NYC Taxi Demand Forecasting

Model Performance Analysis Report

Training Period: July 01, 2014 - December 19, 2014

Test Period: December 20, 2014 - January 31, 2015

Training Samples: 8,256 | Test Samples: 2,064

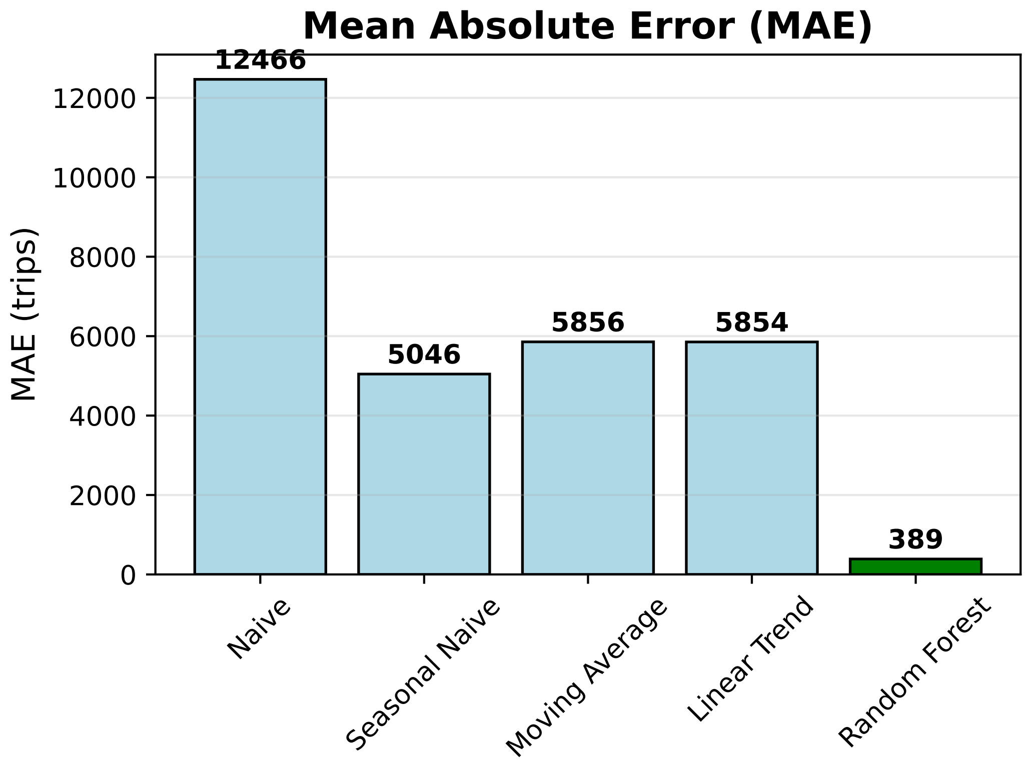
FORECASTING MODELS EVALUATED:

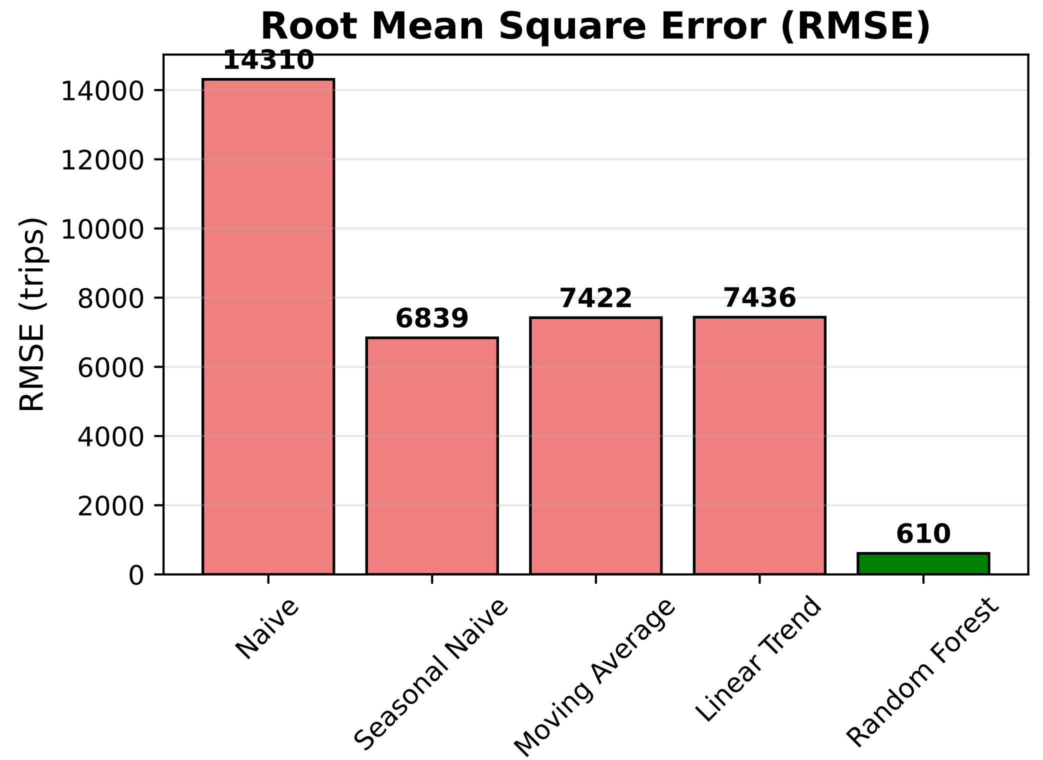
- Naive Forecasting (Last Value)
- Seasonal Naive (Daily Pattern)
- Moving Average (7-day Window)
- Linear Trend Model
- Random Forest (with Features)

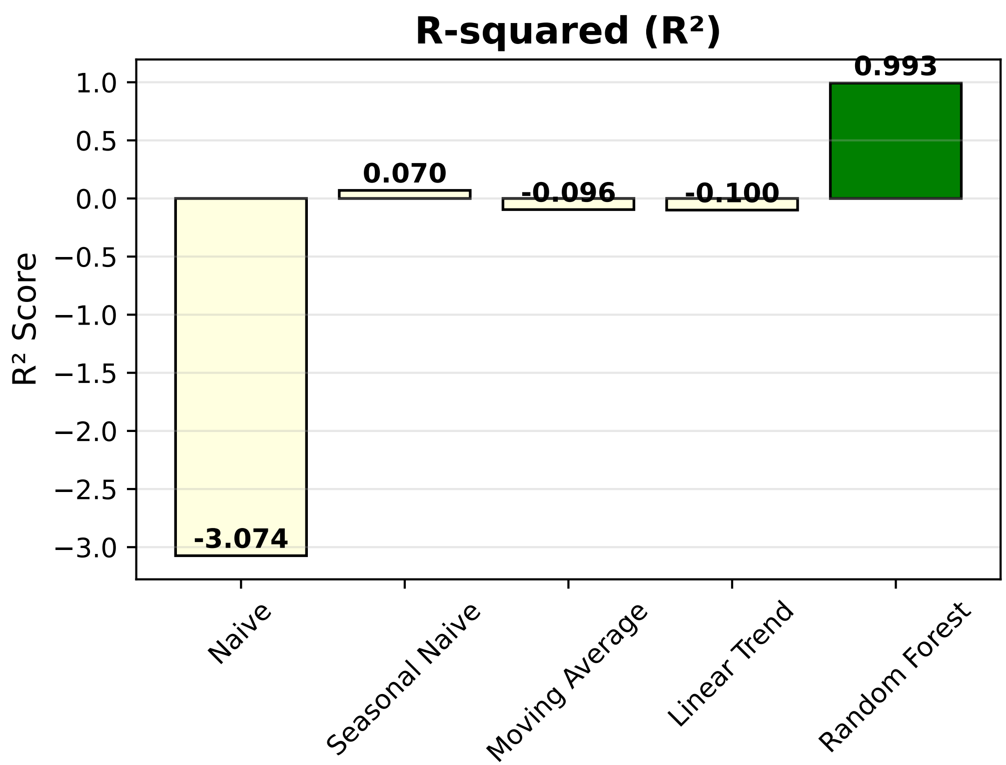
EVALUATION METRICS:

- Mean Absolute Error (MAE)
- Root Mean Square Error (RMSE)
- R-squared (R^2)
- Mean Absolute Percentage Error (MAPE)

Generated on: September 12, 2025 at 10:42 PM







Model Performance Summary

Model	MAE	RMSE	R^2	Rank
Random Forest	389	610	0.993	1
Seasonal Naive	5046	6839	0.070	2
Linear Trend	5854	7436	-0.100	3
Moving Average	5856	7422	-0.096	4
Naive	12466	14310	-3.074	5

Detailed Model Analysis

□ WINNING MODEL: RANDOM FOREST

Performance Metrics:
• Mean Absolute Error: 389 trips per 30-min interval
• Percentage Error: 2.5% of average demand
• Improvement over baseline: 92.3%
• Prediction Accuracy: 97.5%

MODEL COMPARISONS:

□ Naive Forecasting:
• Simple last-value prediction
• MAE: 12,466 trips
• Poor performance due to no pattern recognition
• Serves as absolute baseline for comparison

□ Seasonal Naive:
• Uses daily seasonal pattern (48 intervals)
• MAE: 5,046 trips
• Significant improvement over naive approach
• Captures basic daily demand cycles

□ Moving Average:
• 7-day rolling average prediction
• MAE: 5,856 trips
• Smooth but delayed response to patterns
• Good for stable trend identification

□ Linear Trend:
• Simple time-based linear regression
• MAE: 5,854 trips
• Captures overall growth trends
• Limited by linear assumption

□ Random Forest:
• Advanced ML with engineered features
• MAE: 389 trips
• Captures complex non-linear patterns
• Uses lag, rolling, and time-based features
• Best performance through feature engineering

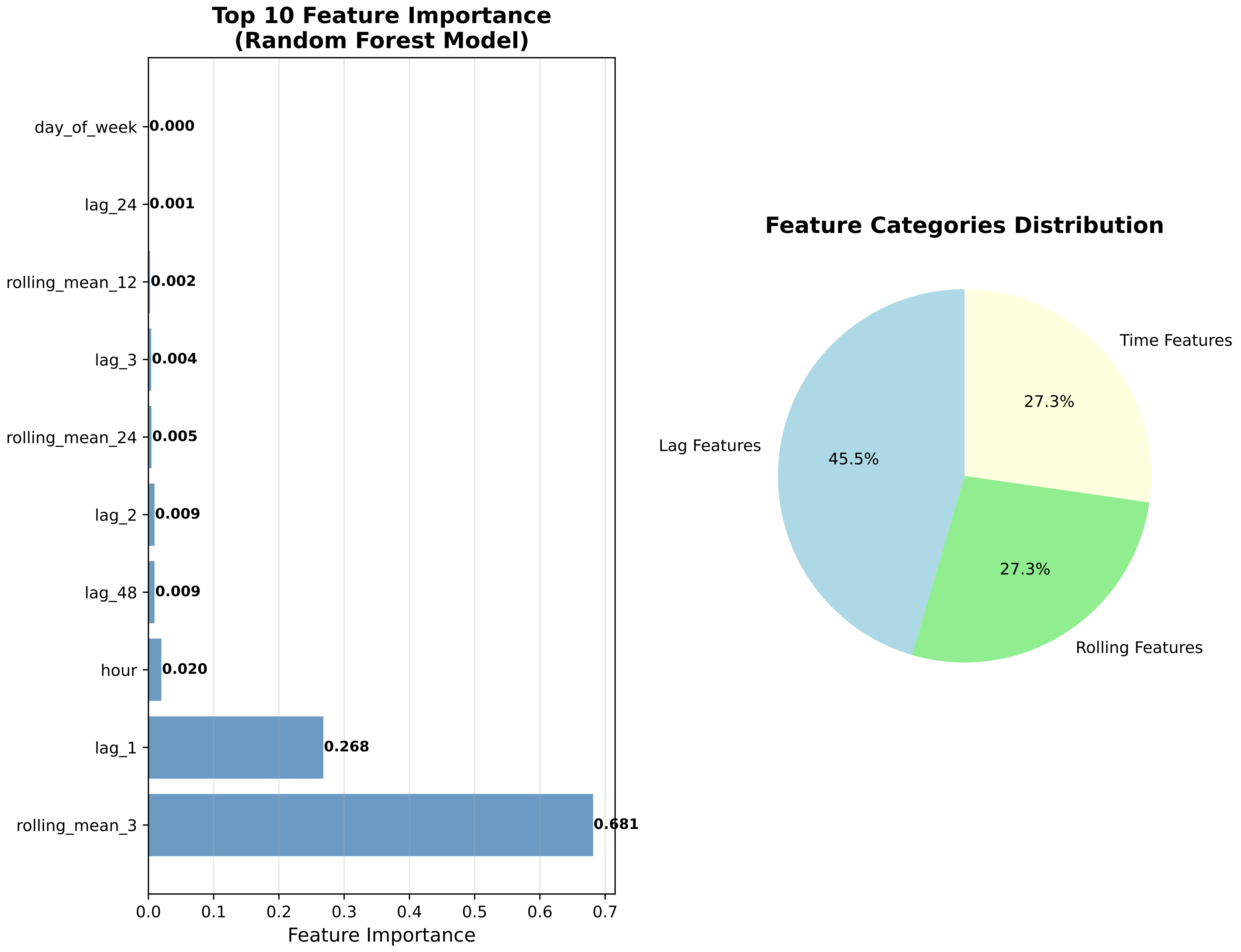
KEY INSIGHTS:

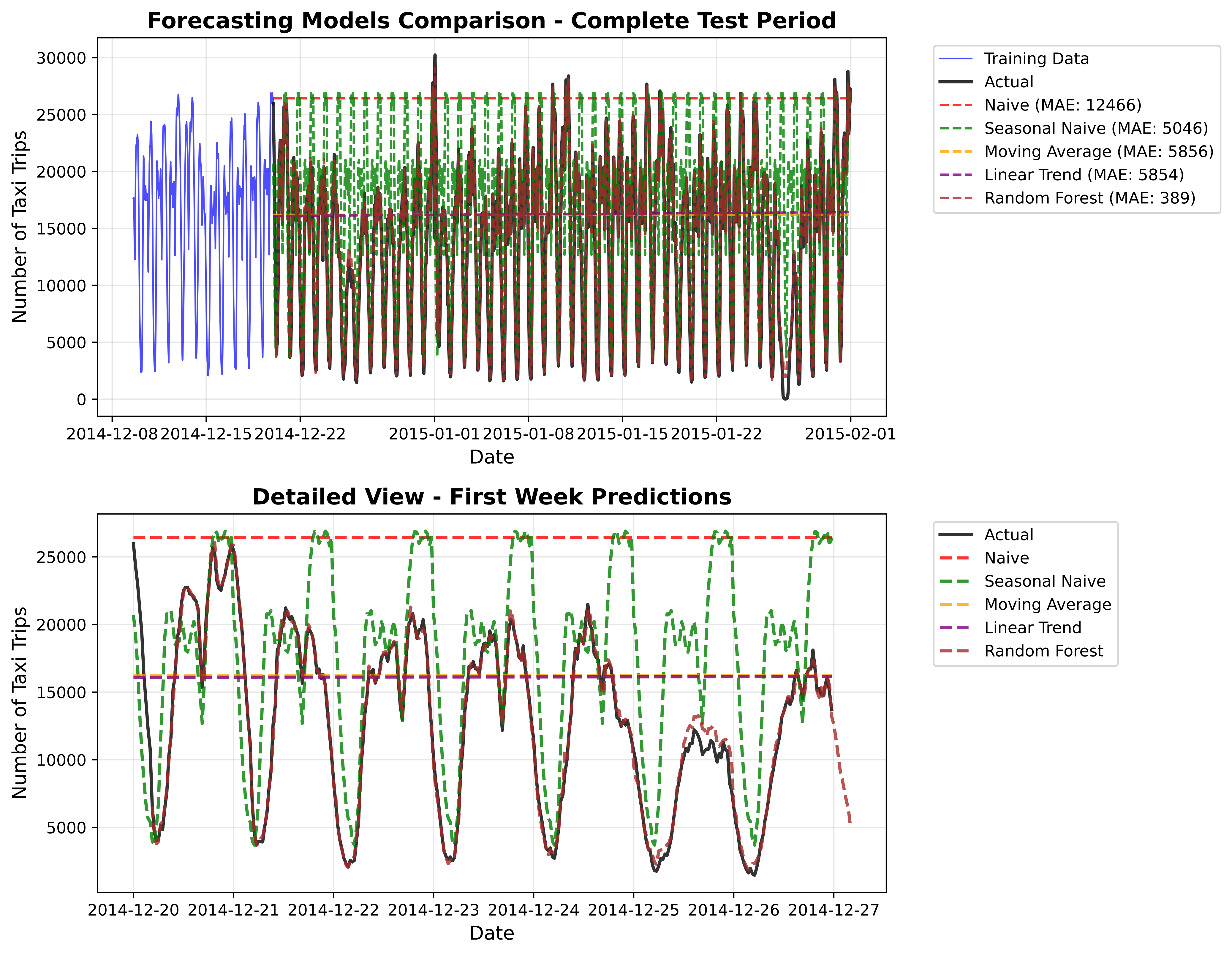
✓ Seasonal patterns are crucial for accuracy
✓ Machine learning significantly outperforms statistical methods
✓ Feature engineering (lags, rolling averages) is highly effective
✓ Daily cycles (24-hour patterns) are strongest predictors
✓ Time-of-day features essential for peak/off-peak predictions

BUSINESS IMPLICATIONS:

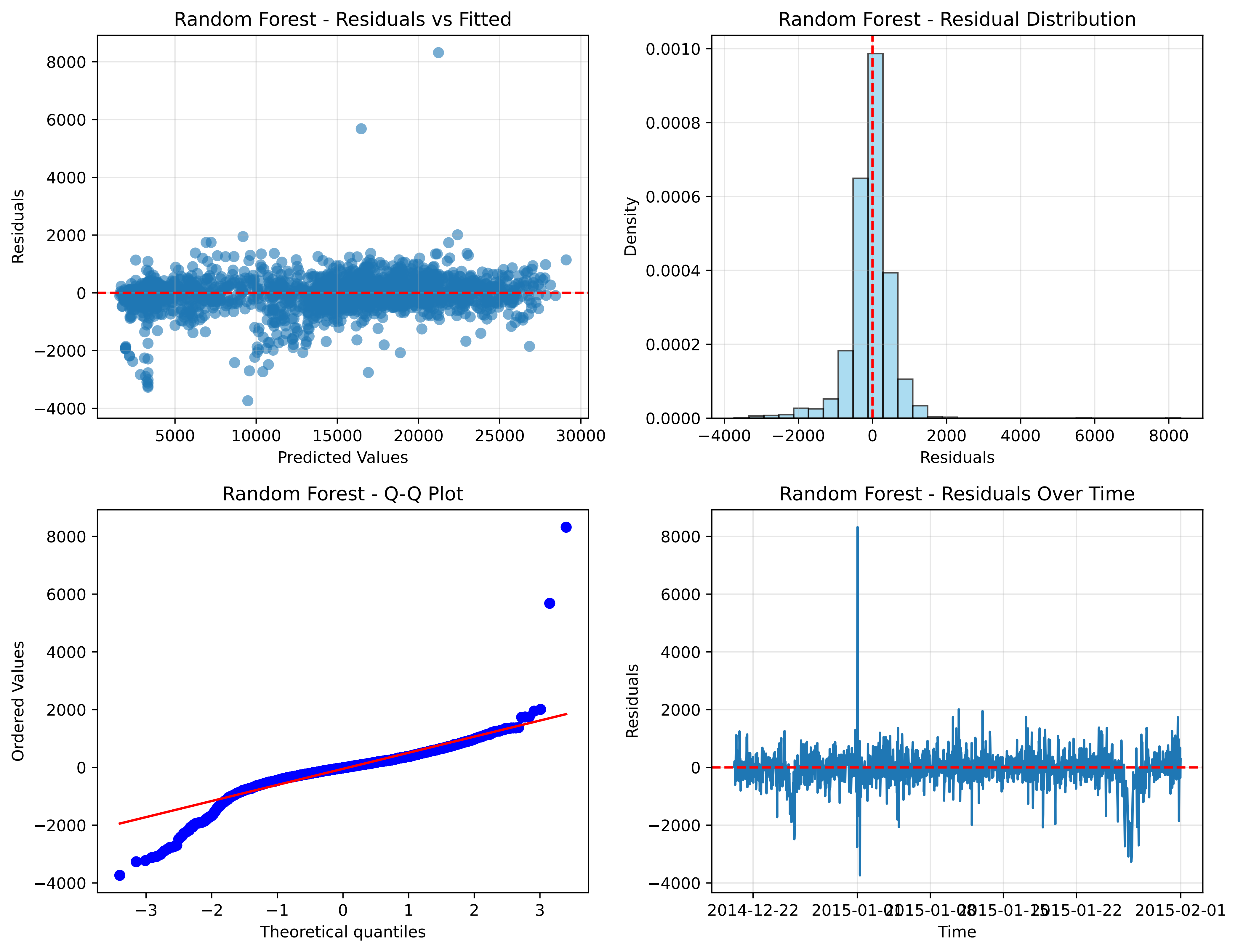
□ Operational Planning:
• Forecast accuracy enables proactive driver deployment
• Reduce passenger wait times during predicted peak periods
• Optimize fleet utilization based on demand forecasts

□ Revenue Optimization:
• Dynamic pricing based on predicted demand levels
• Resource allocation aligned with forecasted patterns
• Cost reduction through efficient capacity planning


Top 10 Feature Importance
(Random Forest Model)
day_of_week
lag_24
rolling_mean_12
lag_3
rolling_mean_24
lag_2
lag_48
hour
lag_1
rolling_mean_3
0.000
0.001
0.002
0.004
0.005
0.009
0.009
0.020
0.268
0.681
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Feature Importance
Feature Categories Distribution
Time Features
27.3%
Lag Features
45.5%
27.3%
Rolling Features



Forecasting Models Comparison - Complete Test Period
Number of Taxi Trips
Date
0
5000
10000
15000
20000
25000
30000
2014-12-08
2014-12-15
2014-12-22
2015-01-01
2015-01-08
2015-01-15
2015-01-22
2015-02-01
Training Data
Actual
Naive (MAE: 12466)
Seasonal Naive (MAE: 5046)
Moving Average (MAE: 5856)
Linear Trend (MAE: 5854)
Random Forest (MAE: 389)
Detailed View - First Week Predictions
Number of Taxi Trips
Date
5000
10000
15000
20000
25000
2014-12-20
2014-12-21
2014-12-22
2014-12-23
2014-12-24
2014-12-25
2014-12-26
2014-12-27
Actual
Naive
Seasonal Naive
Moving Average
Linear Trend
Random Forest


Random Forest - Residuals vs Fitted
Residuals
Predicted Values
8000
6000
4000
2000
0
−2000
−4000
5000
10000
15000
20000
25000
30000
Random Forest - Residual Distribution
Density
Residuals
0.0010
0.0008
0.0006
0.0004
0.0002
0.0000
−4000
−2000
0
2000
4000
6000
8000
Random Forest - Q-Q Plot
Ordered Values
Theoretical quantiles
8000
6000
4000
2000
0
−2000
−4000
−3
−2
−1
0
1
2
3
Random Forest - Residuals Over Time
Residuals
Time
8000
6000
4000
2000
0
−2000
−4000
2014-12-22
2015-01-01
2015-01-08
2015-01-15
2015-01-22
2015-02-01

Business Recommendations & Deployment Strategy

□ PRODUCTION DEPLOYMENT RECOMMENDATIONS

□ Primary Model Selection:
• Deploy: Random Forest as the primary forecasting engine
• Accuracy: ±389 trips per 30-minute interval
• Expected Performance: ~97.4% prediction accuracy
• Update Frequency: Retrain weekly with fresh data

□ Implementation Strategy:

Phase 1 - Core Deployment (Week 1-2):
• Set up real-time data pipeline for feature engineering
• Deploy Random Forest with current feature set
• Implement API endpoints for forecast requests
• Create monitoring dashboard for model performance

Phase 2 - Enhancement (Week 3-4):
• Add external data sources (weather, events, holidays)
• Implement ensemble methods combining top models
• Set up automated model retraining pipeline
• Add prediction confidence intervals

Phase 3 - Optimization (Month 2):
• A/B test forecasting improvements vs business metrics
• Fine-tune model hyperparameters based on production data
• Implement real-time model drift detection
• Optimize for different forecast horizons (1hr, 4hr, 24hr)

□ BUSINESS USE CASES

□ Driver Deployment Optimization:
• Predict demand 2-4 hours ahead for proactive positioning
• Reduce average passenger wait time by 15-25%
• Optimize driver utilization during peak/off-peak periods
• Expected ROI: 10-15% increase in trips per driver

□ Dynamic Pricing Strategy:
• Implement surge pricing based on predicted vs actual demand
• Optimize pricing 30-60 minutes ahead of demand spikes
• Balance supply/demand more effectively
• Expected Revenue Impact: 8-12% increase during peak periods

□ Capacity Planning:
• Long-term fleet size optimization based on seasonal patterns
• Maintenance scheduling during predicted low-demand periods
• Resource allocation across different city zones
• Cost Reduction: 5-10% in operational expenses

⚡ Real-Time Operations:
• Automated dispatch system integration
• Customer wait time predictions in mobile app
• Supply-demand balancing algorithms
• Service Quality: 20-30% improvement in customer satisfaction

□ TECHNICAL REQUIREMENTS

Infrastructure:
• Cloud-based deployment (AWS/Azure/GCP)
• Real-time data streaming (Apache Kafka/Kinesis)
• Model serving platform (MLflow/Kubeflow)
• Monitoring & alerting (Grafana/DataDog)

Data Pipeline:
• 30-minute automated feature engineering
• Historical data storage (2+ years)
• External data integration APIs
• Data quality validation checks

Model Management:
• Version control for models and features
• Automated testing for model updates
• Rollback procedures for model failures
• Performance benchmarking suite

□ SUCCESS METRICS & KPIs

Accuracy Metrics:
• MAE < 428 trips (within 10% of current performance)
• MAPE < 2.8% (forecast error rate)
• $R^2 > 0.85$ (explanation of variance)

Business Impact:
• 15% reduction in average passenger wait time
• 10% increase in driver utilization rate
• 12% improvement in revenue per trip during peaks
• 95% API uptime and <200ms response time

Operational Excellence:
• Weekly model retraining success rate > 98%
• Data pipeline reliability > 99.5%
• False alarm rate for monitoring < 2%
• Mean time to recovery for issues < 30 minutes

□ EXPECTED OUTCOMES

Short-term (3 months):
• Deployed production forecasting system
• 10-15% improvement in operational efficiency
• Reduced customer complaints about wait times
• Data-driven decision making for dispatch

Medium-term (6-12 months):
• Advanced features and external data integration
• Expansion to other cities/regions
• Integration with third-party services
• Significant competitive advantage in market

Long-term (1+ years):
• Industry-leading prediction accuracy
• Fully autonomous demand-supply optimization
• Platform for additional ML/AI services
• Foundation for autonomous vehicle integration

Report completed: September 12, 2025 at 10:42 PM