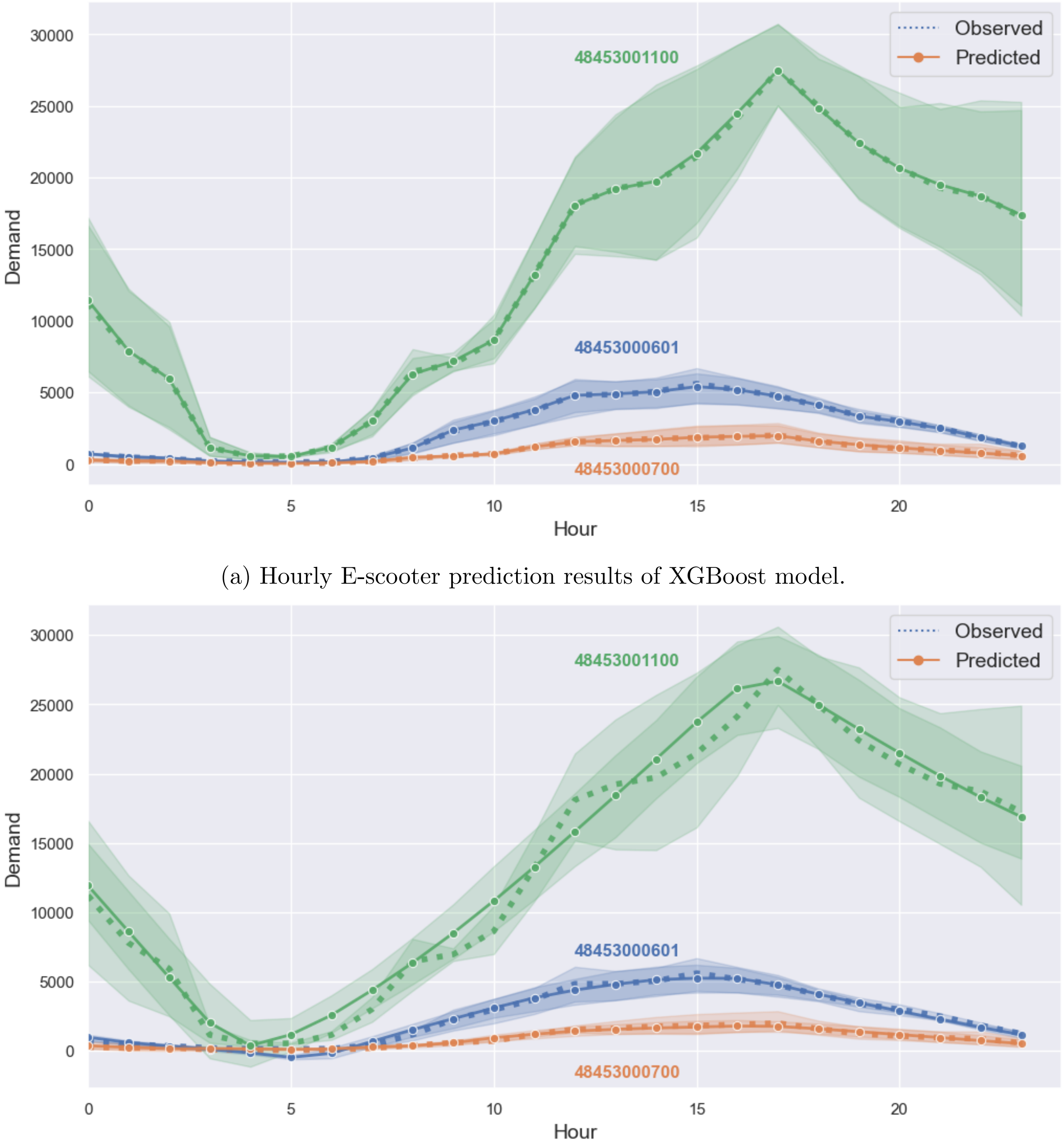
Figure 13: Hourly E-scooter demand prediction performance of XGBoost and Neural network.

Figure 14 illustrates the hourly electric scooter demand predictions generated by the (a) XGBoost and (b) Neural Network models for selected Census Tracts (48453001100, 48453000601, and 48453000700). Both models demonstrate strong performance in predicting demand. The solid line represents the predicted values from the models, while the dotted line corresponds to the actual observed values. It is evident that the demand varies across different regions, yet the machine learning models effectively capture and forecast these variations.

In conclusion, our machine learning models demonstrate promising results in predicting electric scooter demand, with consistent performance across both cross-validation folds and the test set. These findings suggest that the model can be a valuable tool for understanding and managing urban transportation networks, potentially guiding the allocation of resources and infrastructure planning for electric scooters.

6 Discussion

To interpret machine learning models, one can analyze the relationships between feature-target pairs in order to ascertain the significance or contributions of each input feature to the model's predictions. In this study, we examine the best-performing XGBoost model from various angles, employing tree-based feature importance analysis, partial dependence plots (PDPs),

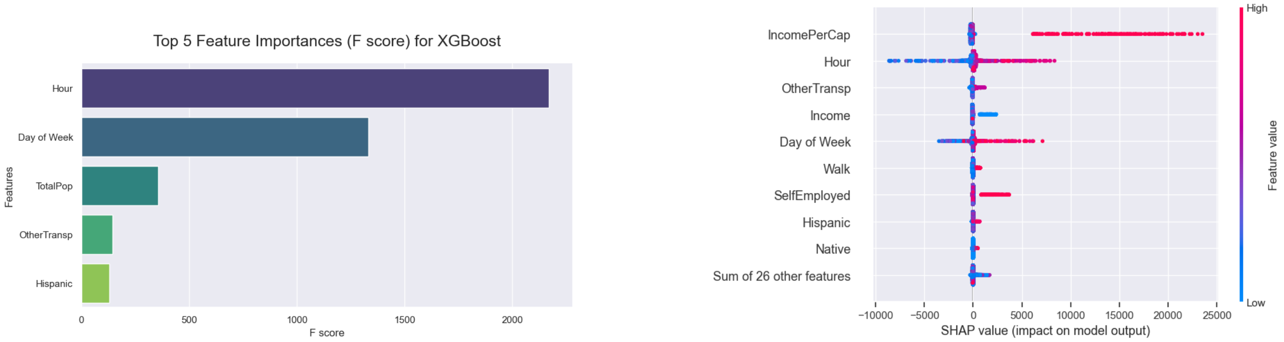


(a) Hourly E-scooter prediction results of XGBoost model.

(b) Hourly E-scooter prediction results of Neural Network model.

Figure 14: Hourly E-scooter prediction results of (a) XGBoost and (b) Neural Network model for the selected Census Tract (48453001100, 48453000601, and 48453000700).

and Shapley additive explanations (SHAPs) for a comprehensive interpretation.



(a) Feature Importance for the best XGBoost model

(b) SHAP summary interpretation plots for the best XGBoost model.

Figure 15: Feature importance of the best XGBoost model and SHAP summary interpretation plots for E-scooter demand.

The concept of feature importance involves assigning a score to each feature in a model, indicating its usefulness and contribution to the construction of boosted decision trees. A feature with a higher score has been utilized more frequently in key decision-making processes, and therefore, is considered more important. As shown in Figure 15a, the five most important features, in descending order, are Hour, Day of Week, Total Population, Other Transportation, and Hispanic ratio. The optimal XGBoost model can be visualized through SHAP summary interpretation plots (Fig. 15b). It depict the features ordered by the mean absolute SHAP value along the y-axis, while the SHAP values are displayed on the x-axis. The color of each feature varies from blue to red based on the SHAP value, with blue representing low values and red representing high values. A positive (negative) SHAP value for a feature indicates that it can increase (decrease) the E-scooter demand value. Each scatter pattern in the plot corresponds to the SHAP values for the training data in that feature. The top five features from the SHAP summary plot were Per capita income > Hour > Other Transportation > Income > Day of Week.

Our feature analysis indicates that time-related factors primarily drive demand. Interestingly, the Hispanic population ratio also emerged as a significant determinant. To further investigate the relationship between race and scooter demand, we employed a Partial Dependence Plot (PDP). Figure 18 illustrates the PDP for hourly demand as it correlates with the proportions of white, black, Hispanic, Asian, and native populations. For most racial groups, demand fluctuation remains relatively minor with changes in racial composition. However, the Hispanic population exhibits a distinct demand pattern based on racial proportions, which is not observed in other groups. Figure 17b highlights an increased demand when the Hispanic population exceeds 60%. We have also conducted a comparison between the partial dependence plot (PDP) and KLLR (Fig. 17). In KLLR, we grouped the data based on composition, but we could not observe a similar pattern. It is worth noting that the partial dependence function at a specific feature value represents the average prediction when all data points assume that feature value [5]. Therefore, the results suggest that areas with more than 60% Hispanic population may not be adequately represented in the distribution of E-scooters, despite having a high potential demand. However, it is important to acknowledge that PDP assumes that the features of interest are not correlated with other features. To make a conclusive statement, further research and validation are necessary.

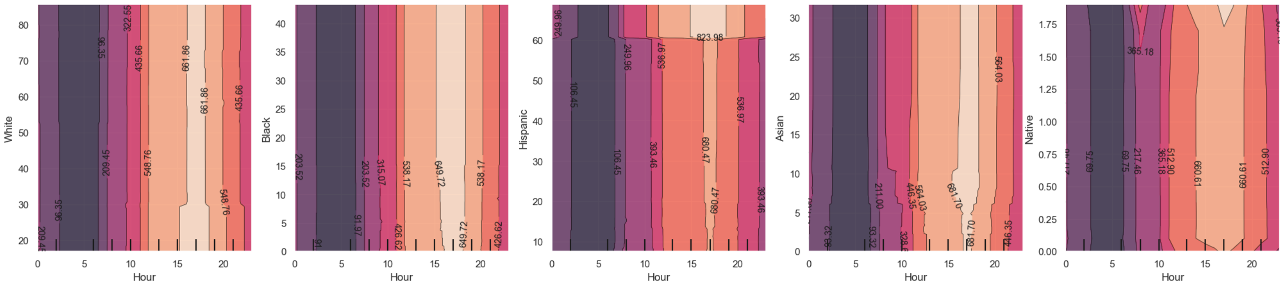
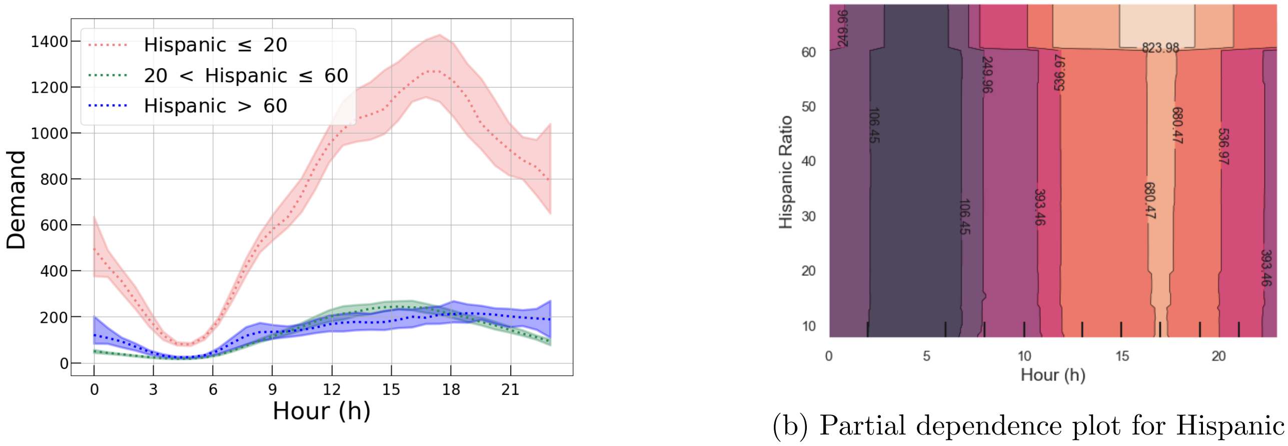


Figure 16: Partial dependence plot of E-scooter demand and the interaction of hour and different race.



(a) KLLR for Hispanic population.

(b) Partial dependence plot for Hispanic population.

Figure 17: Comparison of KLLR and PDP for Hispanic population.

6.1 Limitation of the data

This dataset contains several limitations which include: discrete data, Missing Variables, Potential Multicollinearity, Spatial Autocorrelation and Temporal Variability

6.1.1 Discrete data

This dataset consists of discrete data, which might not adequately capture the nuances in continuous variables like income. This can lead to a loss of information and impact the model's ability to make accurate predictions. As a result, the analysis and conclusions drawn from the regression model should be taken with a grain of caution.

6.1.2 Missing Variables

The dataset does not contain all potentially relevant variables that may influence the number of trips, such as weather conditions, public transportation availability, or nearby attractions. The absence of these variables might result in an incomplete or biased regression model, which could affect the accuracy of the predictions.

6.1.3 Potential Multicollinearity

The demographic variables included in the dataset, such as income, race, and gender distribution, might be correlated with each other since these data are are generate from the TrackId. Multicollinearity can make it difficult to determine the individual effects of these variables on the number of trips and may inflate the variance of the regression coefficients, leading to unstable or unreliable results.

6.1.4 Spatial Autocorrelation

Since some of the data is derived from location information, there might be spatial autocorrelation present, which violates the independence assumption of regression models. This can lead to biased estimates of the regression coefficients and affect the model's predictive accuracy.

6.1.5 Impact of exogenous input (South by Southwest)

During March, the city of Austin experiences a surge in visitors due to the popular South by Southwest conference. Although this event significantly increases the demand for e-scooters, this does not match with the population data for downtown Austin. As a result, we excluded scooter trip data from March to eliminate external influences caused by the large event, South by Southwest. When predicting demand based on race, we found that there was still a high demand in areas with a Hispanic population composition of more than 60%, as was previously observed.

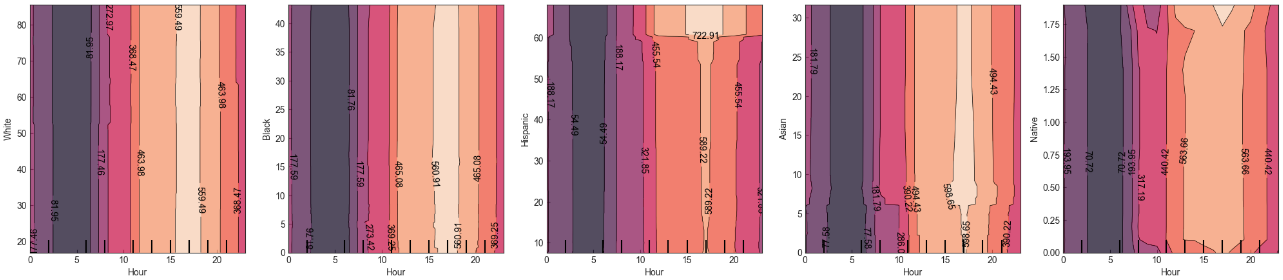


Figure 18: Partial dependence plot of E-scooter demand and the interaction of hour and different race excluding data in March.

7 Conclusion

In conclusion, our research focused on predicting hourly electric scooter demand in Austin through the development of five machine learning models: Linear and Ridge Regression, GPR, XGB, and NN. Among these models, XGB demonstrated the most accurate predictions for e-scooter demand. These high-resolution predictions can assist e-scooter service providers in optimizing device placement and inform public transportation policymakers. Additionally, we investigated the relationship between demographic features and scooter demand using partial dependence plots. While most racial groups displayed consistent demand fluctuations, the Hispanic population exhibited a unique pattern: demand increased when their population surpassed 60%. This finding could imply that areas with over 60% Hispanic population may be underrepresented in e-scooter distribution, despite the heightened demand. However, it is important to consider that partial dependence plots assume feature independence, which may not accurately represent real-world data. Therefore, further research is needed to better understand potential equity issues and ensure that transportation resources are distributed fairly and efficiently across different demographic groups.

8 Acknowledgment

The authors thank Dr. Arya Farahi for his valuable advice and guidance on this project.

Authors		
Ziyu Fan*	Jeil Oh*	Zizhe Jiang*
100%	100%	100%

*These authors contributed equally to this work.

References

[1] Elliot Fishman, Simon Washington, and Narelle Haworth. “Factors affecting the use of dockless bike-share and electric-scooter-share in North America”. In: *Transportation Research Interdisciplinary Perspectives* 8 (2020), p. 100246.

[2] Candace Brakewood, Tim Willems, and David Watling. “A model of the impact of land use and socio-demographics on shared micromobility ridership”. In: *Journal of Transport Geography* 79 (2019), p. 102477.

[3] Trinh Nguyen, Jeffrey R Bigham, and Aun Farooqi. “Data-driven analysis of factors influencing the demand for dockless shared electric scooters”. In: *Transportation Research Part C: Emerging Technologies* 111 (2020), pp. 245–263.

[4] Junhyuk Lee, Yongsuk Kim, and Abhishek Nayak. “Variability in the use of bike-sharing systems: A data-driven analysis”. In: *Journal of Transport Geography* 74 (2019), pp. 123–133.

[5] Jerome H Friedman. “Greedy function approximation: a gradient boosting machine”. In: *Annals of statistics* (2001), pp. 1189–1232.