Understanding Travel Time to Airports in New York City

Sierra Gentry | Dominik Schunack | {sgentry, schunack} @stanford.edu

Introduction and Motivation

Few things are more iconic to New York City than the yellow cabs that dot the city's ever present traffic. Even with the rising competition of rideshare services, many in New York still utilize the classic yellow cabs for their transportation needs, particularly when the airport is involved. However, as rideshare services continue to disrupt this aspect of the transportation industry, taxi services must adapt and transform to better serve their customers. They must be able to better predict traffic given various conditions which will be further discussed in the feature section. This transportation analysis uses machine learning techniques to more accurately predict travel time and trip origin and destination densities to allow the yellow cab fleets to better serve their customers and distribute themselves throughout the city, all while providing additional traffic insight to the city of New York. In particular, this analysis focuses on trips to and from the JFK International Airport. While the overall analysis can be applied to the entirety of the fleet, this subset was chosen to allow for a more meaningful interpretation of results to occur.

Data and Features

Two primary datasets were utilized for this analysis: the *New York City Taxi and Limousine Commission* provided transportation based information, while *Weather Underground* provided weather attributed to New York City between July 2016 and June 2017[1][2].

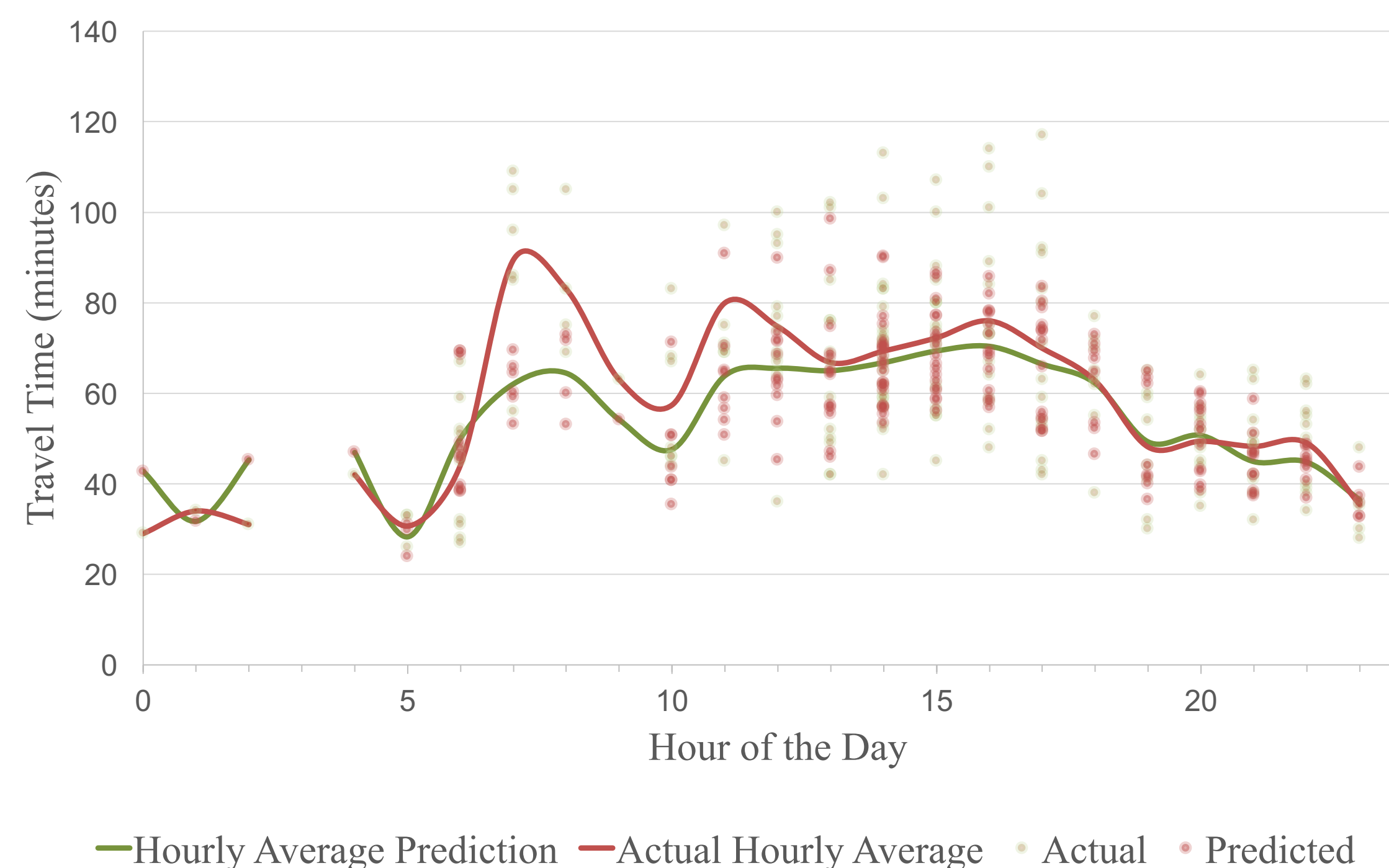| Original	JFK Only
10 million rides per month	400,000 rides per month

- Trip Origin and Destination (by District)
- Travel Time (Minutes)
- Travel Distance (Miles)

- Temperature (°F)
- Precipitation (Inches)
- Wind Speed (mph)
- Weather Condition (e.g. Overcast, Cloudy, Light Rain)

- Nationally Recognized Holidays
- +/- One Week for Holidays
- All Features Associated with Time

- Hour Before Trip Duration Average
- Day Before Trip Duration Average
- Trip Density by Location and Time

Results

Demand Density Map for 7 Districts in New York City For October 5th, 2016 at 3pm

- Actual Demand
- Predicted Demand

Actual and Predicted Travel Time From Midtown Manhattan to JFK in October 2016

— Hourly Average Prediction — Actual Hourly Average · Actual · Predicted

(y-axis: Travel Time (minutes); x-axis: Hour of the Day)

Models

Calculated similarly to an RMSE, the below error metrics are synonymous to an average time error; the respective model tends gives a result within the +/- error, in minutes.
These model results are with regards to travel time.

Model Type	Training Error	Testing Error
Linear Regression	10.64	12.59
Gradient Boosting	4.91	7.60
Principle Component Analysis	5.11	21.41
Single Hidden Layer NN	5.42	5.51

Discussion

- Hourly demand density is more predictable than travel time, as illustrated by the figures in the result section.
- The large number of factors made it difficult for many models to run.
- Airport travel time is more difficult to predict in the morning hours, indicating that the dataset is missing key features that may influence the travel time prediction during that time frame[3].
- The linear regression model does a fair job at picking up the daily trends such as the rush and lunch hours.

Future Work

Beyond helping the cabs, the City of New York could utilize the developed models and apply them to the entirety of the yellow cab fleet to better predict and plan for major events of traffic. If done correctly, trip density could be analyzed at an hourly level, supporting better designed public transit routes, along with promoting actual rideshare services that use predicted consumer demand to generate shared routes.

References

[1]"NYC Taxi & Limousine Commission - Trip Record Data", *Nyc.gov*, 2017. [Online]. Available: http://www.nyc.gov/html/tlc/html/about/trip_record_data.shtml. [Accessed: 11- Dec- 2017].
[2]"New York, NY Forecast | Weather Underground", *Weather Underground*, 2017. [Online]. Available: https://www.wunderground.com/weather/us/ny/new-york-city. [Accessed: 11- Dec- 2017].
[3]N. Laptev, J. Yosinski, E. Li, and S. Smyl, "Time-series extreme event forecasting with neural networks at Uber," International Conference on Machine Learning, 2017.