

KOVAI.CO – ASSIGNMENT

Author: Shreenithii SJ

Date: 03/06/2025

TECHNICAL REPORT

ALGORITHM CHOICE: RIDGE REGRESSION

By using L2 regularisation, Ridge Regression manages multicollinearity well and produces reliable predictions. It is perfect for forecasting transport demand because it captures both service-specific trends and more general calendar effects. Overfitting can be avoided with regularisation, especially for services like School Transport that experience seasonal fluctuations.

Important Model Elements

- Alpha (α) Regularisation
- Prevents overfitting by managing the penalty on large coefficients.
- GridSearchCV was used for optimisation in order to balance variance and bias.
- Cross-service trend learning is made possible by One-Hot Encoding (Service Type) in conjunction with Feature Engineering and Preprocessing.
- DayOfWeek, Month, and is_weekend are temporal features that aid in capturing cyclical changes in demand.
- Train-test split (shuffle=False) preserves the data's time-series integrity.

Overview of Workflow

1. **Data Preprocessing:** Encodes transport categories and extracts calendar-based features.
2. **Model Training:** To maximise performance, Ridge Regression (alpha adjusted through cross-validation) is used.
3. **Forecasting:** By repeating service types for every new date, it creates future demand for the upcoming seven days.
4. **Visualisation:** Line plots show expected ridership for various transport services.

Benefits of the Method

Utilises common travel patterns to manage several transport services collectively.

Documents changes based on time (e.g., trends between weekdays and weekends).

Forecasts are dependable during periods of fluctuation because regularisation prevents instability.

Scalable and lightweight, guaranteeing effective deployment for real-time applications.