# Predicting Shipping Transit Times

## Using Machine Learning & Gradient Boosting

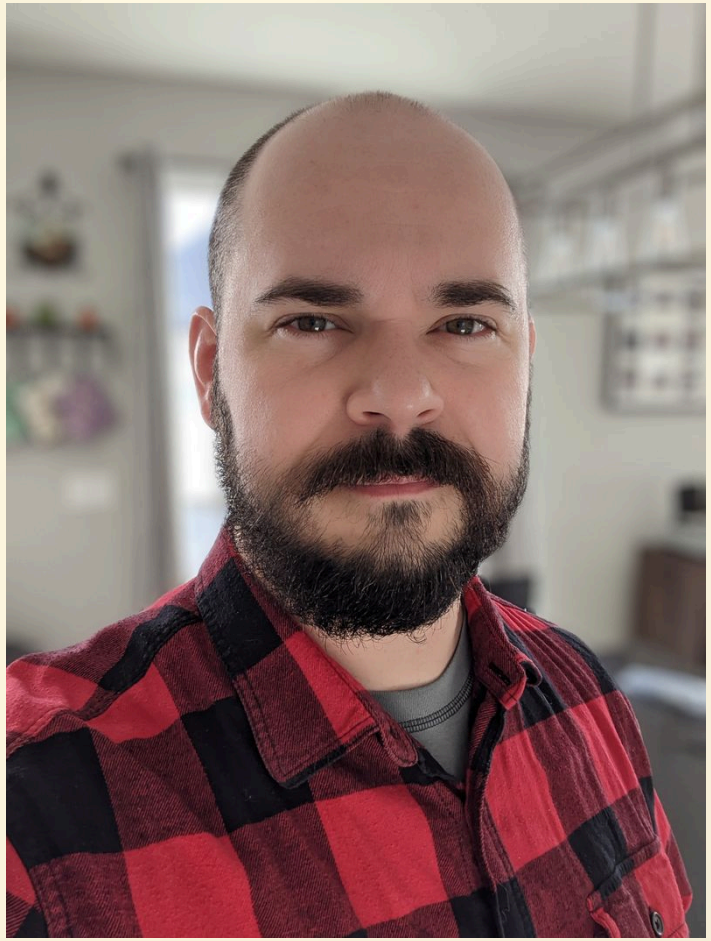
Our shipping data is more valuable than we realize...

## Jakob Anderson

I've been building Machine Learning & other AI software for over 5 years, other software for 30 years.

### AI & ML Apps that:

- Write stories
- Compose music
- Remix image styles
- Invent phrases
- Detect plagiarism



## The Problem

### Traditional shipping estimates fail:

- Based on simple lookup tables
- Inaccurate (±2-3 days variance)
- Lead to customer dissatisfaction
- Can't handle real-world complexity like weekends, holidays, days of the week, seasonal shifts

### Business Impact:

- Lost revenue from missed deliveries
- Higher customer support costs
- Poor operational planning

## Our Solution

**73% more accurate predictions using ML**

| Metric | Traditional | ML-Based | Improvement |
|---|---|---|---|
| Accuracy | ±2.5 days | ±0.68 days | **73% better** |
| Speed | Slow | <50ms (cached) | Real-time |

**How it works:**
Historical Data → Feature Engineering → LightGBM → Predictions

**Powered by:** 29,000+ shipments, 15+ features, gradient boosting

## How Gradient Boosting Works

### Builds trees sequentially, each correcting previous errors

### Why LightGBM?

- Handles complex patterns automatically
- Fast: 10-50ms per prediction (cached)
- Robust to missing data
- Production-ready

### Key Features Used:

- Temporal patterns (day/month)
- Routes & carrier combinations
- Seasonal historical trends
- Package characteristics

## 🚀 Increases Productivity

**Before:** Trusting the inaccurate carrier estimates, using carrier-provided charts that don't accommodate for seasons, weekends, zones
**After:** Data-driven predictions for better carrier comparison

**Benefits:**

- Additional metric for rate shopping decisions
- Compare predicted transit times across carriers
- Make informed tradeoffs between cost and speed
- Batch processing for large shipment volumes

## ✨ Improves Work Quality

### Precision & Reliability:

- 73% more accurate predictions
- Uncertainty ranges (best/expected/worst case)
- Data-driven decision making
- Statistical distribution analysis

### Results:

- Fewer missed delivery windows
- More accurate customer promises
- Better resource allocation

## 🎯 Unlocks New Capabilities

### Uncertainty Quantification:

- Best case: 2.1 days
- Expected: 3.2 days
- Worst case: 4.8 days

### New Services:

- Offer guaranteed delivery windows
- Real-time analytics dashboard
- Carrier performance comparison
- Paid API for third-party apps

## 💰 Makes Money

### Speculative Business Value of Accurate Predictions:

- Paid service for third-party apps. Use predictions from our shipping data as a valuable resource to inform other's decisions.
- Avoid costly late deliveries and lost sales.
- Prevent loss of customer trust and reduce the need for service credits or refunds.
- Less wasted resources, fewer emergency shipments, and improved inventory management.
- On-time shipments can be a key differentiator in competitive markets.
- Loss aversion: Businesses are more sensitive to losses from missed deadlines than to gains from faster delivery—ML helps minimize these risks.

### Bottom Line:

- Accurate transit time prediction is a strategic asset that protects revenue, reduces operational losses, and enables new business opportunities.

## Beyond Transit Time: Future Opportunities

### Logistics:

- Demand forecasting for resource planning
- If we know which carriers/levels slow down for each season, we can route shipments to those that don't slow down
- Dynamic pricing based on capacity

### Business Intelligence Uses for Data Mining & ML predictions:

- Customer churn prediction
- Fraud detection with real-time scoring
- Risk assessments
- Many other paid api prediction services using historical shipping data

## Next Steps

### Current System:

✅ Trained models with 73% better accuracy than carrier charts
✅ REST API
✅ Analytics dashboard deployed

### Your Pilot Project:

1. Integrate with your shipment data
2. live proof of concept, showing informative estimate
3. Measure ROI on real shipments
4. Scale across operations

### Questions?