

PROJECT TYPE

Logistics platform

TECHNOLOGIES

React, React Native, PostgreSQL, Redis, AWS, Docker, Kubernetes, Apache Kafka, Apache Airflow, Nginx

DURATION

20 weeks

METHODOLOGY

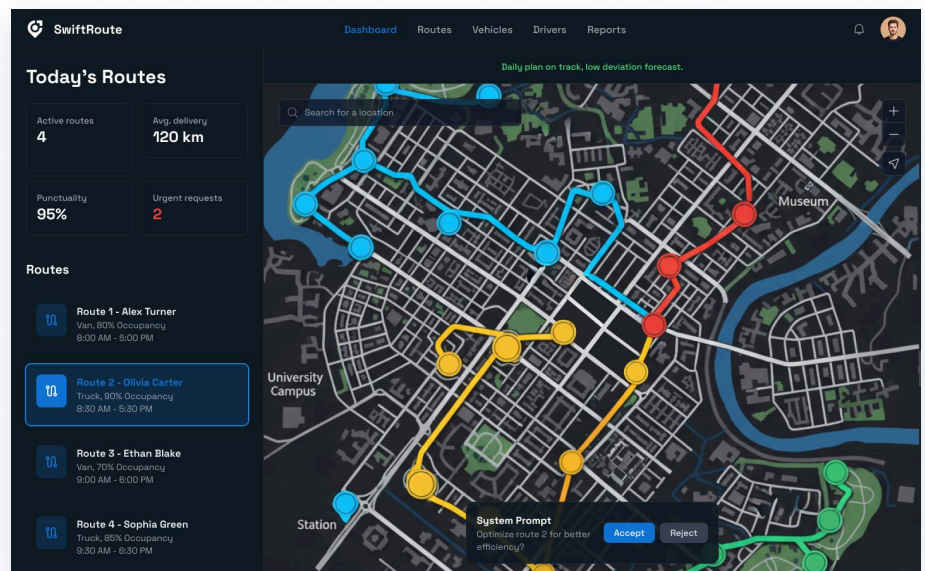
Scrum, Kanban

TEAM

1 Project Manager
1 Business Analyst
1 Solution Architect
2 Backend Developers
1 UX/UI Developer
1 Data and Integration Engineer
1 AI engineer
1 QA Engineer
1 Infrastructure and Automation Engineer

AI/ML route and last-mile optimization for freight and parcel delivery services

An AI/ML-powered logistics platform to plan, optimize, and track B2B and B2C deliveries across a multi-branch network. Features include predictive routing with traffic and weather, capacity-aware vehicle assignment, real-time ETAs and re-optimization, cross-branch load balancing, and customer/driver portals with live notifications.

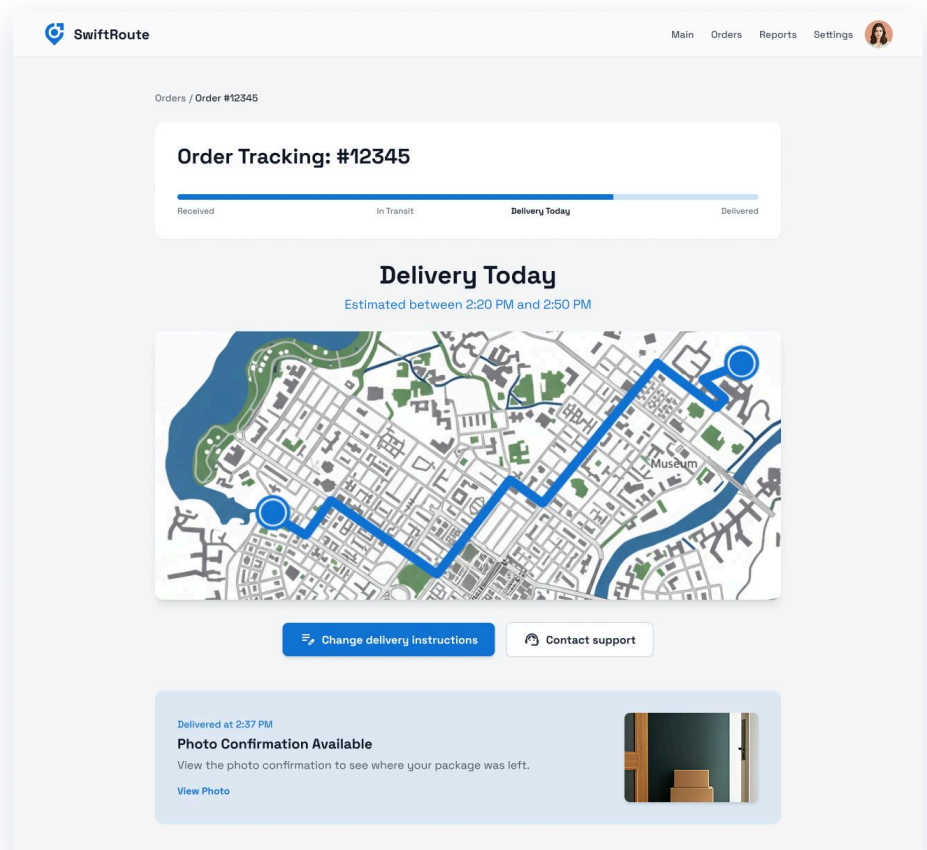


The Client is a freight and parcel delivery service operating in both B2B and B2C formats. The company accepts orders from online stores and corporate shippers (including international deliveries for their end customers), as well as point-to-point orders from individuals. The main challenges were high last-mile costs, inefficient routes, uneven order flows, and repeated trips to the same areas by different drivers in vehicles of inadequate capacity. The Client sought a solution in the form of a unified route and last-mile optimization platform that would serve all parties involved in the delivery and receipt of parcels.

Project Distinctive Features

- ✓ Unified B2B + B2C order intake with data normalization, service-level tiers, and priority rules in one system.
- ✓ AI/ML-driven route planning that accounts for time windows, live traffic, weather, road works, and access constraints, with mid-day re-optimization.

- ✓ Capacity-aware vehicle assignment that right-sizes vans and trucks to volume and weight, reducing “empty air” and missed fit.
- ✓ Hybrid flow orchestration for scheduled corporate pickups and on-demand consumer shipments, including dynamic “micro-slots” during peaks.
- ✓ Cross-depot load balancing that redistributes stops between branches to cut repeat visits to the same area.
- ✓ Real-time ETA with confidence bands for dispatchers, drivers, and customers, plus proactive delay alerts.
- ✓ Driver app with offline mode: turn-by-turn guidance, special access instructions, photo/signature proof of delivery, and failed-attempt workflows.
- ✓ Control-tower suggestions for dispatchers (sequence swaps, consolidations) with accept/override and full audit trail.
- ✓ SLA-aware prioritization that protects corporate delivery windows while absorbing urgent consumer requests.
- ✓ Exception and returns handling (re-routing after failed delivery, consolidation of returns, customer instruction changes).
- ✓ AI-tuned notifications (channel and send-time) to lift response rates and reduce missed handoffs.
- ✓ Geospatial analytics (distance per stop, repeat-visit heatmaps, utilization by area and hour) tied to cost and punctuality KPIs.
- ✓ Secure by design: encrypted data in transit and at rest, role-based access, detailed activity logs.
- ✓ Continuous learning: ML models retrain on new telemetry to improve routing quality and ETA accuracy over time.



Business challenge

To develop a platform that combines B2B and B2C orders and employs AI/ML to calculate routes that take into account time windows, traffic, and weather, select suitable vehicles, and recalculate ETAs in real-time. The solution should reduce total mileage and last-mile costs, increase delivery punctuality, and improve fleet utilization without requiring additional fleet expansion.

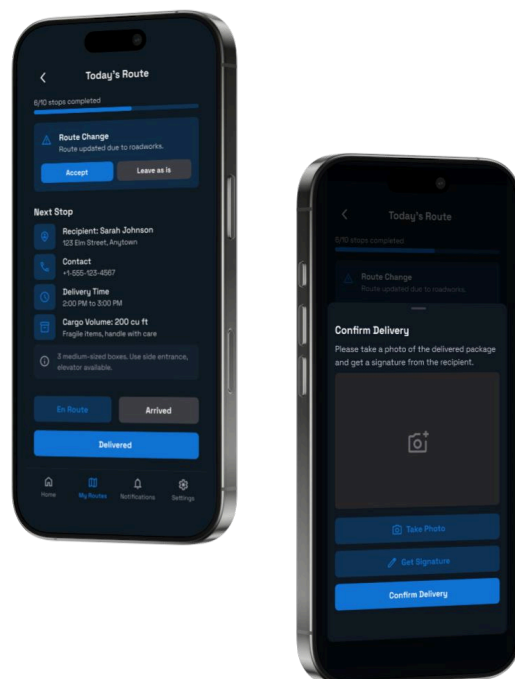
Our solution

Order centralization and data normalization

We have combined sources: corporate integrations of online stores, individual requests from the website and the receiving office, and international shipping tasks. The platform compiles all orders into a unified format, including mandatory attributes such as dimensions and weight, loading requirements, pickup and delivery times, the need for customs clearance at intermediate stages, and recipient contact information.

Intelligent route planning

Planning is AI-driven and based on city and regional maps, taking into account traffic volume, weather factors, and time constraints at pickup and drop-off points. The system groups nearby addresses, consolidates orders in one area, and creates a sequence of stops with estimated arrival times. It also assesses traffic congestion and the likelihood of delays. If changes occur during the day, the platform adjusts routes to minimize total mileage and ensure the agreed-upon deadlines are maintained.



Vehicle selection based on dimensions and restrictions

Operators input data in the system about the cargo volume and weight of each order. A machine learning model assesses the risk of “excess volume” and suggests the optimal vehicle class based on the actual dimensions and weight of the shipment. When creating routes, the system ensures that small shipments don’t get assigned to vehicles with excessive volume, and vice versa. Additionally, it takes into account accessibility features, height restrictions, and unloading times in both business centers and residential areas.

Balancing mixed flows: Corporate and private

Corporate orders have pre-agreed-upon windows and priorities; private orders, on the other hand, are more variable in terms of time and geography. The AI-optimized platform ensures consistent routes by consolidating scheduled pickups for companies into main routes. The closest sections include private addresses, and international shipments are automatically routed to consolidation points, taking into account customs clearance times. If an urgent request arises during the day, the system assigns it to the nearest driver with minimal deviation from the current plan.

Real-time support for dispatchers and drivers

The dispatcher sees the “day's health”: current routes, time deviations, AI-driven delay forecasts, and suggestions for changing the stop sequence. The driver receives the route in the mobile app, which includes a list of points, the optimal order, updated arrival time estimates, loading instructions, recipient contact information, and proof of delivery with a photo or signature. In the event of a traffic jam or street closure, the app suggests a safe detour and syncs with the dispatcher's plan.

Customer communication and delivery transparency

Recipients and senders receive notifications, including pickup confirmation, a tracking link, an updated arrival time estimate, and a warning about rescheduling in the event of unforeseen circumstances. The tracking page displays the route, current status, and estimated arrival time. Upon delivery, a confirmation is generated, including a timestamp and proof of delivery.

Analytics, quality, and security

Managers receive dashboards with key metrics: average delivery mileage, share of repeat trips to a given area, on-time compliance, actual vehicle utilization, and order distribution across areas and times of day. SumatoSoft ensured that data is secure by storing and transmitting it in an encrypted format, restricting access roles, and logging all actions. We also employed ML to improve planning quality gradually using historical route information.

Customer's benefits

Following the platform rollout, total delivery mileage fell by 12-18%, driving a 15-22% reduction in fuel consumption and last-mile costs. On-time performance within agreed windows rose to 96-98% as delays dropped 30-40%. Fleet utilization was rebalanced: vehicle use by volume and time increased 15-20% while downtime decreased 25-35%. Consolidation and smarter stop sequencing cut repeat trips to the same area by 35-45%, and response time to same-day urgent requests improved by 20-30% without compromising corporate delivery commitments.

What's happening with the project right now?

The platform is live across core B2B and B2C flows, with dispatch and driver apps used daily; AI/ML routing and re-optimization run hourly and, after eight weeks in production, KPIs show mileage –14%, last-mile cost –18%, on-time delivery 97%, vehicle utilization +17%, repeat-area visits –41%, and urgent-request response time –24%. On the roadmap: geofencing for high-friction locations, privacy-safe driver safety scoring, and quarterly cost-to-serve analytics by zone and time window.