Dynamic Prediction of 311 Service Request Resolution Using Urban and Weather Data

Xiaotong Li
xiaotoli@kth.se

Dezhi Wen
dezhiw@kth.se

1 Background and Motivation

The New York City (NYC) 311 dataset records millions of service requests submitted by residents since 2010, covering issues related to noise, housing, transportation, sanitation, public safety, and more. Each record includes complaint type, responding agency, timestamps, location, resolution status, and resolution duration. The dataset is updated daily via the NYC Open Data platform and currently contains more than 30 million entries. It has become one of the most widely used open datasets for studying urban governance, service responsiveness, and data-driven public administration.

Since many 311 service requests require field personnel to travel to specific locations, weather conditions can meaningfully alter response efficiency. Poor weather—such as rain, snow, or storms—can delay travel times, hinder outdoor work, and increase the backlog of incidents, making resolution slower than usual. That is why we also add weather data.

Building on this context, the goal is to develop a dynamic, continuously updated AI prediction system that integrates 311 service request data with weather data. The system will produce real-time predictions at the moment a new request is created, estimating whether it will be resolved within 48 hours and the expected resolution time. The system could support government agencies and public service managers in operational prioritization and resource allocation, and also informs clients the estimated executing time.

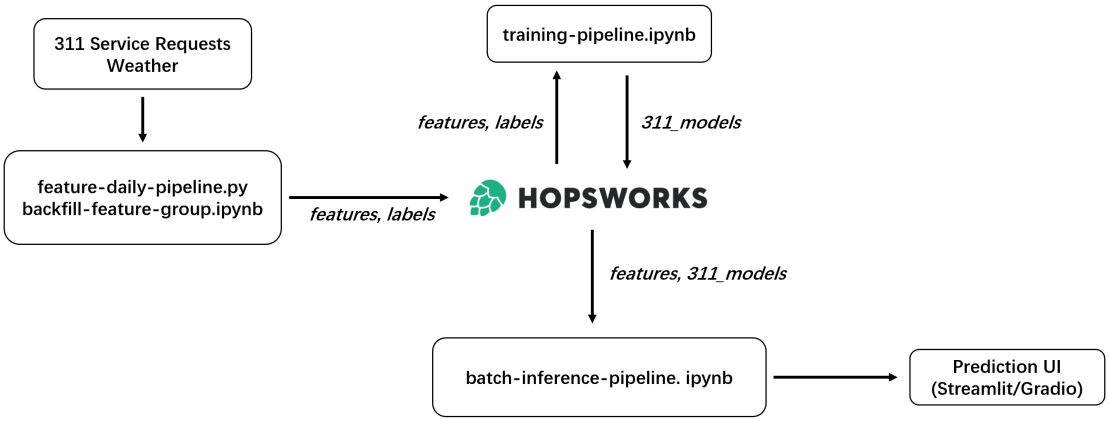

Figure 1: Overview of the data pipeline.

2 Project Objectives

1. Predict whether an individual 311 service request will be resolved within 48 hours by leveraging complaint type, agency, temporal patterns, spatial features, and historical performance.
2. Estimate the expected resolution time (in hours) for each newly created request.
3. Design and implement a dynamic data pipeline that retrieves daily updates from the NYC 311 API and weather APIs, automatically processing incremental data and updating features and models.
4. Develop a lightweight web-based interface (Streamlit/Gradio dashboard) to display prediction outcomes and system updates.

3 Methods and Technical Approach

3.1 Data Preprocessing

- **Missing value handling**: Remove records missing essential timestamp fields; encode missing location or agency information.
- **Time parsing and normalization**: Standardize time zones, convert timestamp strings into `Datetime` format, and compute true resolution duration.
- **Outlier filtering**: Remove or flag cases with negative or extremely long resolution times (e.g., greater than 60 days).

3.2 Classification Model: Predicting 48-Hour Resolution

Candidate models:

- XGBoost Classifier, Random Forest Classifier, CNNs

Evaluation metrics:

- Precision, Recall, F1-score

3.3 Regression Model: Predicting Resolution Duration

Models:

- XGBoost Regressor, Random Forest Regressor and CNNs

Evaluation metrics:

- RMSE, R^2.
- Separate performance for short-duration (< 48h) and long-duration (> 48h) case

4 Expected Contributions and Innovations

4.1 Practical Contributions

- Provide city managers with a prototype system capable of real-time risk assessment at the moment a service request is filed.

4.2 Methodological and Data Contributions

- Move beyond one-time offline analysis by emphasizing dynamic updates and real-time prediction.
- Systematically compare models trained on 311-only features versus 311 + weather features, exploring how environmental conditions interact with service performance.

4.3 Technical Contributions

- Develop an end-to-end pipeline integrating data ingestion, feature engineering, modeling, evaluation, and visualization.
- Provide a reusable template for similar AI systems supporting urban governance and smart-city analytics.