****

**Assessment Report**

on

**“Traffic Volume Prediction”**

submitted as partial fulfillment for the award of

**BACHELOR OF TECHNOLOGY**

SESSION 2024-25

in

**CSE(AI)**

By

Group 9:

Anushka Rajput (202401100300059)

Arpit (202401100300065)

Aajneya Nayak (202401100300001)

Aditya Kanaujiya (202401100300015)

Ankit Singh Yadav (202401100300050)

Section: A

**Under the supervision of**

“Bikki Kumar Gupta”

**KIET Group of Institutions, Ghaziabad**  
 **27 May, 2025**

**KIET Group of Institutions, Ghaziabad  
  
May, 2025**

# 1. Introduction

Urban traffic congestion poses a significant challenge for city planners and commuters. Predictive analytics, especially using machine learning, can help alleviate this by forecasting traffic patterns based on time and weather data. This project builds a regression model to predict traffic volume using historical weather and timestamped traffic data from the Metro Interstate Traffic Volume dataset.

# 2. Problem Statement

To predict the volume of traffic using a regression model based on time-based and weather-related features. The solution must include appropriate feature engineering, model training, performance evaluation, and result visualization.

# 3. Objectives

- Perform feature extraction and encoding on weather and temporal data.

- Build and train a regression model to predict traffic volume.

- Evaluate the model's accuracy using regression metrics.

- Visualize actual vs predicted traffic volumes and residuals.

# 4. Methodology

Tools Used:

Python libraries: pandas, matplotlib, seaborn, scikit-learn

Steps:

Data Loading:

The dataset /content/Metro\_Interstate\_Traffic\_Volume.csv was loaded into a Pandas DataFrame.

Feature Engineering:

- Time data (date\_time) was converted to datetime format.

- Extracted features: year, month, day, hour, weekday.

- Applied one-hot encoding on categorical features like weather type and holiday.

Data Splitting:

The dataset was divided into training and testing sets using an 80:20 ratio.

Model Training:

A Linear Regression model from sklearn.linear\_model was trained.

Evaluation & Visualization:

Used Mean Squared Error and R² score to assess performance. Visualizations included scatter plots, residual plots, and histograms.

# 5. Data Preprocessing

- Converted timestamp (date\_time) into several numeric features like hour, weekday, and month to help model temporal patterns.

- Performed one-hot encoding on categorical variables: weather\_main, weather\_description, and holiday.

- Dropped unused columns such as date\_time and the target variable from feature matrix.

- No null values were found in the final model input features.

# 6. Model Implementation

A Linear Regression model was implemented using sklearn.linear\_model.LinearRegression. The model was trained on 80% of the data and tested on the remaining 20%.

# 7. Evaluation Metrics

- Mean Squared Error (MSE): Measures the average squared difference between predictions and actual values.

- R-squared (R²): Indicates how well data fits the regression model. A value closer to 1.0 suggests a better fit.

Mean Squared Error: 141795.73

R-squared: 0.96

# 8. Results and Analysis

- The R² score of 0.34 shows a moderate ability of the linear regression model to explain variance in traffic volume.

- The scatter plot of actual vs. predicted values shows linear trends but significant variance at certain volume ranges.

- The residual plot shows that predictions are more consistent around lower volumes but deviate more at peak hours.

- The distribution of predicted values shows reasonable spread but indicates room for improvement using more advanced models (e.g., Random Forest, XGBoost).

# 9. Conclusion

The Linear Regression model moderately predicted traffic volume with basic feature engineering. While the results are promising, they suggest that the model’s capacity to capture complex traffic patterns is limited. Future improvements could explore more advanced regression techniques like Random Forest, Gradient Boosting, or Deep Learning.





  






  
