

PERFORMANCE TESTING PHASE

1. Project Title:

Advanced Traffic Volume Estimation Using Machine Learning

2. Objective of Performance Testing:

The objective of performance testing is to evaluate the model’s accuracy, efficiency, and robustness when exposed to real-world scenarios and to ensure it meets expected performance benchmarks under various data conditions.

Types of Performance Testing Conducted:

1. Model Accuracy Testing:

- Assessed using R^2 Score and RMSE (Root Mean Squared Error).*
- Compared the performance of multiple models like Linear Regression, Decision Tree, Random Forest, SVR, and XGBoost.*

2. Model Comparison:

Model	R^2 Score (Train)	R^2 Score (Test)				
-----	-----	-----				
Linear Regression	0.66	0.65				
Decision Tree	0.99	0.89		Random Forest	0.99	0.92
SVR	0.50	0.45		XGBoost	0.95	0.88

Conclusion: Random Forest gave the best performance among all tested models.

3. Execution Time Testing:

- Verified model response time after form submission using Flask Web Application.*
- Output generated within 1–3 seconds for each prediction request.*

4. Load Testing (Manual):

- *Manually tested form submissions with different sets of inputs.*
- *Model handled multiple requests sequentially without errors.*

Testing Tools Used:

- *Python Libraries: scikit-learn, xgboost, pandas, numpy*
- *Web Framework: Flask*
- *IDE: VS Code*
- *Browser Testing: Google Chrome*

Challenges Identified:

- *Encoding issues for categorical values like holiday and weather.*
- *Errors due to unseen labels in form inputs which were not part of training data.*
- *Order of input features did not match during prediction phase initially.*

Actions Taken to Fix Issues:

- *Used LabelEncoder and saved encoder using pickle to maintain consistent label handling.*
- *Re-trained model including all required fields.*
- *Ensured consistent column order in training and prediction.*
- *Cleaned and preprocessed data including missing values and formatting.*

Conclusion:

The Traffic Volume Estimation model was successfully tested for performance. Based on accuracy and stability, the Random Forest Regressor was selected for deployment. The application performs reliably and generates predictions with minimal latency.