Transit Hacks Datathon

RideReady: Predicting Chicago Bus Delays

Daniel, Jazil, Karim, Sakshi

RideReady: Predicting Chicago Bus Delays

Public transit reliability is a major concern for urban commuters. Understanding and predicting these delays can significantly improve commuter experiences and help the Chicago Transit Authority optimize its operations.

We aim to **predict whether a bus will be delayed**, leveraging data from multiple sources:

- **Bus GPS data** capturing movement from **April 10–24, 2024**
- **Bus stop information** (location name)
- **Weather data** (temperature, conditions like rain, snow, etc.)

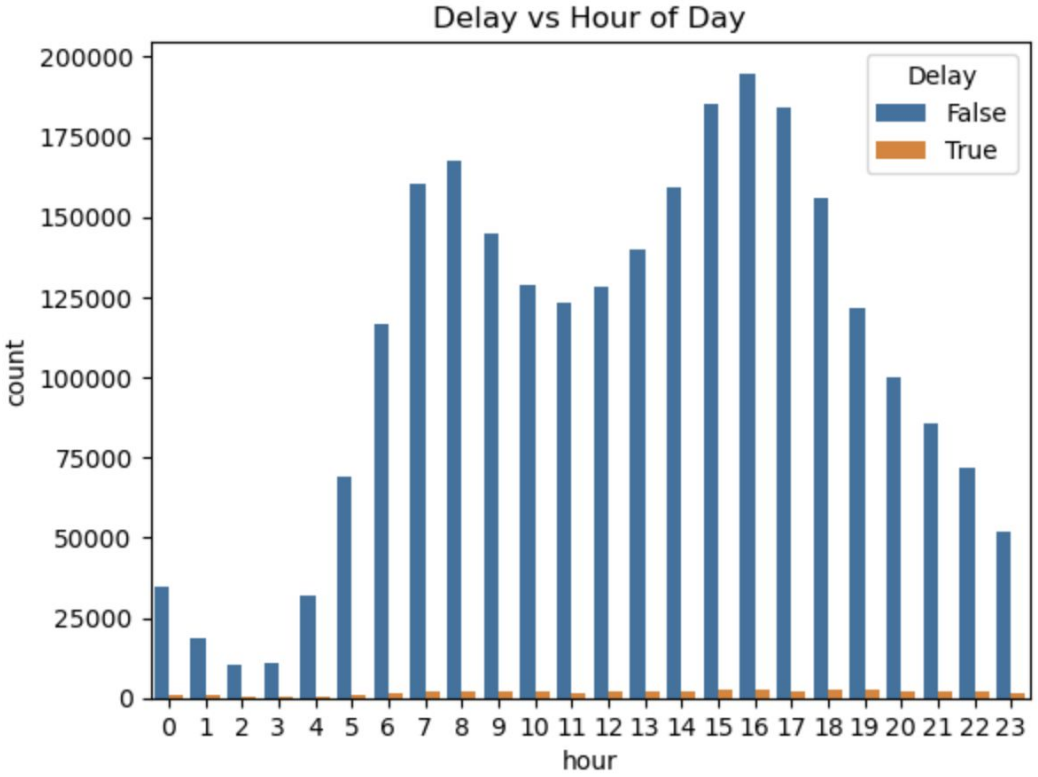
The objective is to build an accurate predictive model that allows the CTA and commuters to anticipate delays and adjust plans proactively.

Better predictions → Better decisions → Happier riders + More efficient city

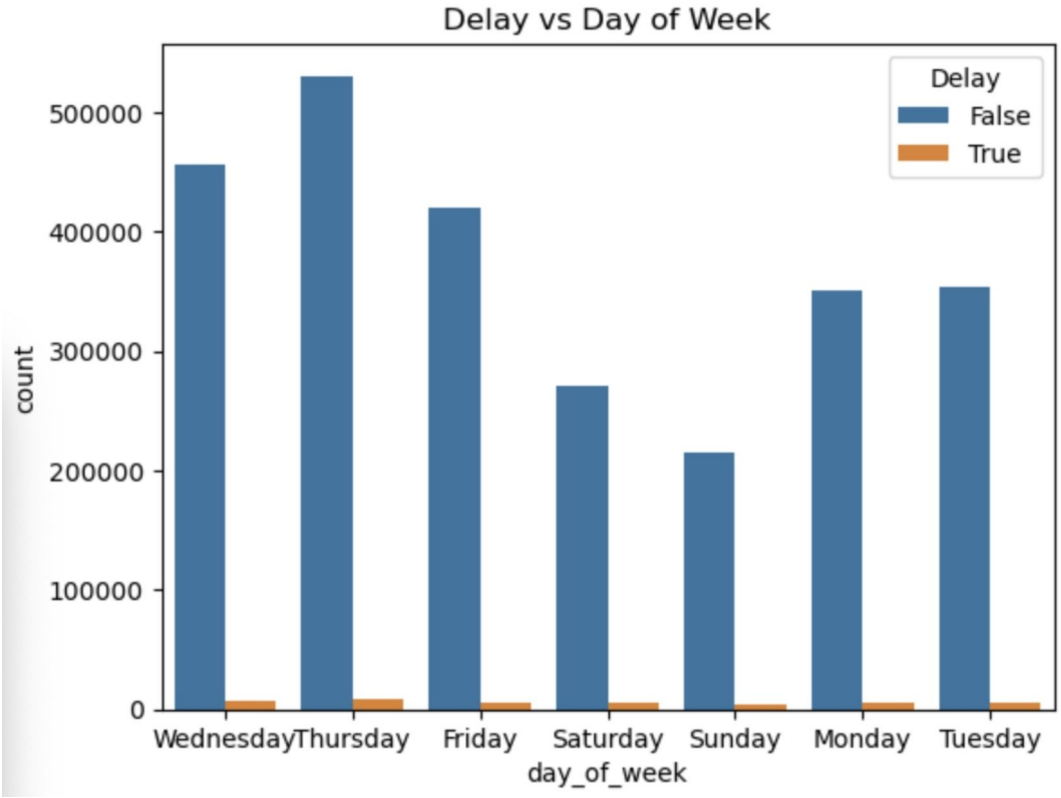
Data Glance

Category	Example Columns
📍 Location Data	`Route`, `Destination`, `nearest_stop_name`, `distance_to_stop`
🕒 Time Data	`Timestamp`, `time`, `year_month`, `is_day()`
🌦️ Weather Data	`temperature_2m (°C)`, `rain (mm)`, `snowfall (cm)`, `snow_depth (m)`, `wind_speed_10m (km/h)`, `cloud_cover (%)`, etc.
🎯 Target	`Delay` (boolean)

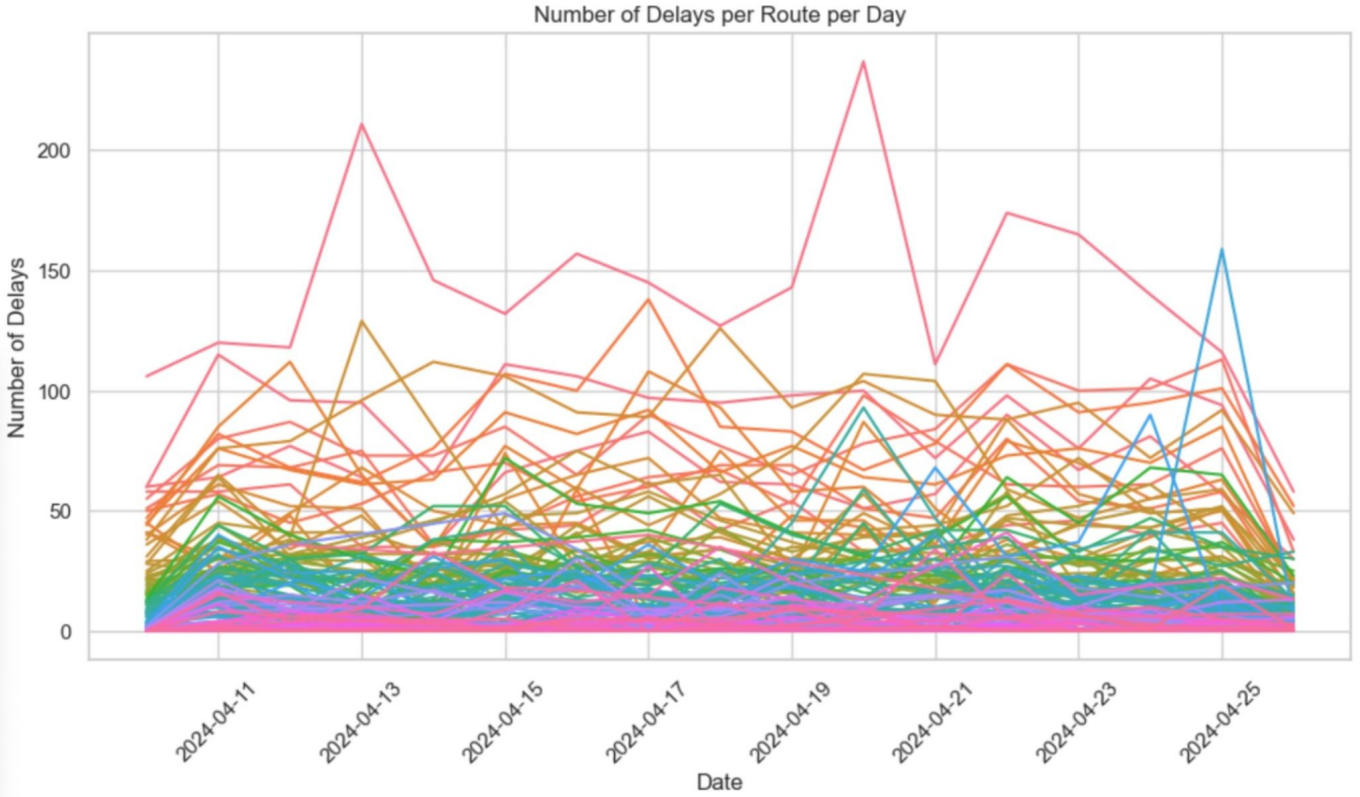
Delay Trends by Hour of Day



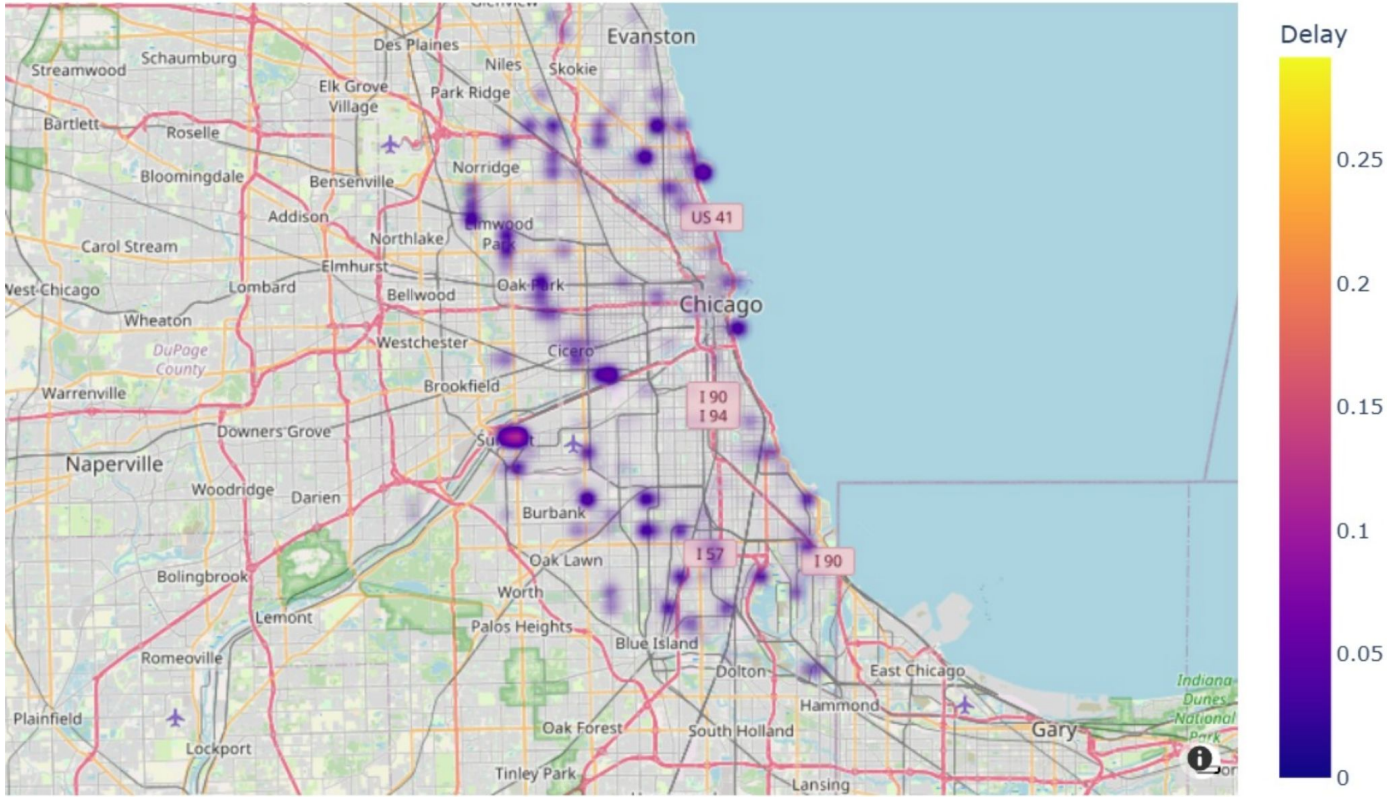
Delay Trends by Day of Week



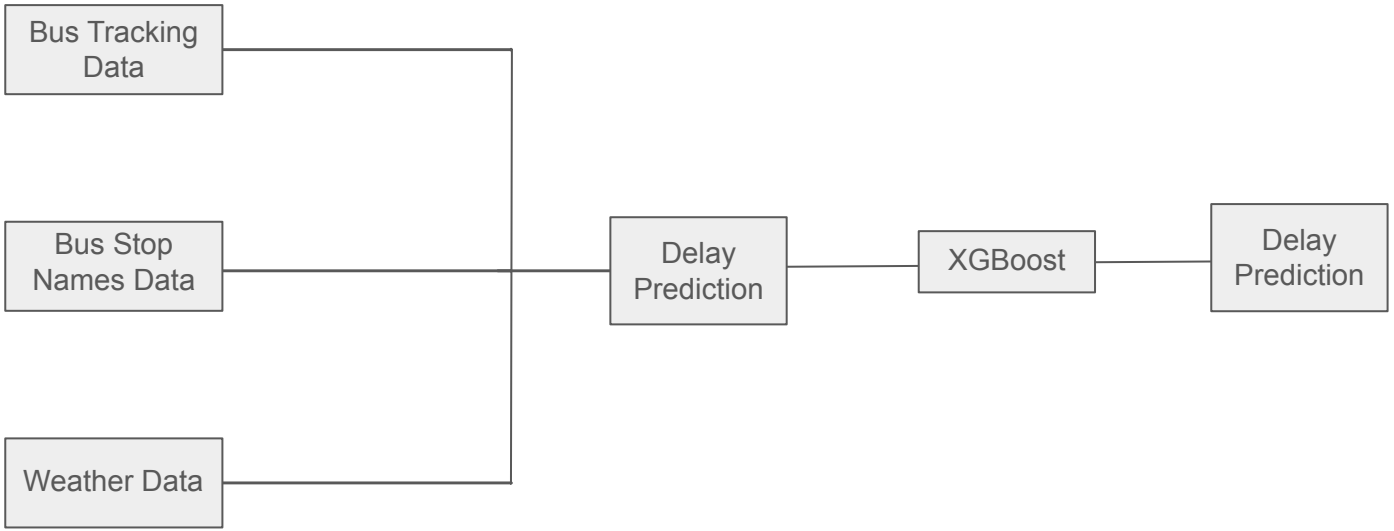
Delay Trends by Route Over Time



Geographic Distribution of Bus Delays in Chicago



Our Approach



Conclusion

Our model demonstrates strong performance in predicting delays, achieving a recall of **93%**, which indicates its high effectiveness in correctly identifying delayed instances.

We plan to develop a mobile application that allows users to input key information such as Route, Temperature, and Time to predict whether their bus is likely to be delayed.