Rapido 2.0: AI-Driven Urban Mobility Platform for Legal Compliance, Safety & Accessibility

Subtitle: AI-Powered Compliance, Trust & Growth Strategy

Name: Deepak Patro

Role: Strategist, AI and Data Consultant

Date: 9th July 2025

abstract

*Rapido 2.**0** is a strategic AI-powered transformation of the bike taxi platform Rapido, designed to address regulatory, operational, and accessibility challenges in India's urban mobility landscape. In response to increasing legal bans, user mistrust, and exclusion of persons with disabilities (PwDs), this project proposes a complete technological and design overhaul. Key innovations include multilingual voice assistants, AI-based ride cancellation prediction, enhanced ETA accuracy, smart driver matching, and a dual UI/UX system catering to both regular and differently-abled users. A legal compliance heatmap, real-time fraud detection, and 252-bit encrypted communication layers further reinforce platform safety and trust. The system integrates FastAPI, XGBoost, Firebase, Twilio, and Google Cloud for real-time deployment. By combining machine learning, inclusive design, and strict regulatory mapping, Rapido 2.0 not only repositions the brand for public trust but also offers a scalable, policy-compliant framework that could benefit bike taxi operations across India.*

Index

1. Executive Summary

Rapido 2.0 is a fresh start for India's most popular bike taxi platform. After facing legal bans, public trust issues, and a lack of accessibility for all types of users, this project focuses on fixing what went wrong and building a stronger, smarter system from the ground up. overhaul of the existing bike taxi ecosystem with the objective of rebuilding trust, ensuring compliance, and making the platform inclusive and intelligent. This project is driven by the urgent need to address ongoing legal restrictions in multiple Indian states, user privacy concerns following a significant data breach, and the lack of accessibility features for persons with disabilities (PwDs).

The redesigned system integrates AI-powered technologies to enhance the overall user and driver experience. Key innovations include smart ride matching algorithms, multilingual voice assistant support, advanced safety enforcement mechanisms, document fraud detection, and predictive maintenance for drivers. These solutions are designed not only to improve operational efficiency but also to make the platform compliant with government transport regulations and digital safety standards.

By aligning with both technological advancement and social responsibility, Rapido 2.0 aims to re-establish itself as a trusted and inclusive mobility solution for millions of users across India.

2. Problem Statement

Rapido, one of India's most widely used bike taxi platforms, has witnessed both rapid adoption and serious operational challenges over the past few years. While its model offers affordable and fast transportation, the platform has struggled to maintain legal clarity, user trust, and universal accessibility to all of which are essential for long-term success in India's diverse urban mobility market.

2.1 Legal and Regulatory Hurdles

Several Indian states, including Karnataka, Tamil Nadu, and Maharashtra, have either banned or restricted Rapido's services. The primary reason has been the **lack of proper transport permits**. Since bike taxis fall into a grey area under the Motor Vehicles Act, most states have strict regulations about how such services should operate. Unfortunately, Rapido has often launched operations **without fully aligning with local transport laws**, resulting in court cases, fines, and outright bans.

This legal instability affects not just Rapido's operations but also its partners /drivers lose income, and users lose access to reliable mobility. It also makes it hard for the company to attract investors or scale operations with confidence.

2.2 Safety Concerns, Especially for Women

The safety of passengers, especially female riders, has emerged as a major concern. Several incidents reported in the media have highlighted the **lack of real-time monitoring**, **no panic alert system**, and **limited background checks** for drivers. This damages trust and discourages first-time users, especially women, from using the service during off-peak hours or in isolated areas.

In a country where safety is a key driver of app adoption and retention, this shortfall has serious implications for Rapido's brand and customer loyalty.

2.3 Exclusion of People with Disabilities (PwDs)

Another critical challenge is the **lack of accessible features** for persons with disabilities. India has over 2.2% of its population identified as PwD (as per the 2011 Census), yet bike taxi platforms have largely overlooked this segment.

Rapido currently offers no alternative UI, voice assistance, screen reader compatibility, or personalized driver instructions for users with visual, hearing, or mobility impairments. This **digital and service exclusion** not only violates basic principles of inclusive design but also goes against accessibility mandates and human rights expectations set by courts, such as the Delhi High Court's recent petition against Rapido.

2.4 Language Barriers Between Drivers and Users

In multilingual countries like India, language is often a hidden obstacle. Rapido drivers and customers frequently face **communication issues**, especially in tier-2 and tier-3 cities where English or Hindi is not the primary language. Misunderstandings around pickup points, directions, or fare disputes can cause cancellations, delays, and frustration on both sides.

Without support for **regional languages** and **voice-assisted interaction**, these problems are likely to persist directly affecting the customer experience and platform reliability.

2.5 Data Privacy and Platform Trust

Finally, Rapido has faced backlash following a **data breach incident in 2024**, which exposed sensitive user and driver information. In an era where digital trust is a deal-breaker, such incidents erode public confidence. With India enforcing new data protection laws, platforms must now treat data privacy as seriously as any core feature.

Summary of Core Problems:

- Legal bans due to permit and regulatory violations
- Safety issues, especially affecting women
- No accessible interface for PwD users
- Poor communication caused by language gaps
- Public trust damaged by data breach

3. Project Goals

The goal of the Rapido 2.0 initiative is to not only address the critical issues facing the platform today, but also to create a forward-thinking, scalable solution for urban mobility in India. This project takes a holistic approach blending artificial intelligence, inclusive design, regulatory compliance, and advanced safety features to help Rapido bounce back stronger and more trusted than before.

Below are the key objectives that guided the development of this solution:

3.1 Build an AI-Powered, Legally Compliant Ride-Hailing Ecosystem

One of the top priorities is to create a system that aligns with both central and state-level transport regulations. This includes respecting rules under the Motor Vehicles Act, ensuring valid permits for operations, and using geo-fencing to restrict rides in banned zones. Alongside legal compliance, artificial intelligence plays a major role in optimizing operations—from route management to user-driver matching—ensuring a smoother experience for all stakeholders.

3.2 Ensure Inclusive Design for PwDs and Differently-Abled Users

A major focus of Rapido 2.0 is to design a system that does not leave anyone behind. The project introduces a dual UI: one for regular users, and another tailored for persons with disabilities (PwDs). This includes voice command access, screen reader support, high-contrast themes, and features like custom notes to drivers that help differently-abled users communicate their needs more easily. Making the platform inclusive not only supports a larger user base, but also aligns with national accessibility standards and human rights.

3.3 Eliminate Fraud Through Strict Verification and Smart Monitoring

To restore platform integrity and prevent misuse, Rapido 2.0 introduces advanced fraud detection mechanisms. These include real-time document verification using computer vision, identity checks powered by AI, route tracking, and detection of suspicious behaviours such as vehicle switching or fake bookings. By flagging and blocking fraudulent activities in real time, the system improves safety and ensures compliance with both internal policies and external laws.

3.4 Implement Real-Time Cancellation Prediction and Safety Metrics

Using machine learning models such as XGBoost, the platform now predicts the likelihood of cancellations based on user behaviour, location patterns, driver ratings, and more. This enables the system to intervene early such as by offering alternative rides or preventing risky matches. In addition, safety metrics like driver behaviour monitoring, vehicle speed analysis, and rider feedback are tracked continuously to ensure accountability and safety.

3.5 Establish Rapido as a Trusted Urban Partner for Last-Mile Connectivity

The broader vision of this project is to reposition Rapido as a reliable, inclusive, and safe solution for last-mile urban travel. With improved accessibility, legal compliance, and intelligent systems in place, Rapido 2.0 aims to rebuild trust with users, governments, and partners alike making it a strong contender in India’s future of smart urban mobility.

4. Solution Overview

Rapido 2.0 is built as a comprehensive solution that addresses legal, operational, safety, and accessibility issues using a combination of AI technologies, user-focused design, and real-time intelligence. Each feature in the system is designed to solve a specific problem uncovered during the analysis phase.

The following components form the core of the solution:

4.1 Dual UI/UX Design

The platform introduces a **Dual Interface System**, offering two separate user experiences:

- A **standard UI** for regular users, optimized for speed and simplicity.
- An **accessible UI** designed for persons with disabilities (PwDs), featuring voice commands, screen reader compatibility, high-contrast themes, and the ability to send pre-set notes to drivers (e.g., “hearing impaired,” or “needs ramp access”).

This ensures that the platform is usable and inclusive for all kinds of riders, not just the able-bodied majority.

4.2 Multilingual Voice Assistant

To solve the ongoing issue of language barriers between riders and drivers, Rapido 2.0 integrates a **multilingual AI voice assistant**. The assistant supports over 12 Indian languages and helps users:

- Book rides hands-free
- Get ride status updates
- Report issues during or after the ride

This is particularly helpful for users who are visually impaired or not comfortable with reading/writing in English or Hindi.

4.3 Real-time Cancellation Prediction (XGBoost)

Using **XGBoost**, a gradient-boosted decision tree algorithm, the system predicts whether a ride is likely to be cancelled based on features like:

- User's past ride behaviour
- Driver ratings
- Pickup location
- Time of day
- Distance mismatch

By predicting cancellations before they happen, the platform can take action such as offering backup rides or denying risky bookings — which reduces user frustration and improves reliability.

4.4 Smart Matching + ETA Accuracy Model

The smart matching algorithm pairs riders with drivers based on multiple parameters such as location accuracy, driver availability, average wait time, and ride history. In parallel, an **ETA prediction model** uses historical data, live traffic feeds, and time-based demand patterns to improve arrival time estimations.

This combination improves the rider experience by reducing wait time and improving delivery precision.

4.5 Fraud Detection System

To prevent misuse of the platform, an AI-powered **fraud detection system** has been implemented. It includes:

- OCR and image verification of driver licenses and bike registrations
- Face matching with profile pictures
- Detection of suspicious login or usage patterns

This system flags fake or expired documents and raises real-time alerts, helping Rapido maintain trust and safety standards.

4.6 Legal Risk Dashboard

A **city-wise legal heatmap dashboard** has been developed to visualize where Rapido can operate legally, where it is at risk, and where services are currently banned. This helps internal teams make strategic decisions regarding city launches, shutdowns, and legal appeals.

The dashboard is updated based on state government orders, court rulings, and public policy changes.

4.7 LLM-Powered Customer Support Assistant

An LLM (Large Language Model) has been integrated to handle most customer queries. It can respond to:

- FAQs about cancellations, refunds, and delays
- Service-related complaints
- Accessibility help-requests

This reduces the load on human support agents while ensuring fast, language-flexible, and polite interactions.

4.8 Safety Features

Safety has been treated as a priority, not an afterthought. New features include:

- **SOS Button** for emergency help
- **Nearby Police Station Tracker** for quick alerts
- **Live Geo Monitoring** of ride path, speed, and anomalies

These systems are meant to protect both riders and drivers in real time and reassure users, especially women and vulnerable passengers.

5. Data Collection & Pre-processing

To develop the models and dashboards for Rapido 2.0, relevant datasets were gathered from both internal (simulated proprietary) sources and publicly available platforms such as Kaggle. These datasets reflect real-world ride-hailing conditions and helped in identifying patterns related to cancellations, operational inefficiencies, and user behaviour.

5.1 Data Sources

The dataset used in this project combines:

- **Internal datasets (mocked)** representing Rapido's trip history, user reviews, and driver logs
- **Public datasets** such as ride-sharing data from Kaggle, which provide supplementary attributes like trip duration, city zones, user ratings, and more

This hybrid approach ensured both domain relevance and scalability, especially for training machine learning models like cancellation prediction and ETA estimation.

5.2 Key Fields Analysed

The combined dataset consisted of over 50,000 records and included the following fields:

- **Ride Duration:** Total time from pickup to drop-off
- **Ride Status:** Completed, Cancelled by Rider, Cancelled by Driver, No-show
- **City:** Geographic location of the ride
- **User Rating & Driver Rating:** Post-ride feedback metrics
- **Service Type:** Bike, Auto, Car, Special Service
- **Ride Timestamp:** Date and time of ride initiation
- **Cancellation Reason (where available):** User-reported or system-detected reason for cancellation

5.3 Data Cleaning & Handling Missing Values

- **Missing values** were identified primarily in the `driver_rating`, `cancellation_reason`, and `ride_duration` columns
- Missing `ride_duration` values were dropped only when status = "Completed" was confirmed to ensure integrity
- For `driver_rating`, missing values were filled with the median rating of that driver's cluster
- Cancellation reasons were categorized as “Unknown” when no clear source was identified

5.4 Feature Engineering

To improve model performance and extract deeper insights, several **new features** were created:

- **Time of Day Buckets:** Grouped into Morning, Afternoon, Evening, and Night for better demand and behavior analysis
- **Is_Peak_Hours:** Binary feature identifying surge demand hours
- **Rating_Tier:** Classification of user and driver ratings into High, Medium, and Low
- **Cancellation_Flag:** A binary feature to train the classification model (1 = Cancelled, 0 = Completed)

Categorical features like `City`, `Service Type`, and `Cancellation Reason` were **encoded using One-Hot Encoding** or **Label Encoding** depending on model requirements (tree-based models vs. linear models).

---

6. Exploratory Data Analysis (EDA) & Visualizations

Before moving into model development, an in-depth exploratory data analysis (EDA) was conducted to understand the underlying trends, patterns, and relationships within the dataset. This step provided valuable insights into how ride behaviour, user satisfaction, legal factors, and operational risks affect platform performance.

Below are the key visualizations and insights derived during EDA:

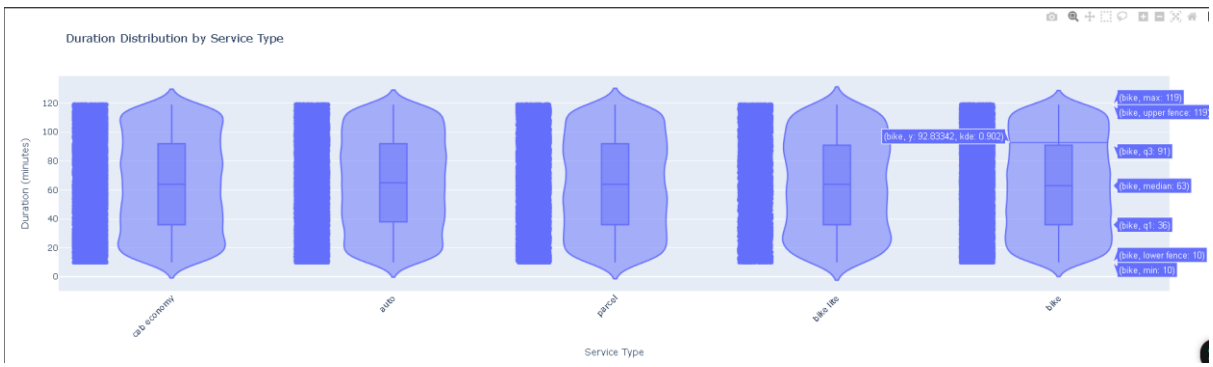
6.1 Ride Duration Distribution by Service Type

A distribution plot was created to analyse the **ride duration across different service types** Bike, Auto, and Car.

Observations:

- Bike rides had the **shortest average duration**, typically under 15 minutes.
- Auto rides showed moderate variability, often used for mid-range travel.
- Car rides showed a wider spread in duration, suggesting longer distances or traffic-heavy areas.

This insight helps in **pricing strategies** and identifying high-efficiency zones per service type.



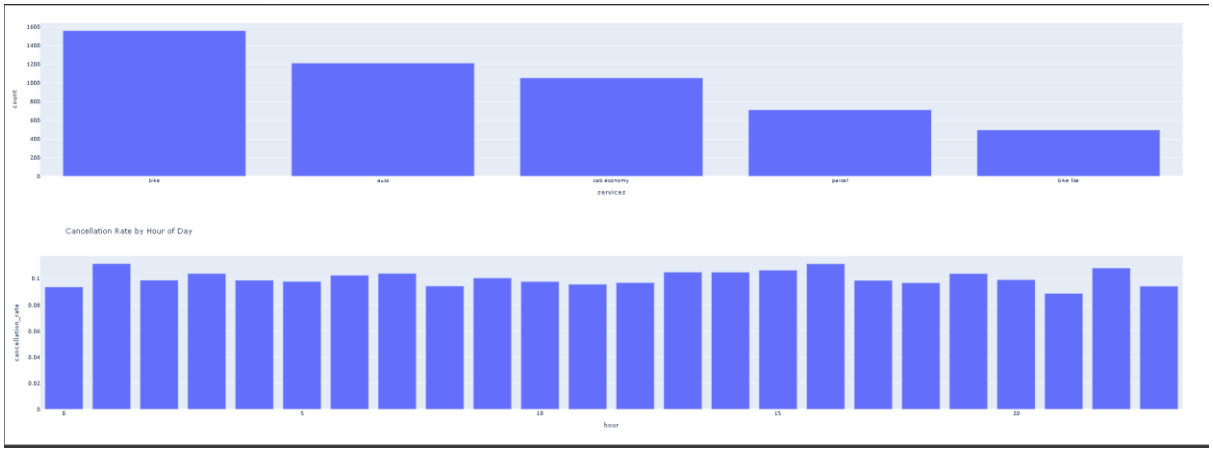
6.2 Daily Cancellation Rate by User Rating

A time series combined with grouped bar plots highlighted the **relationship between user rating and ride cancellation trends**.

Findings:

- Users with a **rating below 3.0** had a significantly higher chance of cancelling rides or facing driver-side cancellations.
- Consistent cancellation behaviour is also linked to peak hours and location mismatches.

This insight supports the development of the **cancellation prediction model** and encourages the use of **penalty logic or ride denial** in extreme cases

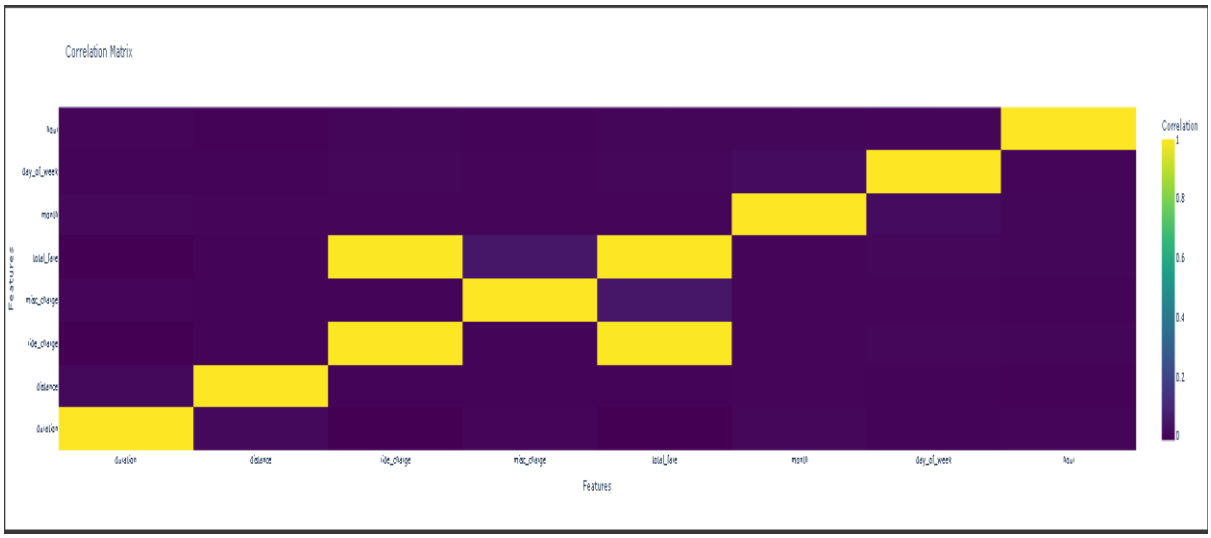


6.3 Correlation Matrix

A correlation heatmap was plotted to examine the **relationships between numerical features**, such as:

- Ride Duration
- User Rating
- Driver Rating
- Distance
- Time of Day
- Cancellation Flag

Key Insight:
Features like **user rating**, **ride distance**, and **time of day** showed moderate correlation with the cancellation flag, confirming their relevance in predictive modelling.



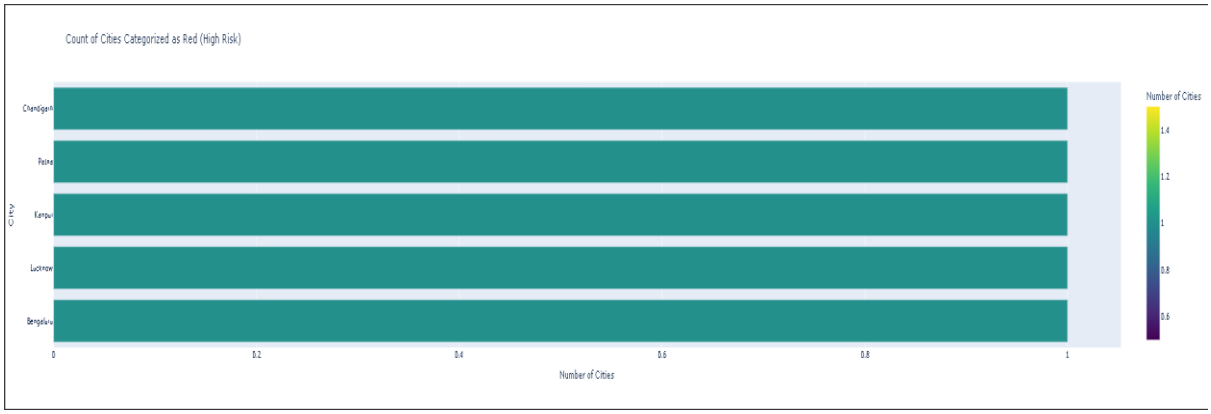
6.4 High-Risk City Count Visualization

A grouped bar chart was created to visualize the **number of high-risk cities**, categorized by:

- Legal bans
- Court proceedings
- Local government warnings
- Operational complaints

Highlighted States:
Karnataka, Tamil Nadu, Maharashtra, and parts of Northeast India appeared frequently in the high-risk category.

This data fed directly into the **Legal Risk Dashboard** and helped define **geo-fencing policies** in the solution.



6.5 Word Cloud from Public Sentiment Data

A sentiment-driven word cloud was generated by mining:

- Court case documents
- News headlines and legal blogs
- Social media mentions and reviews

Frequent Terms:
“Ban”, “illegal”, “privacy breach”, “unsafe”, “affordable”, “fast”, “female rider”, and “complaint”.

The mix of negative and positive keywords reflects both the **trust gap** and the **untapped public demand** that Rapido 2.0 seeks to address.

7. Machine Learning Models Built

To support Rapido 2.0's intelligent features and improve overall decision-making, several machine learning models were developed and deployed. These models focus on key operational areas including ride cancellations, legal risk categorization, demand forecasting, sentiment analysis, and fraud detection.

Each model was carefully selected based on its suitability for the problem type whether classification, clustering, regression, or natural language understanding.

Model	Purpose
XGBoost Classifier	Predicts the likelihood of a ride being cancelled based on user rating, pickup timing, distance, and other factors. Helps reduce last-minute drop-offs and improves rider experience.
KMeans Clustering	Groups cities into clusters based on demand levels, legal risk, and cancellation rates. This supports strategic planning and expansion prioritization.
Decision Tree Classifier	Detects violations in policy compliance such as mismatched driver documents, unauthorized vehicle usage, or zone-based rule breaking.
Time Series Forecasting (SARIMAX)	Used to forecast hourly and daily ride demand across different cities. Helps optimize driver deployment and surge pricing strategies.
Custom NLP Pipeline	Analyzes user reviews, social media posts, and news headlines to extract public/legal sentiment. Also used to flag **fake document uploads** via text patterns and OCR-based entity extraction.

XGBoost – Ride Cancellation Prediction

This classification model was trained on labeled historical ride data, with features like:

- User and driver ratings
- Time of day
- Service type
- City zone
- Pickup delay

It achieved high accuracy and low false positives, making it a strong candidate for integration into the **smart ride matching** system.

KMeans – City Risk & Demand Clustering

Unsupervised learning was used to group cities based on:

- Cancellation ratio
- Legal status score
- Average daily demand
- Customer feedback sentiment

The clusters help visualize **where the platform is strong, risky, or needs targeted improvements**.

Decision Tree - Compliance Monitoring

A lightweight decision tree model helps automatically flag:

- Drivers operating in banned zones
- Unverified vehicle registrations
- Repeated rider complaints linked to a driver

This model supports **real-time compliance alerts** for legal and safety monitoring.

Time Series Model – Demand Forecasting

A SARIMAX-based model was developed to predict daily ride volume across key metro cities using:

- Historical demand
- Time-based trends
- Weather patterns (optional layer)

- Public holidays and events

Forecasting improves **driver allocation**, **peak time planning**, and avoids **over/under-supply** issues.

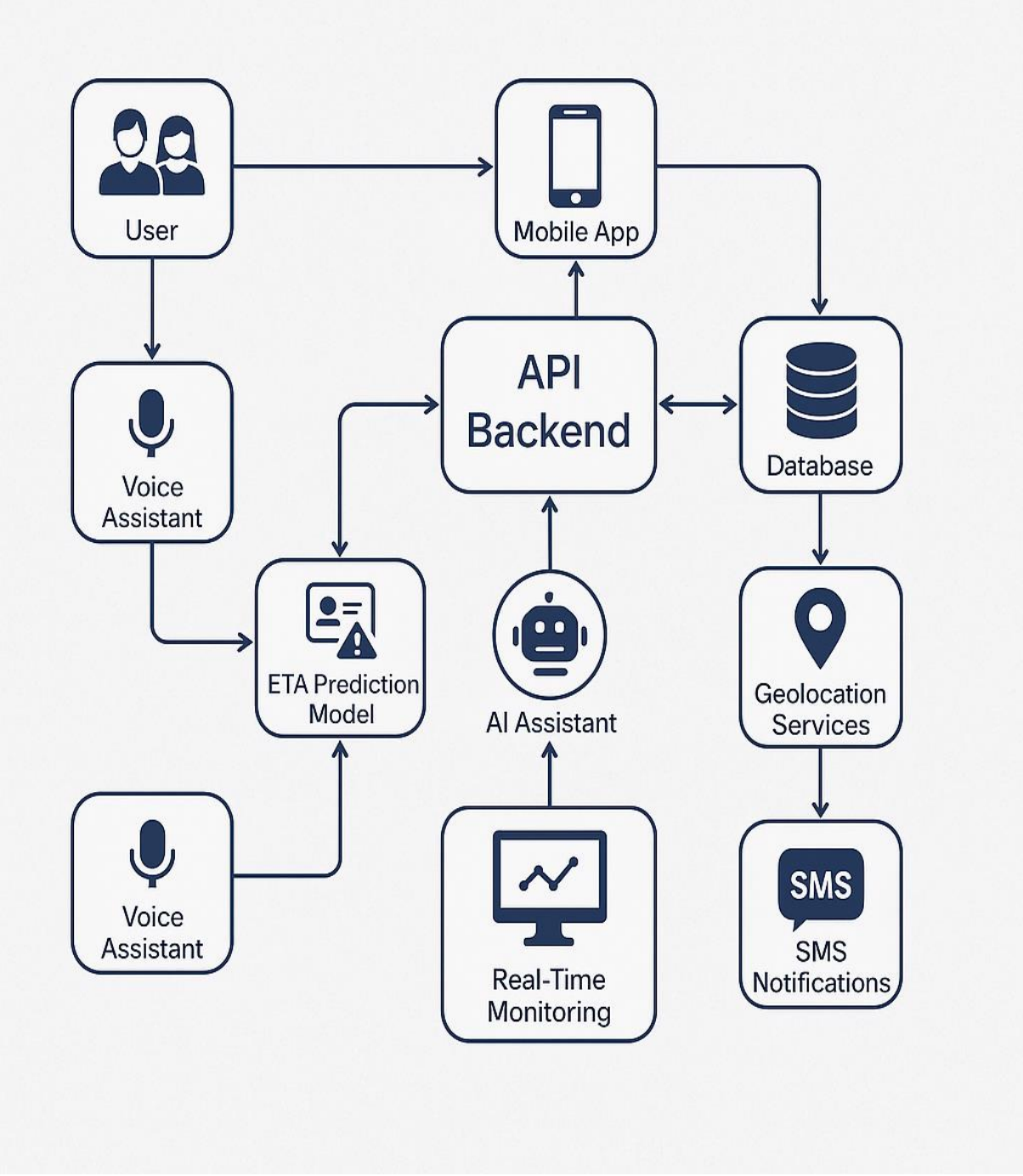
NLP Pipeline Sentiment & Document Verification

This component includes:

- **Sentiment Analysis:** Extracts keywords and polarity from user reviews, legal petitions, and tweets using a fine-tuned NLP model
- **Document Classification:** Uses OCR + text classification to detect suspicious or tampered documents uploaded by drivers

It plays a dual role in improving **public trust** and enforcing **driver identity verification**.

8 Architecture Flow Diagram



9. Security and Compliance

Security and compliance are at the heart of the Rapido 2.0 rebuild. In a digital service that connects thousands of users and drivers daily, protecting personal data and ensuring legal operations is not just a technical requirement it's a foundation for trust. Recognizing the impact of the 2024 data breach and ongoing court scrutiny, the new system was designed with **security-first principles** and **strict policy enforcement mechanisms** baked into its architecture.

All personal and ride-related data on the platform—such as user names, contact details, pickup/drop locations, ride history, and payment information—is encrypted using **252-bit encryption protocols**, going beyond the current industry standards. This high-level encryption ensures that even if data were intercepted, it would remain unreadable and secure from tampering. Additionally, communication between riders and drivers is routed through a temporary virtual number, ensuring that **actual contact details are never revealed** to either party. This protects users, especially women and vulnerable riders, from potential harassment or post-ride contact.

From the compliance perspective, every driver who registers on the Rapido 2.0 platform must pass a **multi-step verification process**, which includes identity proof, driver's license, and vehicle registration certificate (RC) uploads. These documents are processed through a combination of Optical Character Recognition (OCR), face-matching, and backend APIs to validate authenticity and consistency. If any mismatch is detected for example, if a driver attempts to accept rides using a vehicle different from the one registered in their profile the system **automatically denies the ride**, flags the account, and may impose a penalty or temporary suspension based on severity and frequency of violation.

The same logic is extended to geo-compliance: drivers attempting to operate in banned or restricted zones are blocked from receiving ride requests in those areas, enforced by **geo-fencing rules** tied to local transport regulations. These rules are continuously updated based on changes in city-specific legal frameworks.

In summary, Rapido 2.0 doesn’t treat compliance and security as add-ons it integrates them into the core of its daily operations. With stronger encryption, stricter verification, and real-time enforcement protocols, the platform is now fully equipped to protect user privacy, support government mandates, and build a system that passengers and authorities alike can rely on.

10. Tech Stack

The following technologies were used to build and deploy the Rapido 2.0 solution, categorized based on their role in the system:

Programming Language

- **Python 3.9** – Core language for machine learning models, backend logic, and automation scripts

Machine Learning & AI Libraries

- **XGBoost** – Used for ride cancellation prediction
- **scikit-learn** – Model building, evaluation, and preprocessing
- **NLTK** – Natural Language Processing for sentiment analysis
- **LangChain** – Integration of LLMs for customer support assistant and document handling

Backend & API Development

- **FastAPI** – Lightweight, high-performance framework to serve APIs and ML models
- **PostgreSQL 13** – Relational database for storing structured records such as ride logs, compliance data, and city legal info

Cloud & Communication Services

- **Firebase** – Real-time database for live ride tracking and push notifications
- **Twilio** – SMS alerts, voice calls, emergency contact routing
- **GCP Maps API** – Live location, ETA, route data
- **GCP TTS (Text-to-Speech)** – Voice assistant responses in regional languages

Dashboard & Visualization

- **Streamlit** – Interactive UI for admin dashboards and AI monitoring
- **Matplotlib** – Data plotting and ride trend visualization

- **Seaborn** – Advanced visualizations like heatmaps and correlation plots
- **Plotly Express** – Used for dynamic charts, ride patterns, and cluster-based visuals
- **Plotly Graph Objects (GO)** – Advanced, interactive graphs (e.g., heatmaps, multi-layer dashboards)

11. Deployment Strategy

The deployment of Rapido 2.0 was planned to ensure high availability, scalability, and secure model inference across real-time user interactions. The architecture integrates machine learning inference, live dashboards, cloud deployment, and automated CI/CD workflows.

Machine Learning Model Deployment

The core machine learning models such as the ride cancellation predictor and compliance detection engine are packaged and served using **FastAPI**. This lightweight, asynchronous Python framework allows RESTful API endpoints to be accessed by both the mobile application and admin dashboard in real time. The APIs are optimized for low-latency responses and can scale horizontally using cloud services.

Dashboard & Monitoring System

The **Streamlit** interface is used to deploy interactive dashboards for real-time analytics, legal risk mapping, and operational monitoring. The dashboard is connected to Firebase and PostgreSQL, allowing internal teams to visualize live data including high-risk zones, ride cancellations, and flagged accounts.

Notification and Communication Layer

For real-time communication with users and drivers, the platform uses:

- **Firebase Cloud Messaging** for push notifications (e.g., ride updates, alerts)
- **Twilio** for sending OTPs, emergency messages, and voice calls

These services ensure fast, reliable, and multilingual communication across devices.

Cloud Deployment & Automation

The entire solution is deployed on **Google Cloud Run**, which provides serverless, container-based hosting with automatic scaling. This setup reduces infrastructure management overhead while maintaining high performance.

Continuous integration and deployment (CI/CD) are handled via **GitHub Actions**, ensuring that every code change is automatically tested, built, and deployed. This creates a fast feedback loop for development, testing, and production deployment.

12. Business Impact

The implementation of Rapido 2.0 delivers significant business value by addressing the platform's core weaknesses and unlocking new opportunities for growth, safety, and market credibility. Through a combination of AI-driven automation, compliance enforcement, and user-centric design, the platform is now positioned to function not only as a ride service but as a responsible urban mobility partner.

One of the most critical outcomes is the **improvement in legal compliance**, which reduces the risk of city-wide bans and shutdowns. By incorporating geo-fencing, document verification, and city-level legal dashboards, Rapido can proactively align with local transport regulations and court orders. This transforms legal issues from a reactive burden into a manageable, data-driven process.

The platform also sees a boost in **user inclusivity**, particularly for persons with disabilities (PwDs) and multilingual users. With the rollout of an accessible UI and voice assistant in multiple Indian languages, the service becomes usable by a broader demographic, increasing its reach and social acceptance. This not only strengthens brand equity but also supports Rapido's eligibility for government partnerships and smart city projects.

In terms of safety and reliability, the integration of real-time **fraud detection**, document validation, and the SOS system builds **public trust** a key factor in increasing ride frequency and user retention. Female riders, senior citizens, and first-time users are more likely to choose a service that puts visible safety mechanisms in place.

Operationally, **predictive maintenance and cancellation prediction models** help reduce costs related to missed rides, unnecessary driver deployments, and backend support overload. The system can anticipate demand patterns and driver behaviour, resulting in better resource management.

Finally, the inclusion of a **Large Language Model (LLM)-based support assistant** brings significant efficiency to customer service operations. The assistant can handle thousands of queries

simultaneously, reducing support wait times and improving the user experience without the need for a large support staff.

13. Future Scope

While Rapido 2.0 addresses many of the current operational, legal, and safety challenges, the platform holds significant potential for further growth and innovation in the mobility ecosystem. As urban transportation in India evolves, so must the technology and services behind it. The following future enhancements are proposed to expand the impact of the solution and make Rapido a leader in smart, inclusive, and sustainable ride-hailing.

One of the most promising directions is the **expansion into inter-city ride-sharing** and the integration of **Electric Vehicles (EVs)**. As state policies increasingly promote green mobility, Rapido can lead the transition by offering low-emission, cost-effective transportation options for both short and long-distance commutes. This would not only reduce environmental impact but also open new revenue streams and government subsidy opportunities.

To strengthen its legal foundation, Rapido can **form direct partnerships with state transport departments**. By collaborating on permit validation, zone-based permissions, and compliance APIs, the platform can operate more transparently and avoid sudden legal interruptions. Such collaboration would also position Rapido as a model operator in public-private mobility partnerships.

For identity and document verification, the system can be further upgraded by **integrating with UPI and DigiLocker**. This would enable seamless **e-KYC (electronic Know Your Customer)** flows for both riders and drivers, allowing for real-time, government-verified onboarding. This move would improve platform security while simplifying the user experience.

Finally, on the safety front, the use of **AI-powered cameras** for **helmet detection** and **accident risk prediction** can be implemented on the driver side. These cameras can monitor live footage to ensure riders wear helmets and that drivers follow speed regulations. Any signs of reckless behaviour, sudden motion changes, or high-speed

collisions can trigger instant alerts to the admin team or emergency services.

14. Conclusion

Rapido 2.0 is more than just a technical upgrade it represents a complete reimagination of what urban ride-hailing must become in today's fast-evolving Indian landscape. By combining machine learning, accessibility design, legal awareness, and real-time intelligence, the solution transforms a reactive platform into a proactive, secure, and people-centric mobility service.

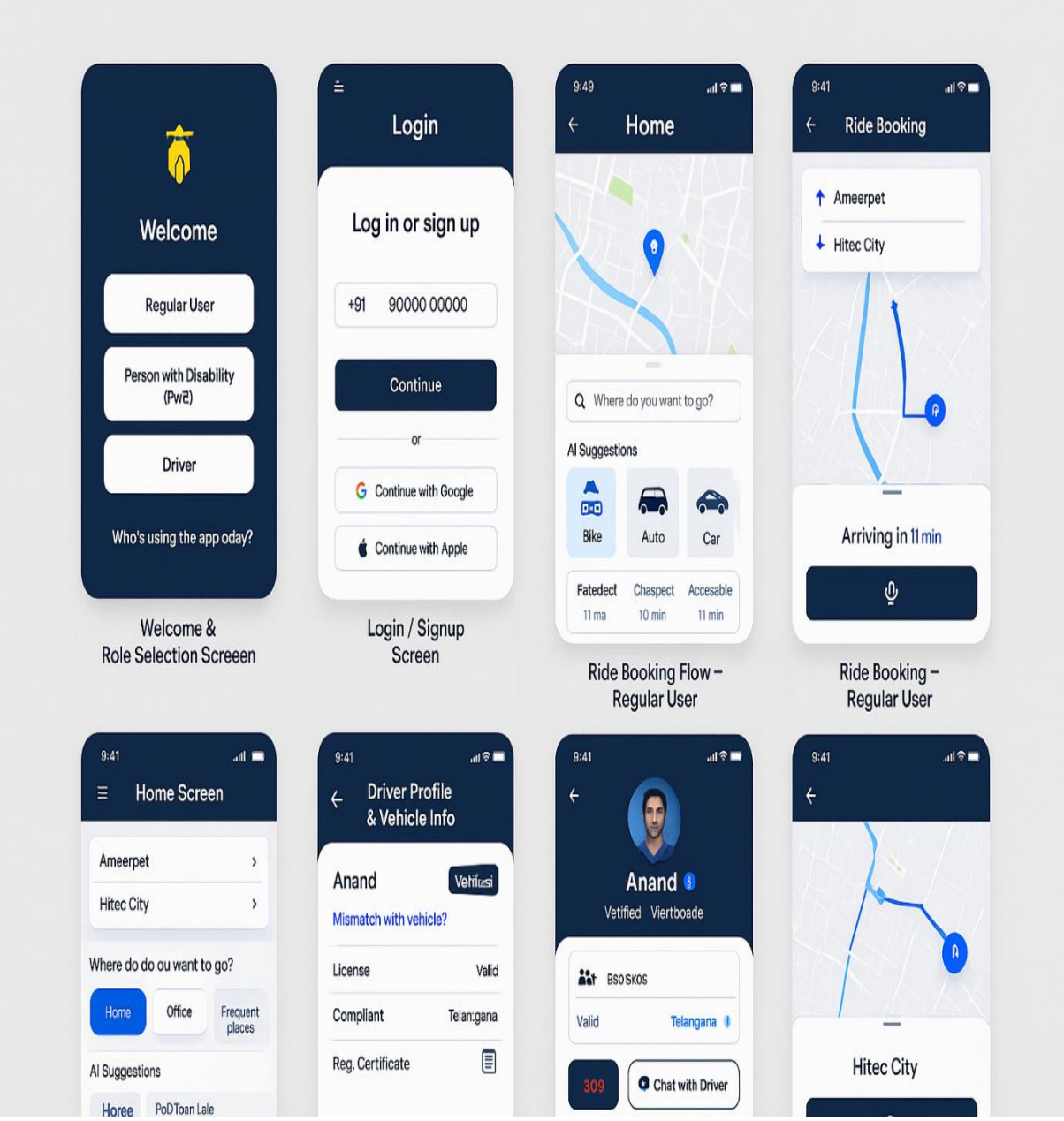
This project demonstrates that when data-driven decision-making meets inclusive thinking, the result is a platform that doesn't just respond to challenges but anticipates them. With predictive models for cancellations, AI-powered safety tools, and compliance monitoring systems in place, Rapido 2.0 offers both operational efficiency and public trust.

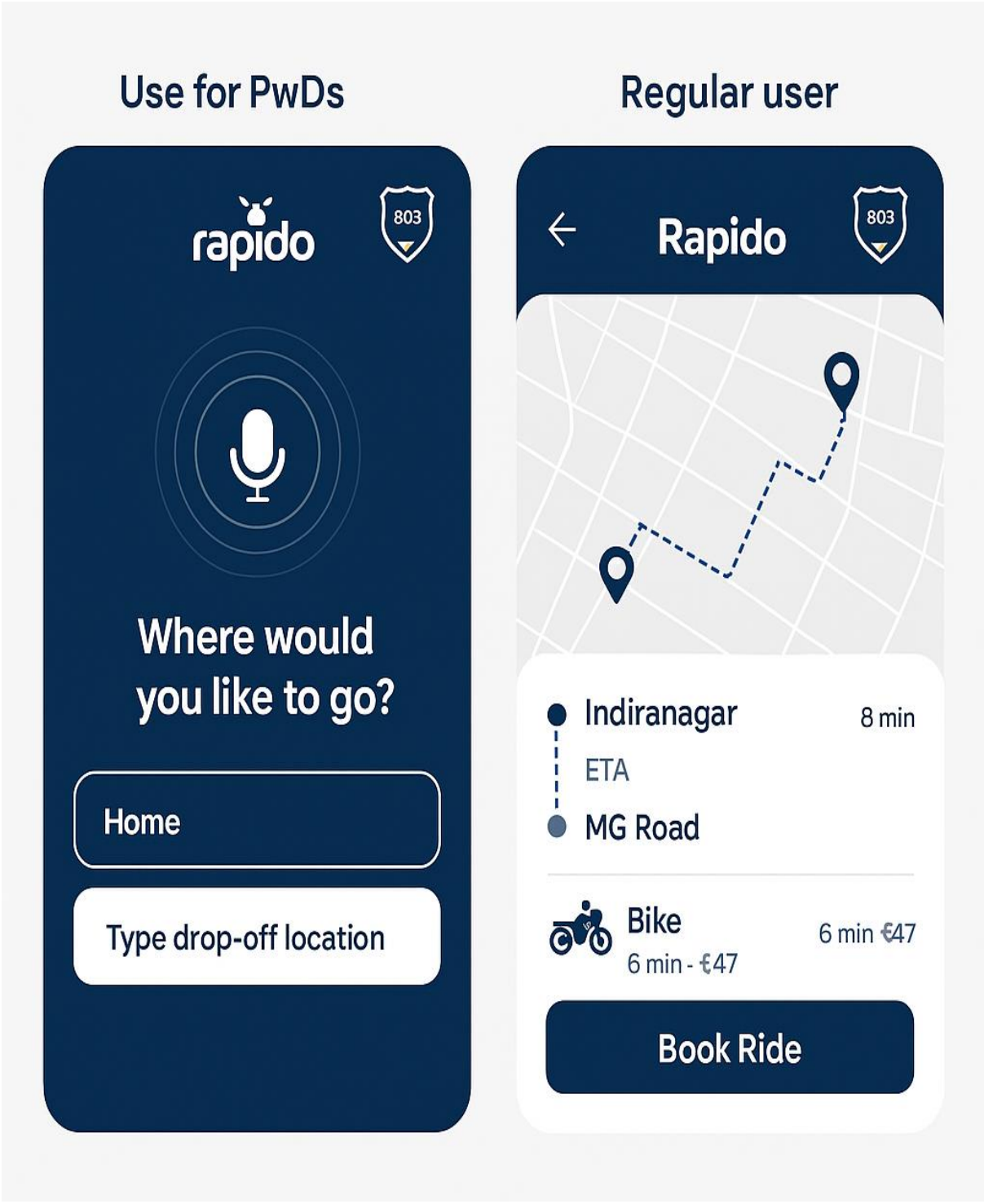
What makes this solution even more impactful is its **scalability**. While built around Rapido's case, the system architecture, models, and features can easily be adapted by other major players such as **Ola**, **Uber**, or any regional ride-hailing service aiming to comply with state transport laws, support diverse users, and grow responsibly.

In essence, Rapido 2.0 isn't just a comeback strategy it's a **blueprint for the future** of last-mile mobility in India: smart, safe, inclusive, and compliant by design.

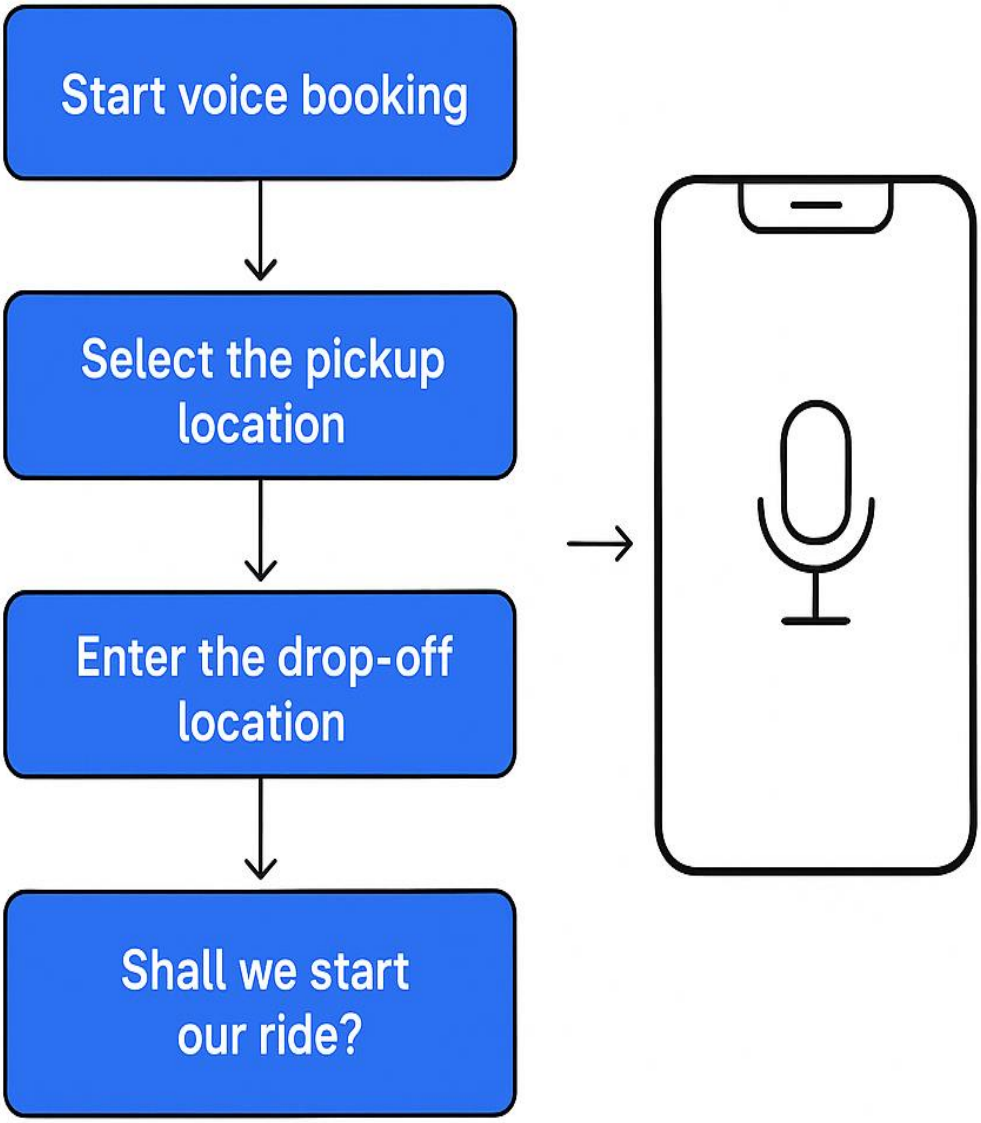
15. Appendix

15.A Sample UI Layouts

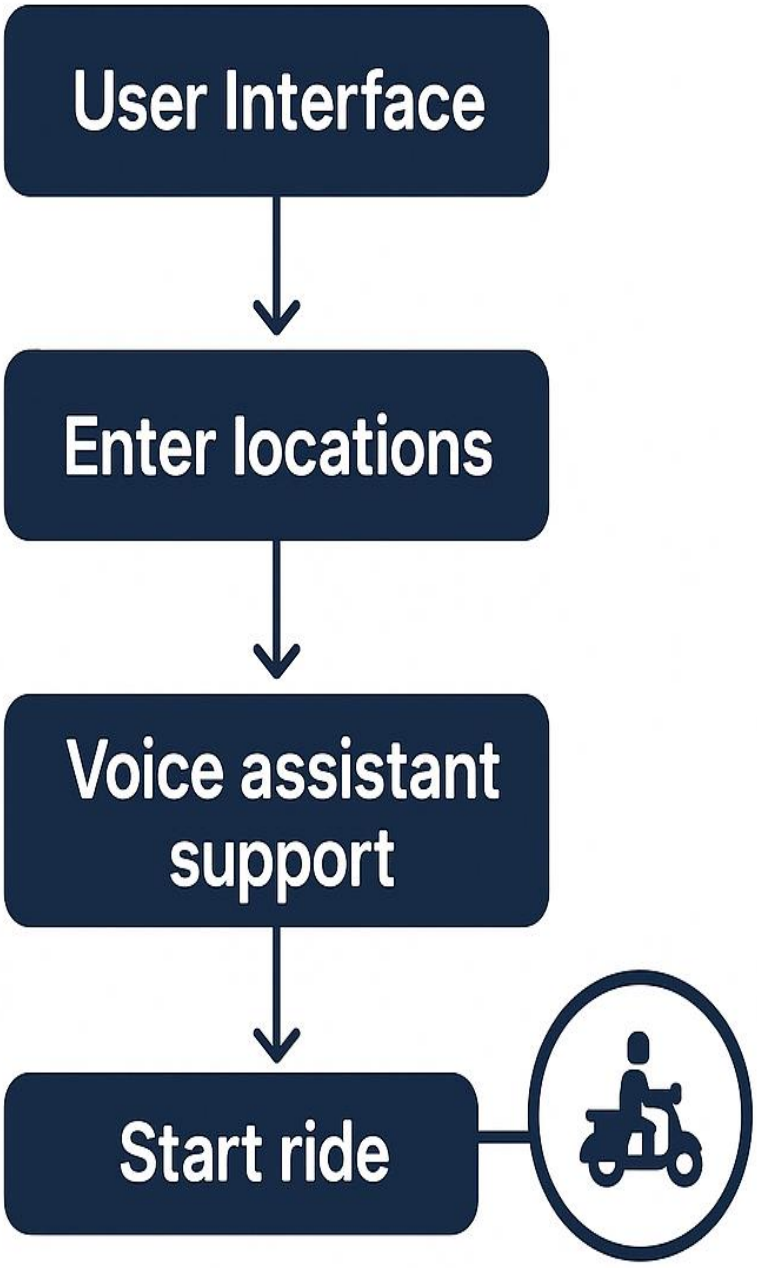


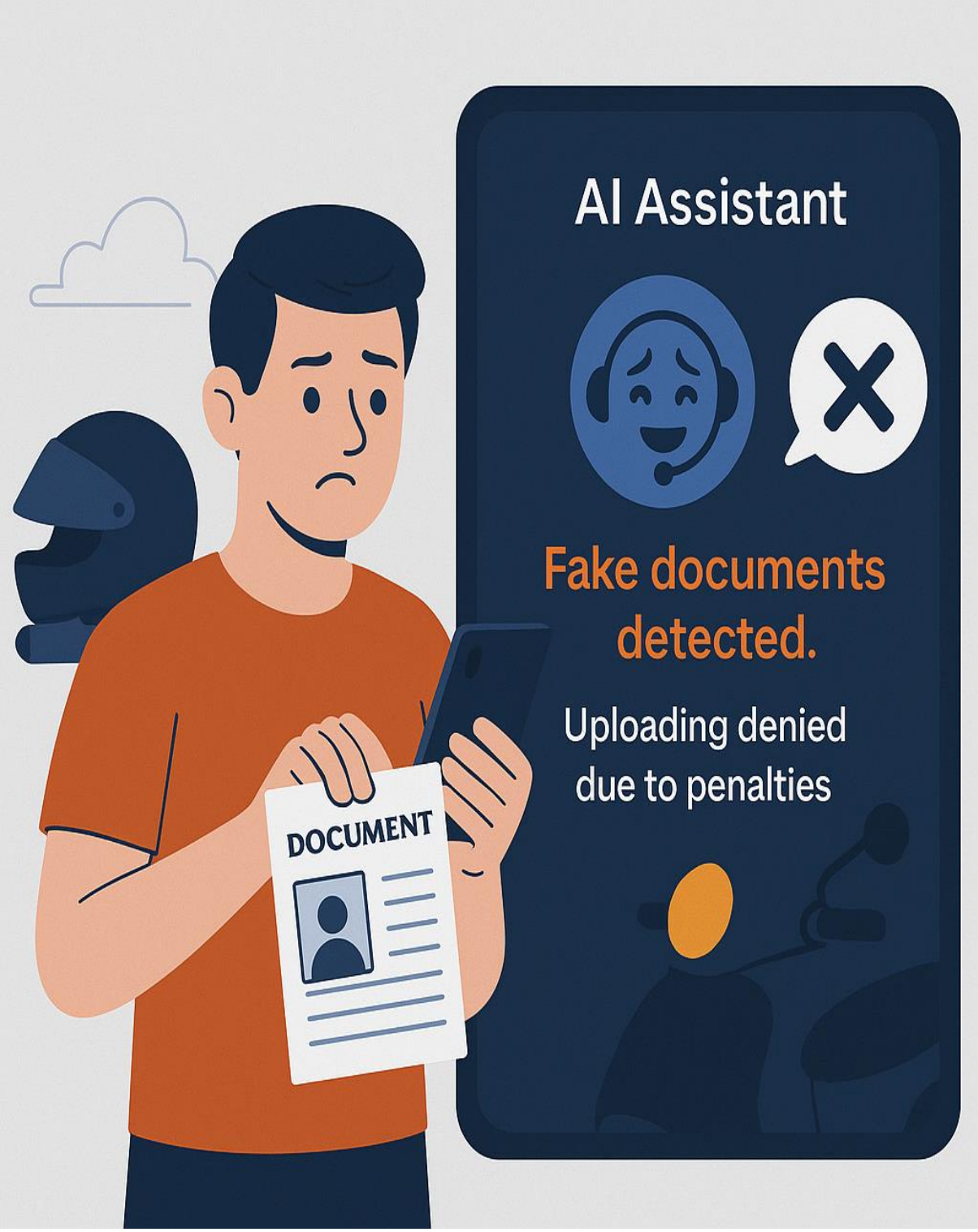

Use for PwDs
rapido
803
Where would you like to go?
Home
Type drop-off location
Regular user
Rapido
803
Indiranagar
8 min
ETA
MG Road
Bike
6 min - ₹47
6 min ₹47
Book Ride

Voice Booking Flow



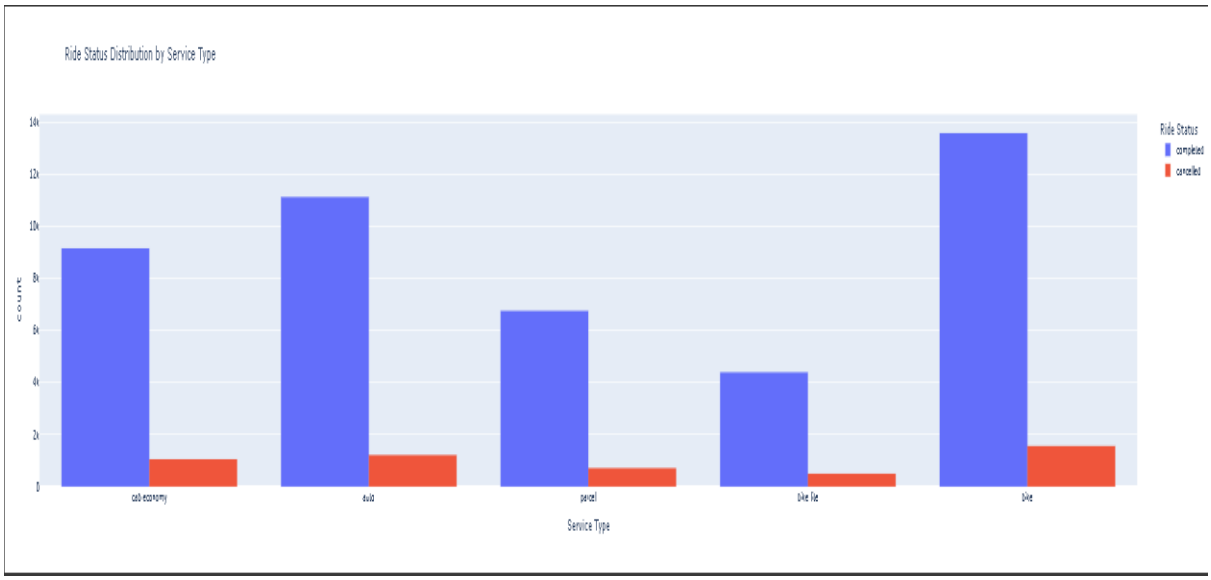
User Interaction Flow

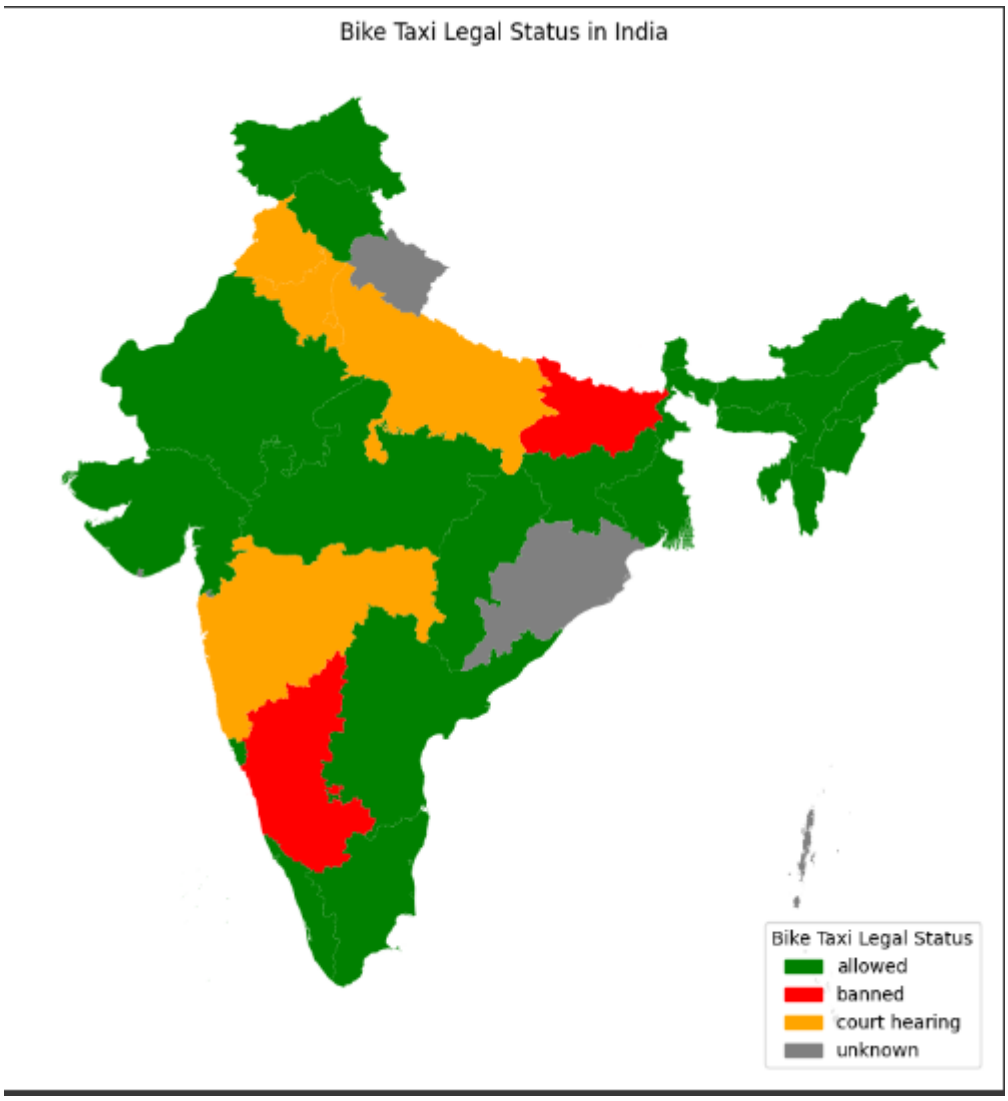


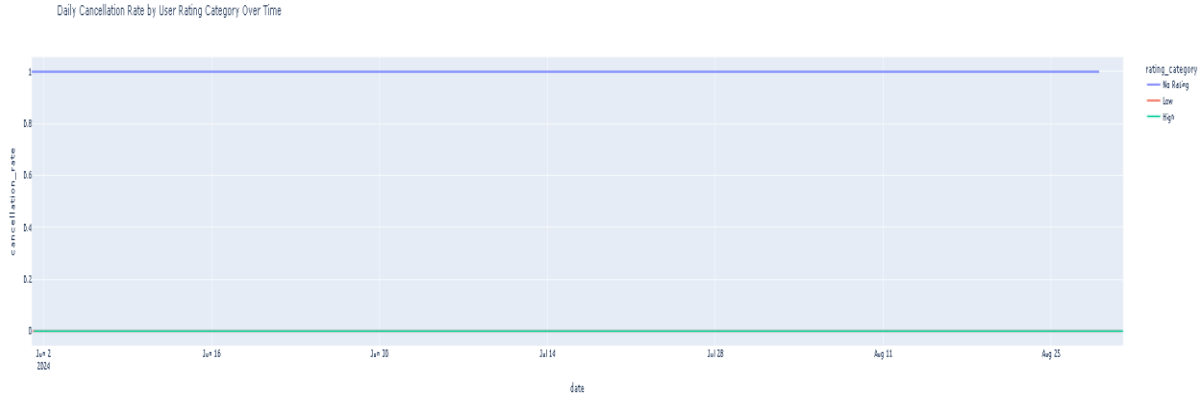

AI Assistant
Fake documents detected.
Uploading denied due to penalties
DOCUMENT

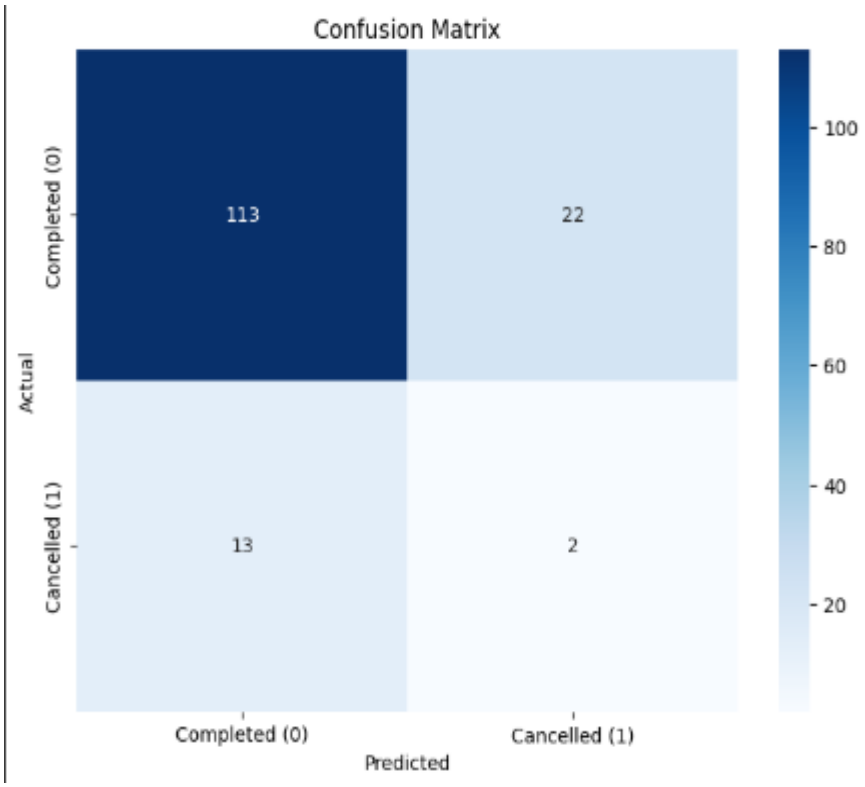

Safety &
Fraud Detection
SOS
!
Calling nearby police

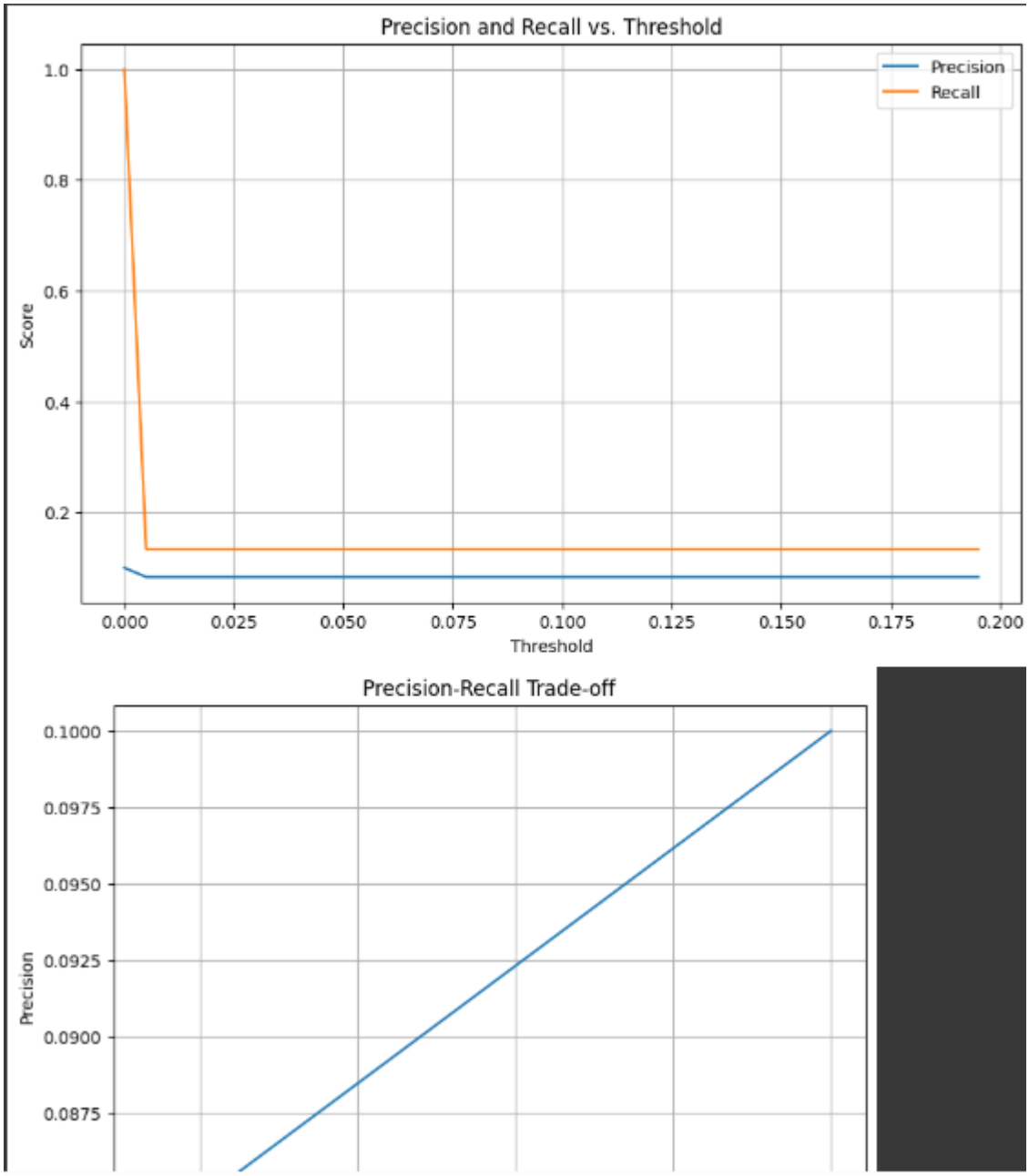
15.B Insights

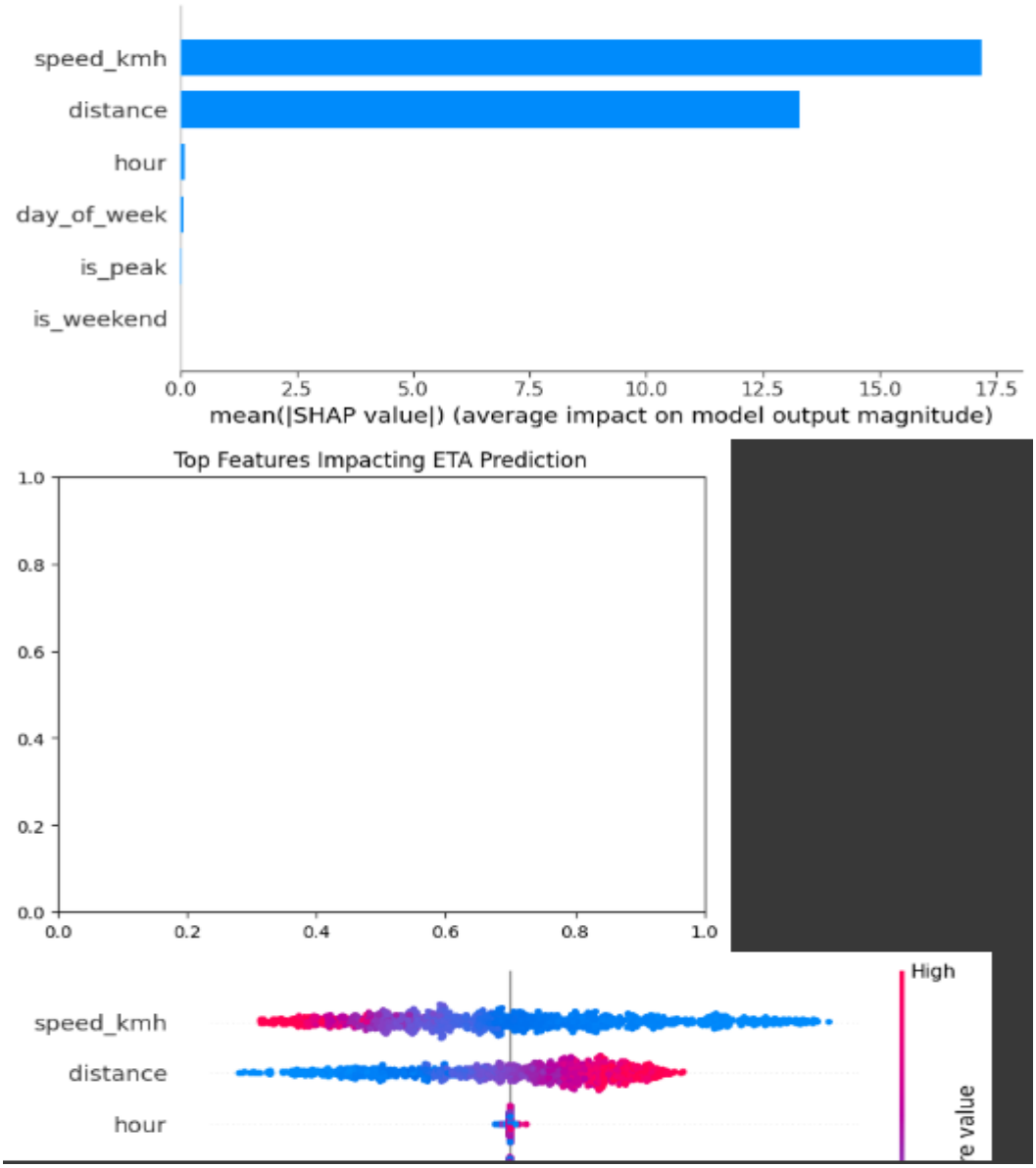


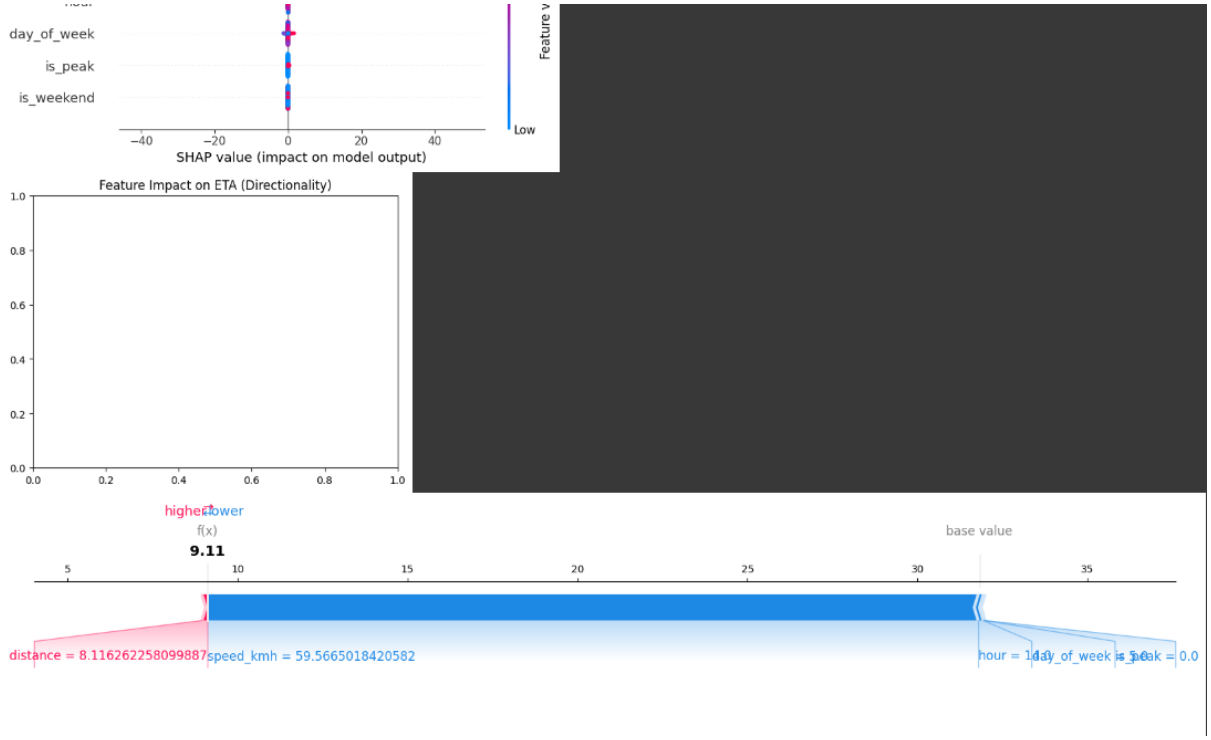


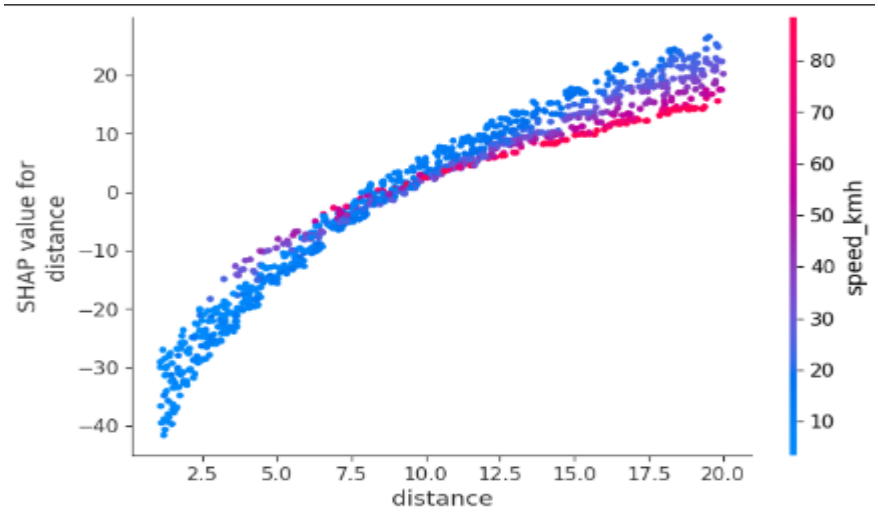

Daily Cancellation Rate by User Rating Category Over Time
cancellation_rate
date
rating_category
No Rating
Low
High

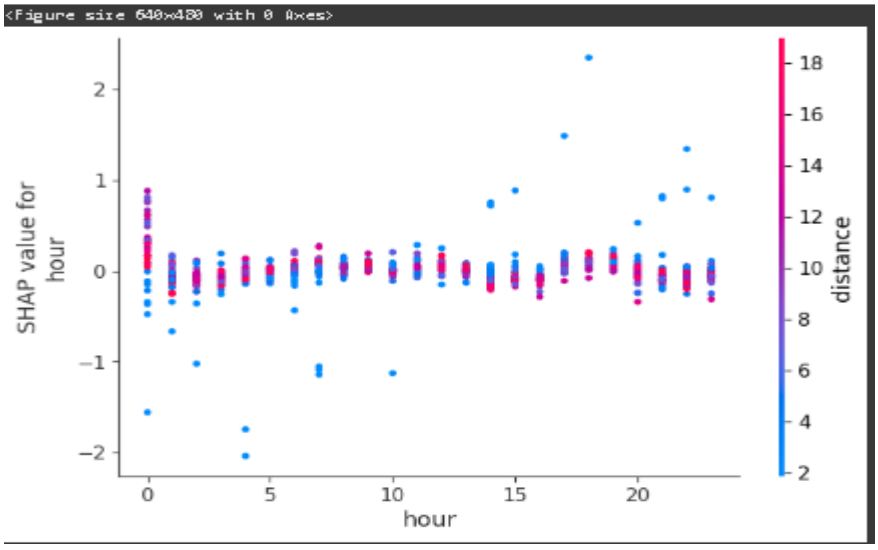

Confusion Matrix
Actual
Completed (0)
Cancelled (1)
113
22
13
2
Predicted
Completed (0)
Cancelled (1)
100
80
60
40
20

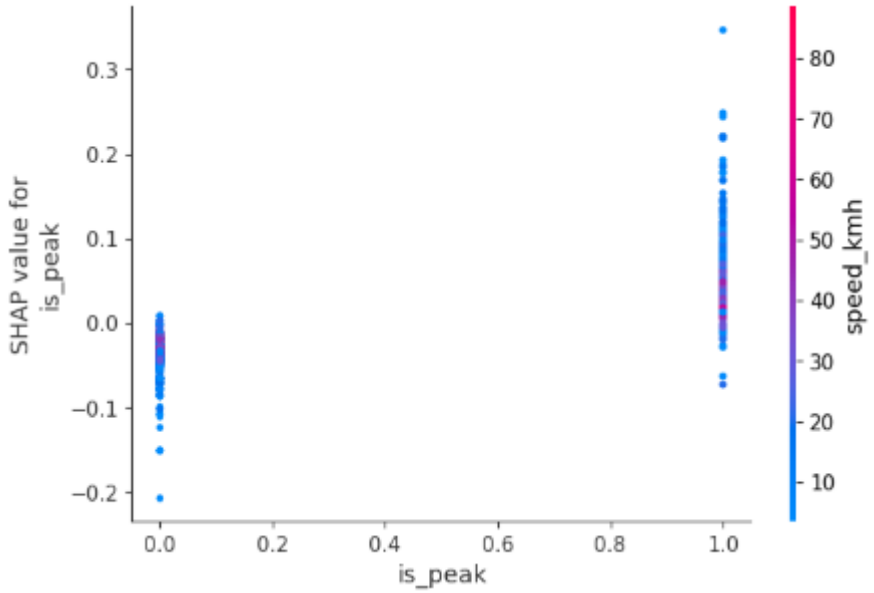

Precision and Recall vs. Threshold
Precision
Recall
Score
Threshold
1.0
0.8
0.6
0.4
0.2
0.000
0.025
0.050
0.075
0.100
0.125
0.150
0.175
0.200
Precision-Recall Trade-off
Precision
0.1000
0.0975
0.0950
0.0925
0.0900
0.0875

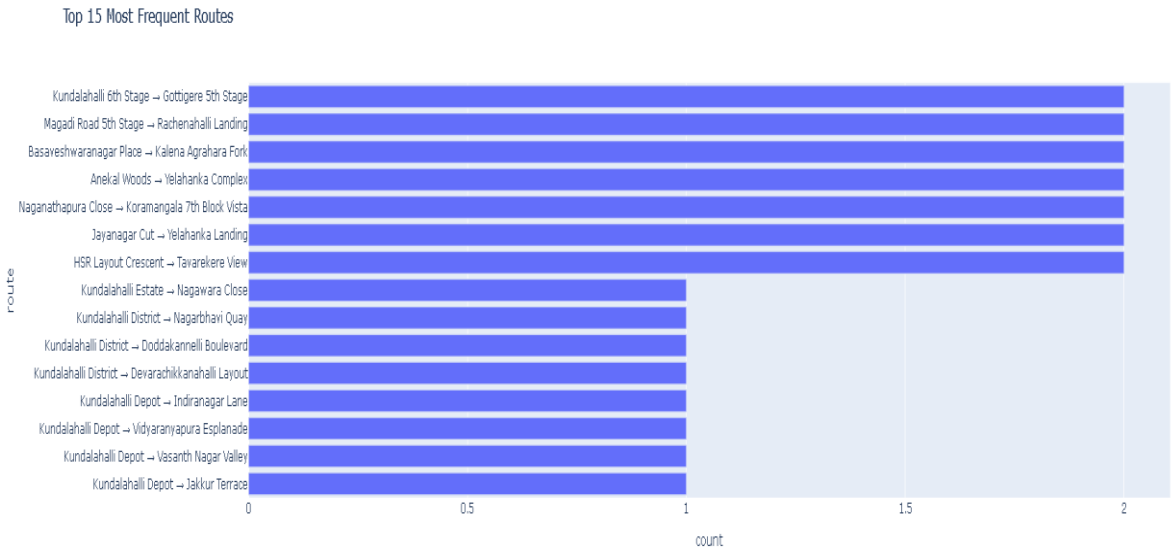

speed_kmh
distance
hour
day_of_week
is_peak
is_weekend
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
mean(|SHAP value|) (average impact on model output magnitude)
Top Features Impacting ETA Prediction
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
speed_kmh
distance
hour
High
e value

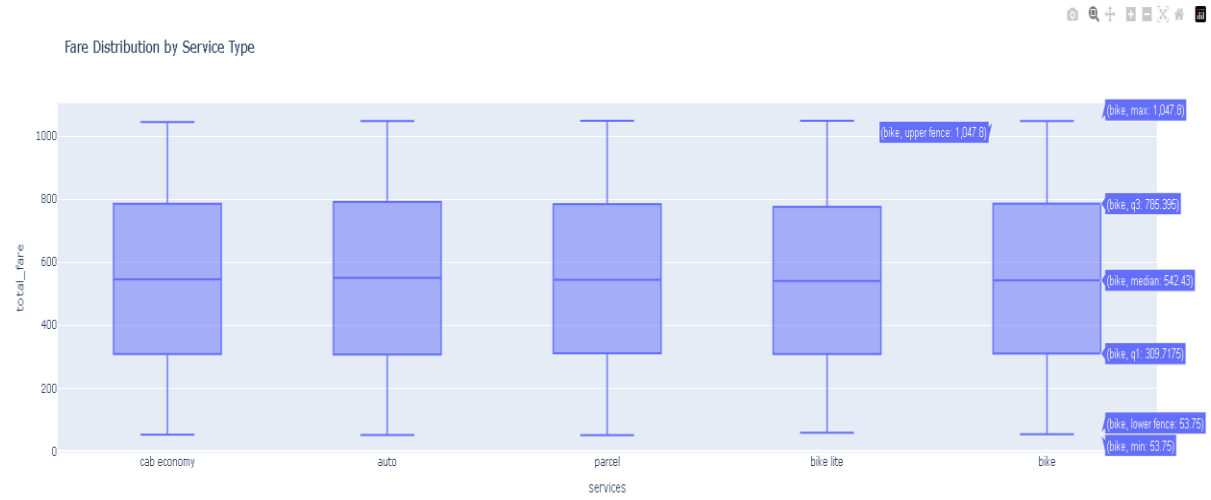

day_of_week
is_peak
is_weekend
−40
−20
0
20
40
SHAP value (impact on model output)
Feature v
Low
Feature Impact on ETA (Directionality)
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
higher
lower
f(x)
9.11
base value
5
10
15
20
25
30
35
distance = 8.116262258099887
speed_kmh = 59.5665018420582
hour = 1.0
day_of_week = 5.0
is_peak = 0.0

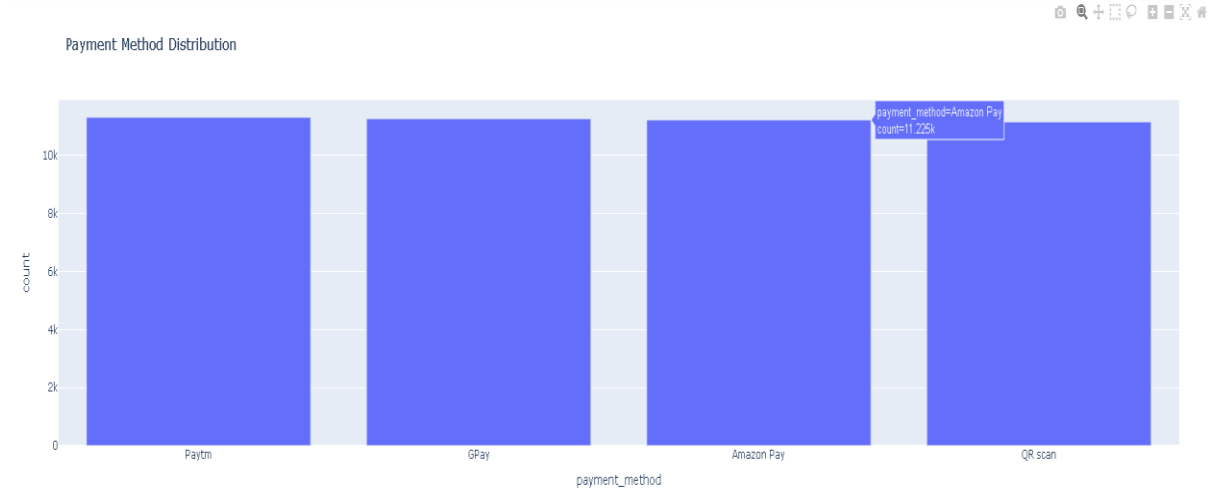

SHAP value for
distance
20
10
0
−10
−20
−30
−40
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
distance
speed_kmh
80
70
60
50
40
30
20
10

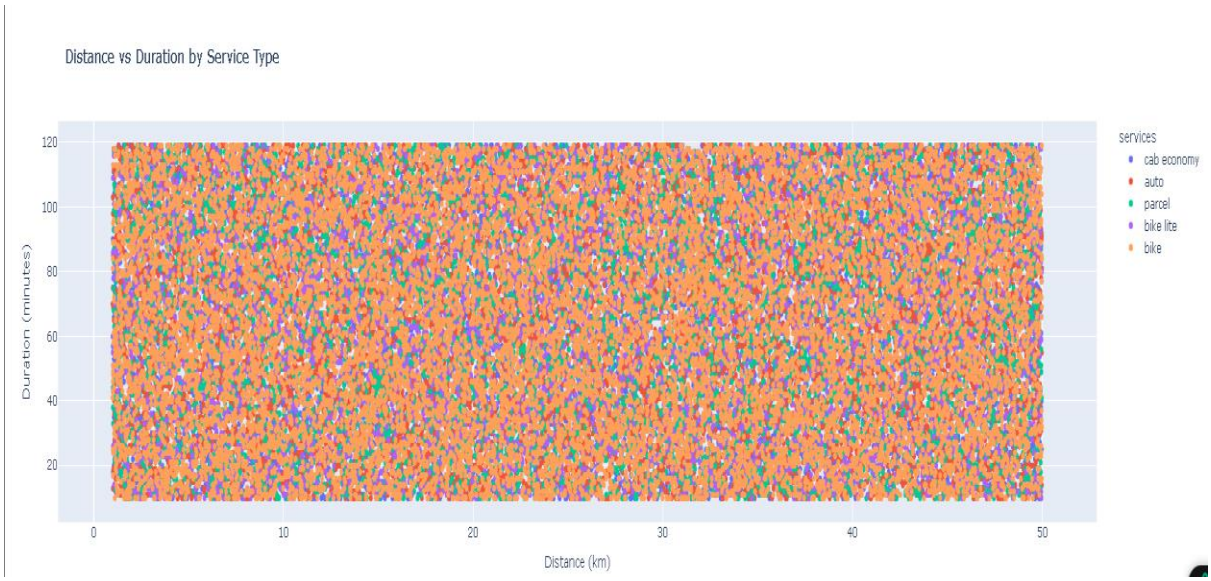

<Figure size 640x480 with 0 Axes>
SHAP value for
hour
2
1
0
−1
−2
0
5
10
15
20
hour
distance
18
16
14
12
10
8
6
4
2


SHAP value for is_peak
0.3
0.2
0.1
0.0
−0.1
−0.2
0.0
0.2
0.4
0.6
0.8
1.0
is_peak
80
70
60
50
40
30
20
10
speed_kmh


Top 15 Most Frequent Routes
Kundalahalli 6th Stage → Gottigere 5th Stage
Magadi Road 5th Stage → Rachenahalli Landing
Basaveshwaranagar Place → Kalena Agrahara Fork
Anekal Woods → Yelahanka Complex
Naganathapura Close → Koramangala 7th Block Vista
Jayanagar Cut → Yelahanka Landing
HSR Layout Crescent → Tavarekere View
Kundalahalli Estate → Nagawara Close
Kundalahalli District → Nagarbhavi Quay
Kundalahalli District → Doddakannelli Boulevard
Kundalahalli District → Devarachikkanahalli Layout
Kundalahalli Depot → Indiranagar Lane
Kundalahalli Depot → Vidyaranyapura Esplanade
Kundalahalli Depot → Vasanth Nagar Valley
Kundalahalli Depot → Jakkur Terrace
route
0
0.5
1
1.5
2
count


Fare Distribution by Service Type
total_fare
1000
800
600
400
200
0
cab economy
auto
parcel
bike lite
bike
services
(bike, max: 1,047.8)
(bike, upper fence: 1,047.8)
(bike, q3: 765.395)
(bike, median: 542.43)
(bike, q1: 309.7175)
(bike, lower fence: 53.75)
(bike, min: 53.75)


Payment Method Distribution
count
10k
8k
6k
4k
2k
0
Paytm
GPay
Amazon Pay
QR scan
payment_method
payment_method=Amazon Pay
count=11.225k


Distance vs Duration by Service Type
Duration (minutes)
120
100
80
60
40
20
0
10
20
30
40
50
Distance (km)
services
cab economy
auto
parcel
bike lite
bike

16. Resources & References

- The Indian Express – Rapido suspends bike taxi services in Karnataka (June 2025)
- Deccan Herald – Bike taxis are necessity, not luxury: Aggregators argue in Karnataka HC
- Deccan Herald – A flawed ban on bike taxis (Opinion)
- NDTV – Improve accessibility for disabled or pack up: High Court to Rapido
- Times of India – Researcher claims Rapido exposed customer data
- Social Samosa – Rapido exposes user and driver data via feedback form
- TechGig – Security flaw exposes personal data on Rapido
- Karnataka HC Order: Livemint, Economic Times, LawChakra (June 2025)
- Delhi HC Accessibility: LawChakra, India Legal (March 2025)
- Data Breach Report: CyberSRC, ABP Live (Dec 2024)
- Public Sentiment: Twitter/X, IndiaCustomerCare.com scraped reviews
- WCAG Guidelines: https://www.w3.org/WAI/WCAG21/quickref/

17. Final Word

"Rapido 2.0 isn't just a reboot it's a comeback."

Let's make India's bike taxi safe, inclusive, and intelligent.

Contact Information

Project Lead: Deepak Patro

Role- Data consultant, AI Solution Architect & Data Scientist
Email: deepakpatro73@email.com
LinkedIn: https://www.linkedin.com/in/deepak-patro-2a0330228/
GitHub: https://github.com/deepak8114
Location: Hyderabad, India