Model Evaluation & Refinement Report

Overview

The model developed uses a gradient boosting framework (XGBoost) tailored for the time series forecasting of vehicle traffic volume. The evaluation focused on quantifying prediction accuracy, understanding model behaviour via feature importance and SHAP analysis, and iteratively refining hyperparameters using Optuna optimisation to minimise forecast errors.

Evaluation Metrics

To assess the model's performance comprehensively, the following key metrics were used:

- **Mean Absolute Error (MAE):** Measures the average magnitude of errors between predicted and actual values, providing an intuitive error scale in vehicle counts.
- **Mean Squared Error (MSE) and Root Mean Squared Error (RMSE):** Penalise large deviations more severely, useful for capturing outlier impacts on forecasts.
- **R-squared (R^2):** Indicates the proportion of variance in the observed data explained by the model predictions. Values close to 1 denote strong predictive power.
- **Mean Absolute Percentage Error (MAPE):** Provides a scale-invariant measure of prediction accuracy, useful to compare errors across different traffic volumes.

This battery of metrics ensures nuanced insight into accuracy, robustness to outliers, and explanatory power.

Performance Summary

Upon training with optimised hyperparameters, the XGBoost model showed strong predictive abilities reflected in:

- Low MAE and RMSE values indicate minimal average prediction errors and robustness to occasional large deviations.
- High R^2 scores, confirming the model captures most traffic variance, are important given the inherent noise and variability in traffic data.

These results align well with the literature on XGBoost applications in traffic prediction, where it consistently outperforms classical linear models and some neural network approaches in both precision and computational efficiency.

Model Interpretation & Insights

- **Feature Importance:**

  The analysis of feature importance provided clear interpretability, showing that lag features (previous hour/day traffic volumes) and cyclical time features (hour of day, day of week) are primary drivers in forecasting. Rolling mean and standard deviation features contributed to smoothing noise and improving generalisation.

- **SHAP Analysis:**
  SHAP values helped uncover non-linear interactions and individual feature impacts on predictions, visually summarising influential factors and their directional effects. Such insights validate domain knowledge and identify opportunities for feature engineering.

Refinement Steps & Hyperparameter Tuning

- The model underwent iterative hyperparameter optimisation via Optuna, using time series cross-validation splits to respect temporal dependencies.
- Parameters such as learning rate, tree depth, subsampling ratios, and number of estimators were tuned to balance bias-variance and avoid overfitting.
- Visualisation of optimisation history aided in diagnosing convergence and stability of performance improvements.

Error Analysis & Model Limitations

- **Temporal Error Patterns:**
  Mean absolute error analysed by hour and day-of-week revealed predictable periods of higher forecasting uncertainty, typically corresponding to rush hours or weekends with irregular traffic behaviour.
- **Error vs. Traffic Volume:**
  Residual plots showed that prediction errors slightly increased with higher actual traffic volumes, signalling room for refinement, possibly through additional explanatory variables or advanced ensembles.
- **Data Limitations:**
  The model performance is constrained by the features provided—weather, special events, and other external factors could enrich the dataset for better accuracy.

Conclusion & Future Directions

The XGBoost-based TrafficPredictor demonstrates strong promise for short-term traffic volume forecasting, achieving low errors and interpretable models with automated feature transformation and robust tuning.

Next steps recommended:

- Incorporate additional contextual features like weather, holidays, and incidents.
- Explore ensemble approaches blending XGBoost with neural network models for capturing long-range temporal dependencies.
- Extend the geographic scope of training data for more generalised models.
- Incrementally retrain models with fresh data for adapting to evolving traffic patterns.

This rigorous approach to model evaluation and refinement establishes a foundation for reliable real-world traffic prediction systems that can support better urban mobility and congestion management.