

BUSINESS PROBLEM — Ride Demand Prediction System

Ride-hailing companies must match driver supply with passenger demand.

Incorrect demand estimation leads to:

- a).Long passenger wait times**
- b).Driver idle time**
- c).Surge pricing imbalance**
- d).Revenue loss and poor customer experience**

A Ride Demand Prediction System helps forecast ride requests in advance across locations and time periods.

1. Business Objective:

Primary Goal:

Predict future ride demand to optimize fleet allocation and operational efficiency.

Objectives:

- Forecast ride requests by location and time
- Reduce passenger waiting time
- Improve driver utilization
- Optimize dynamic pricing strategies
- Support operational planning

Business Actions Based on Predictions:

- Reposition drivers to high-demand zones
- Activate surge pricing when needed
- Offer driver incentives in predicted hotspots
- Improve dispatch planning

Business Success Criteria:

- Reduction in average passenger wait time (e.g., 15–25%)
- Increased ride completion rate
- Higher driver utilization rate
- Improved customer satisfaction ratings
- Reduced cancellation rates

2. Assess Situation:

Inventory of Resources:

- Historical ride booking data
- GPS location data
- Timestamp information
- Driver availability records
- Weather data
- Traffic data
- Event and holiday information
- Existing analytics and operations teams
- Cloud infrastructure and databases

Requirements:

- Accurate short-term demand prediction
- Near real-time prediction capability
- Scalable system for multiple cities
- Interpretable insights for operations teams
- Fast inference for live dispatch systems

Assumptions:

- Historical demand patterns reflect future trends
- External factors (weather/events) influence demand

- Location and time strongly affect ride requests
- Data collection systems are reliable

Constraints:

- Data privacy and location security regulations
- Sudden demand spikes (concerts, emergencies)
- Cold-start problem in new areas
- Missing or noisy GPS data
- Real-time computation limitations

Costs and Benefits:

Costs

- Data storage and processing infrastructure
- Model development and maintenance
- Engineering and data science manpower
- Real-time deployment systems

Benefits

- Better supply-demand balance
- Increased revenue
- Improved customer experience
- Reduced driver idle time
- Efficient operational planning

3. Determine Data Science Goals

Data Science Objective (Technical View):

Build a prediction model to estimate: Number of ride requests per location & time interval

(e.g., next 15 minutes / next hour)

Data Science Tasks:

- Data preprocessing and cleaning
- Feature engineering (time, location, weather)
- Handling missing data
- Time-series modeling
- Train models such as:
 - Linear Regression
 - Random Forest
 - XGBoost
 - LSTM / Time-Series models

Data Science Success Criteria:

- Low prediction error (RMSE / MAE)
- Accurate peak-demand prediction
- Stable performance across locations
- Real-time prediction latency within limit.

4. Produce Project plan (overview):

Stages	Activity	Duration
1.	Business Understanding	1 week
2.	Data collection & understanding	2 weeks
3.	Data cleaning & preparation	2 weeks
4.	Model Building	3 weeks
5.	Model Evaluation	1 week
6.	Deployment	1 week
7.	Monitoring and Maintenance	continuous

5.Final Outcome:

- A deployed Ride Demand Prediction System
- Improved driver allocation decisions
- Reduced waiting time
- Data-driven operational optimization.

Ride demand Heatmap visualization:

