The 411 on 311:

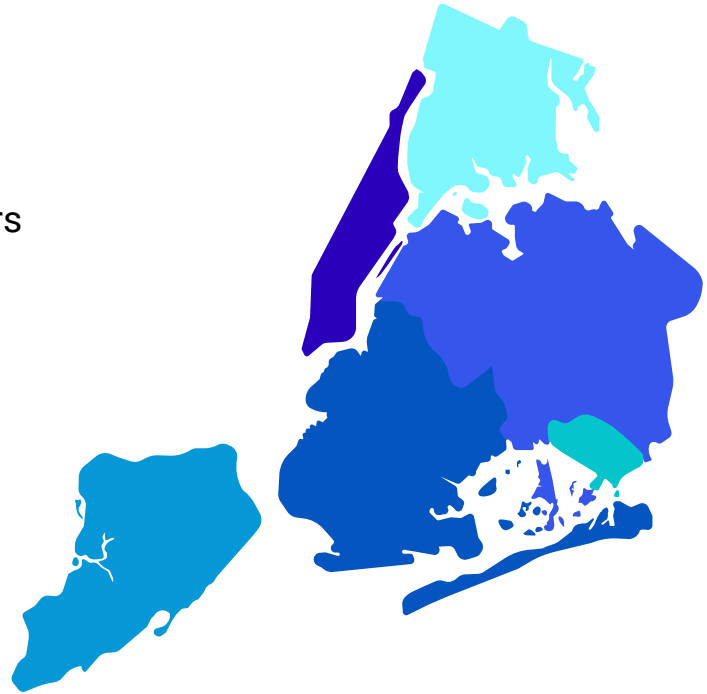
Predicting the Responding Government Agency for NYC 311 Service Requests

Data Sources

NYC OPEN DATA: 311 Service Requests from 2010 to Present

- Sample included 1.5 million requests (Jun.-mid Nov. 2020)

NYC DPT. OF CITY PLANNING: Community District Indicators



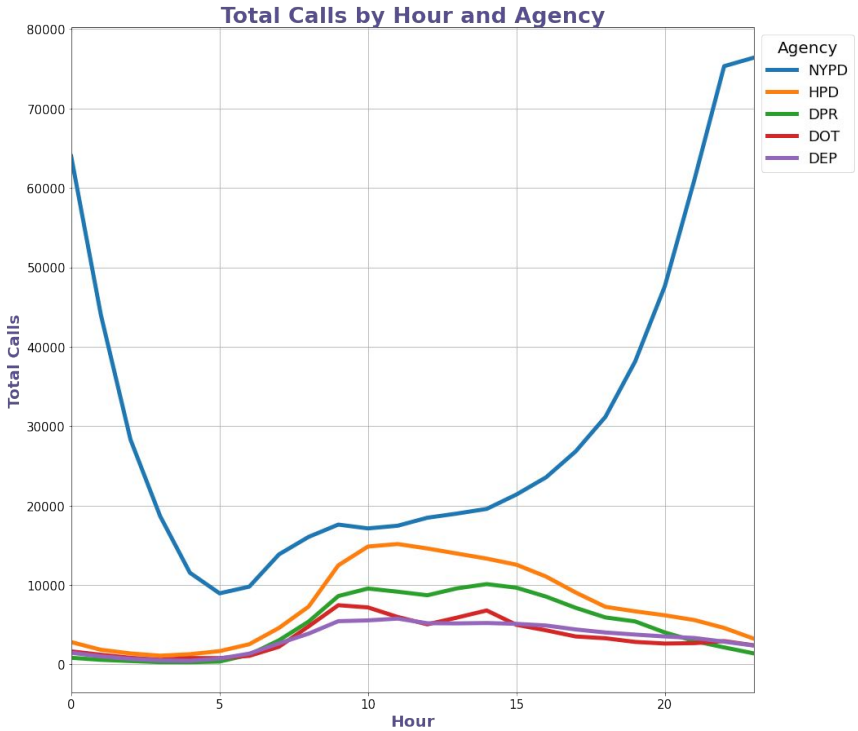
Most Frequent Words in Call Descriptors

311 Call Descriptor Word Cloud

Words related to the following appeared most frequently:

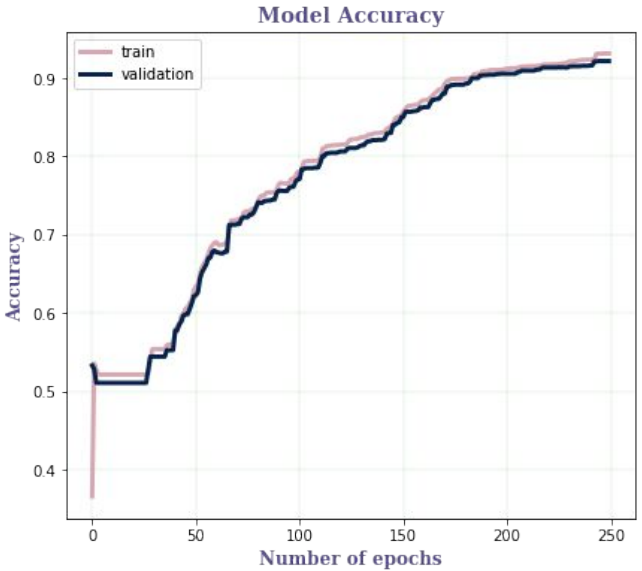
- noise
- illegal parking
- tree condition
- sidewalk/street condition

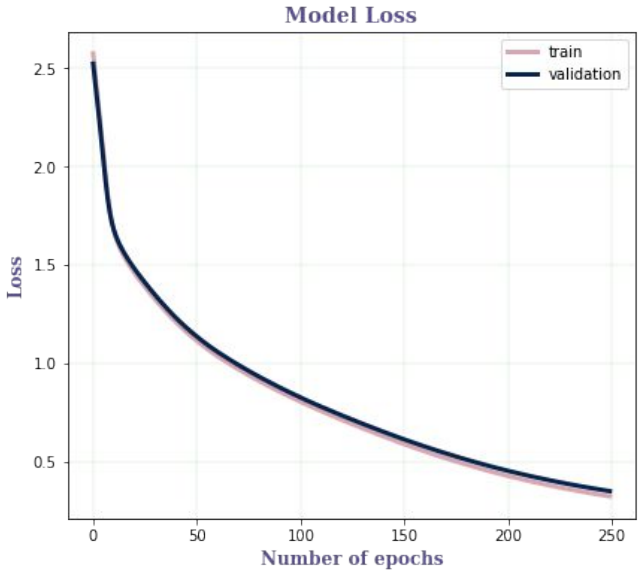
Hours with the Highest Call Volume



- Most calls assigned to the NYPD are made between 8PM and 5AM
- Most calls to other top agencies are made during the day

Keras Model Accuracy and Loss





- 92.6% accuracy on test data
- 72.8% accuracy on random subset
- Most successful with distinguishing between majority classes (NYPD, HPD, and DPR).

Dashboard

https://nyc-311.herokuapp.com/

Recommendations

Develop Classifier for Automatic Agency Referral

Using a larger, more diverse set of description training data

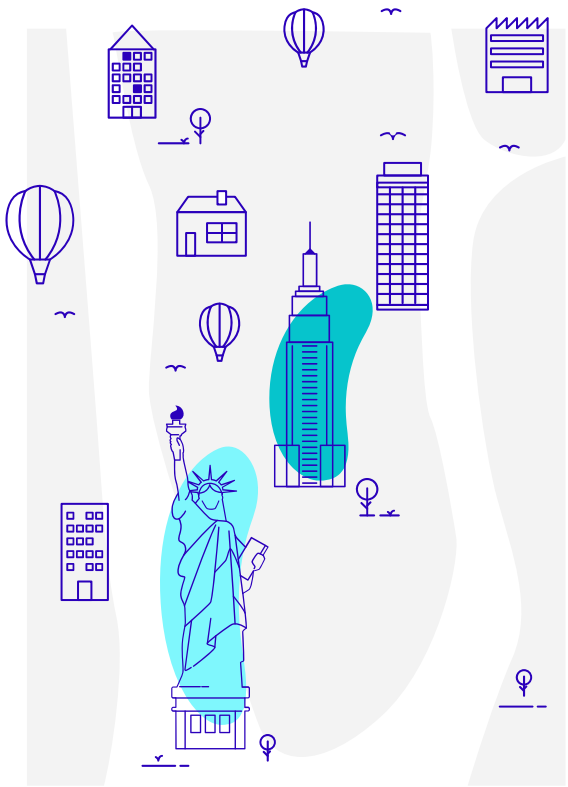
Prepare for Peaks in Call Volume

Following extreme weather events and at very early/late hours

Evaluate Current Distribution of Calls to Gov. Agencies

Should agencies other than the NYPD handle a larger share of requests?

Suggestions for Future Work



Evaluate Model Performance Using Additional Metrics

Even with all unique descriptions represented in the training and test samples, accuracy is an incomplete metric.

Collect Descriptor Data Generated by Requestors

This will facilitate the inclusion of a broader diversity of descriptions in the training data.

Train a Model to Predict Call Volume Based on Variables like Time and Location

This would allow responding agencies to prepare more effectively for high-volume days/times.

Thank you!

Presentation Template Credit

- **Slidesgo**
- **Freepik**