# Traffic Prediction Model Analysis & Recommendations

INSIGHTS FROM GRADIENT BOOSTING REGRESSOR PERFORMANCE

# Introduction & Objective

- Objective: Predict traffic flow using historical time-series data

- Focus on accurate forecasting for junction-level traffic patterns

- Model used: Gradient Boosting Regressor

# Data Preparation Overview

- Data source: processed_traffic_data.csv
- Time-based features: hour, day, month, weekend indicator
- Lag features: 1, 24, 168-hour traffic history
- Time-series split: 80% training, 20% validation
- Feature scaling using StandardScale

# Model Training & Tuning

- ▶ Model: Gradient Boosting Regressor (ensemble-based)

- ▶ Cross-validation: TimeSeriesSplit with 5 folds

- ▶ Hyperparameter tuning via GridSearchCV

- ▶ Optimized using negative mean squared error

# Model Performance – Key Metrics

- Mean Absolute Error (MAE): 0.9996

- Root Mean Squared Error (RMSE): 1.3923

- R-squared ($R^2$): -0.1139

- Negative $R^2$ indicates poor predictive performance

# Visual Analysis of Predictions

- Predictions are overly smooth compared to actual traffic
- Fails to capture sharp peaks and sudden drops
- Under-predicts during high traffic, over-predicts during low traffic
- Visuals confirm numerical evaluation results

# Recommendations – Feature Engineering

- Add advanced lag and peak-hour-specific features
- Use rolling mean and rolling standard deviation features
- Apply cyclical transformations for time variables
- Incorporate external data such as weather and holidays

# Recommendations – Model & Tuning

- Try advanced models: XGBoost, LightGBM, CatBoost
- Explore deep learning models: LSTM, GRU
- Use ARIMA/SARIMA as baseline models
- Adopt Randomized or Bayesian hyperparameter optimization

# Recommendations – Analysis & Preprocessing

- ▶ Perform detailed error analysis
- ▶ Detect and handle outliers appropriately
- ▶ Check stationarity and apply differencing if required
- ▶ Review data quality and consistency

# Conclusion

- Current model performance is inadequate (negative $R^2$)

- Significant scope for improvement through feature engineering and modeling

# Next Steps

- Next steps: phased model redevelopment and evaluation
- Focus on accuracy, robustness, and real-world applicability