**Technical Report: Time Series Forecasting of Transport Modes Using Prophet**

**GitHub Repository:** [akshayaasd/kovai.co](https://github.com/akshayaasd/kovai.co)

**Key Insights from the Dataset**

1. **Clear Weekly and Seasonal Trends**  
   All five transport categories — *Local Route, Light Rail, Peak Service, Rapid Route,* and *School* — show strong weekly seasonality. School and Peak Services, in particular, peak during weekdays and drop on weekends, suggesting commuter-centric usage. This pattern can help optimize resource allocation by reducing fleet sizes or frequency on weekends.
2. **Differential Demand Across Transport Modes**
   * **Rapid Routes** show consistent usage, indicating a stable base of riders relying on high-speed transit options.
   * **School routes** have sharp, periodic peaks aligned with school calendars, suggesting demand forecasting should incorporate academic schedules.
   * **Light Rail** usage is more stable but shows slight dips during holidays, implying mixed-use between daily commuters and occasional travelers.
3. **Anomalies Indicate External Events or Disruptions**  
   Visualizations reveal spikes or drops on specific dates across modes — possibly due to public holidays, weather conditions, strikes, or local events. Integrating external datasets (e.g., event calendars, weather data) could improve predictive accuracy.
4. **Underutilized Capacity in Local Routes**  
   Local Routes show considerable volatility with lower averages and high variance. This may indicate either underutilized capacity or inconsistent service reliability. A targeted survey or real-time tracking could clarify whether route restructuring or marketing is needed.
5. **Opportunity for Dynamic Scheduling and Demand-Aware Planning**  
   Given the predictable nature of usage patterns, the city transport authority can implement dynamic scheduling. For example:
   * Increase **Peak Service** frequency during morning and evening rush hours.
   * Reduce **School Services** during exam seasons or holidays based on predicted drops.
   * Use **Light Rail** consistently, but optimize train lengths during low-demand windows.

### Forecast Summary

### ****1. Data Loading & Cleaning****

* Loaded a dataset containing daily passenger counts for different transport services.
* Converted the date column into a datetime format to support time-based analysis.
* Removed the 'Other' column due to missing values and irrelevance.
* Ensured that the remaining service columns had no missing data and were ready for forecasting.

### ****2. Data Visualization****

* Plotted time series graphs to show how passenger usage changed over time for each service.
* Created a correlation heatmap to observe how different transport services are statistically related.

### ****3. Forecasting Each Service****

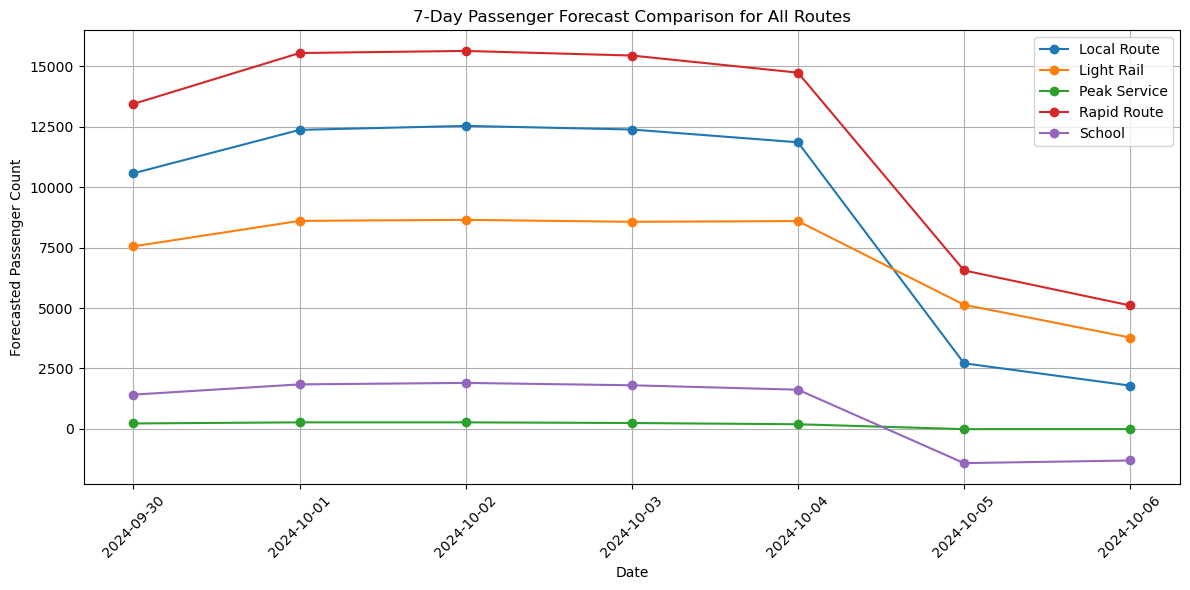
* Built a reusable function that:
  + Fits a time series model (Prophet) for each transport service.
  + Displays both the forecasted trends and their seasonal components (weekly, yearly, trend).
  + Saves the forecasted results as CSV files for later analysis.

### ****4. Detailed Forecast for One Service****

* Ran a focused 7-day forecast for "Local Route" separately.
* Visualized predicted values along with confidence intervals.
* Displayed exact forecast numbers to show future demand.

### ****5. Combined Forecast Comparison****

* Ran 7-day forecasts for all five key services.
* Merged all forecasts into one view to compare predictions across services.
* Plotted a combined line chart to visualize how each service is expected to perform in the coming week.

****