TÉCNICO LISBOA

TRAFFIC VOLUME PREDICTION USING DEEP LEARNING

INTELLIGENT SYSTEMS

Francisco Carvalho
111000

Faculty:
João Sousa

04/02/2025

DATA OVERVIEW

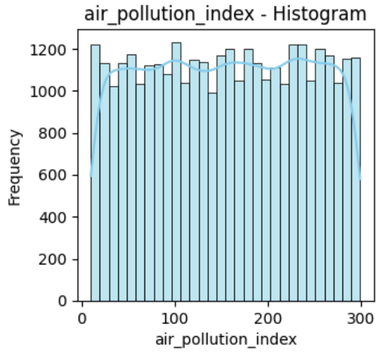
01. DATATIME

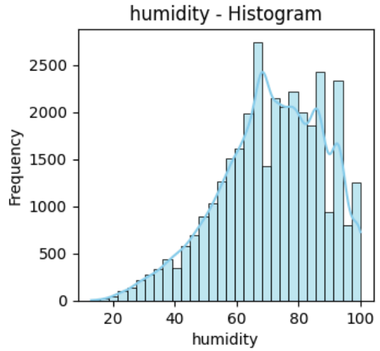
02. WEATHER DATA

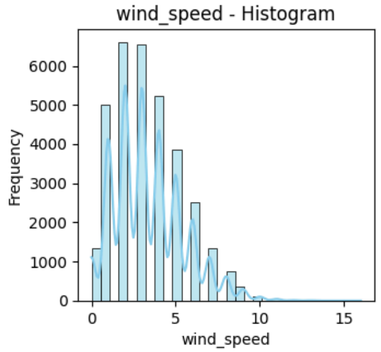
- Temperature
- Weather Description

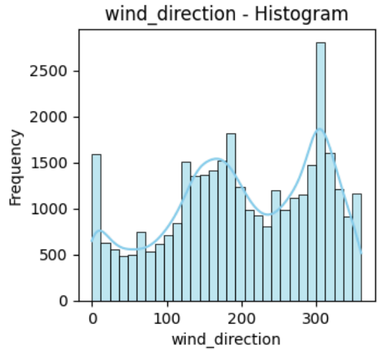
03. TRAFFIC VOLUME

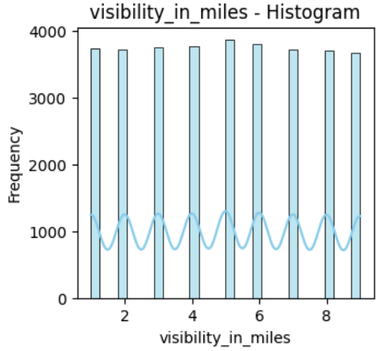
INDIVIDUAL FEATURE ANALYSIS - NUMERICAL FEATURES

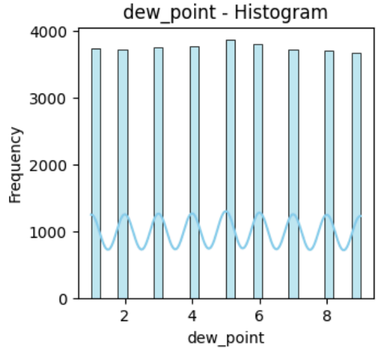


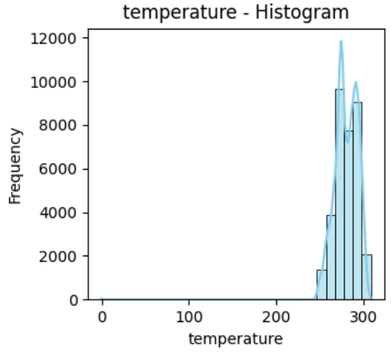


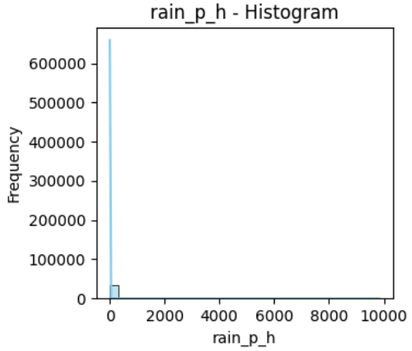


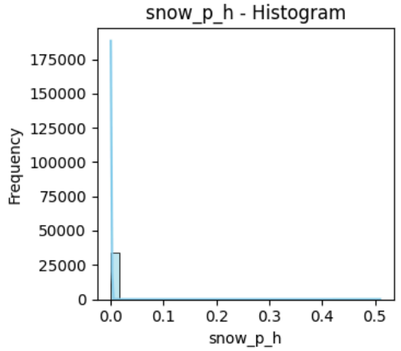


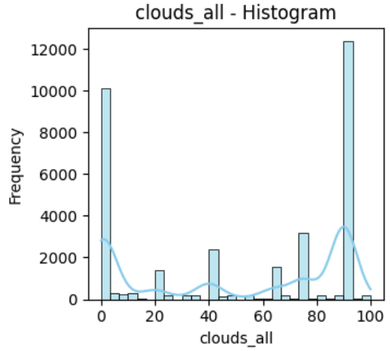


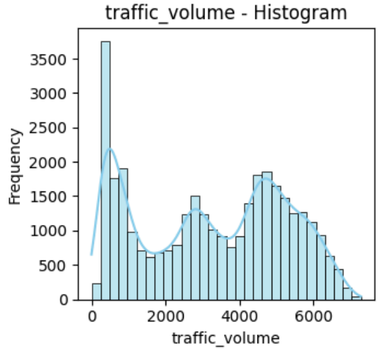


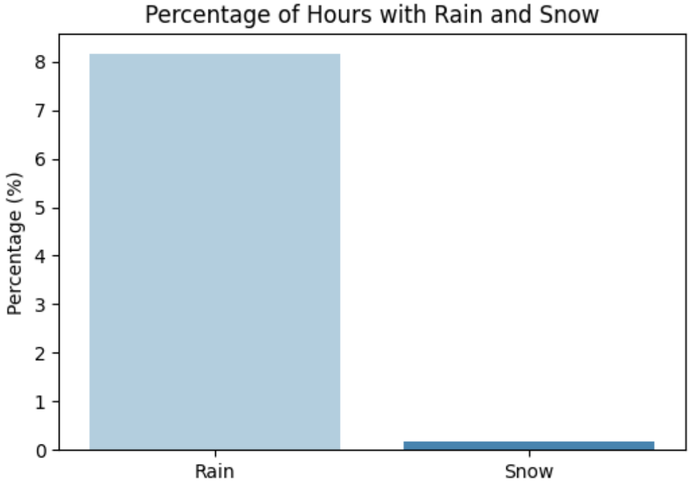




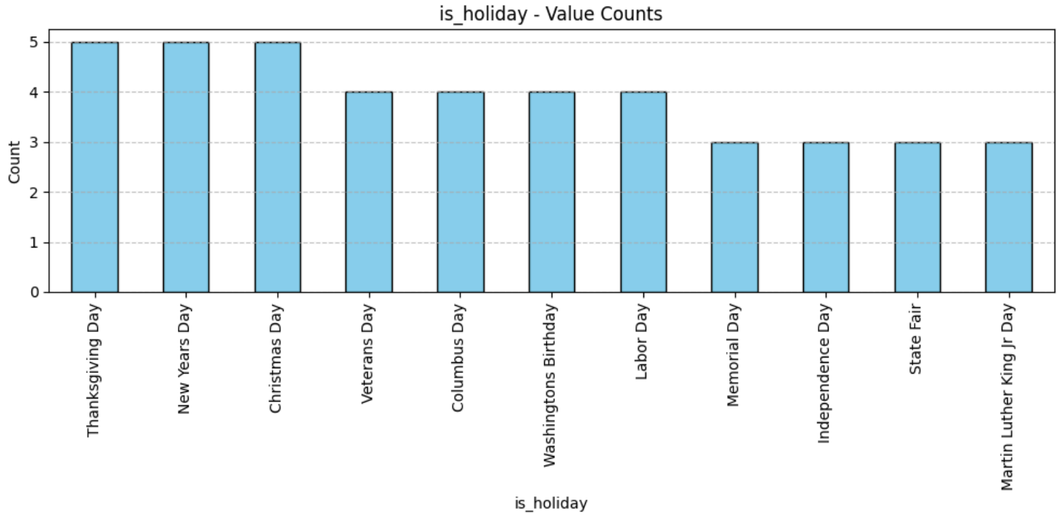


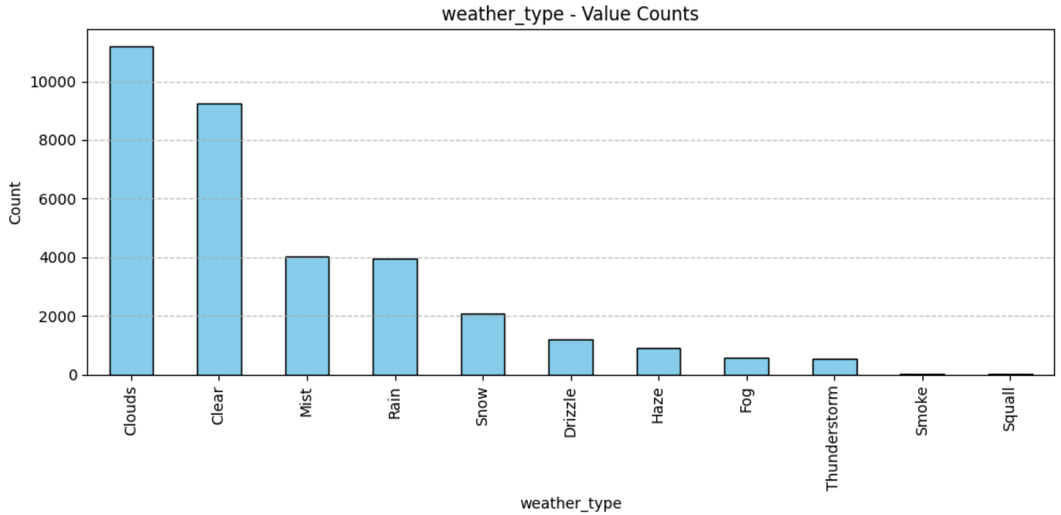


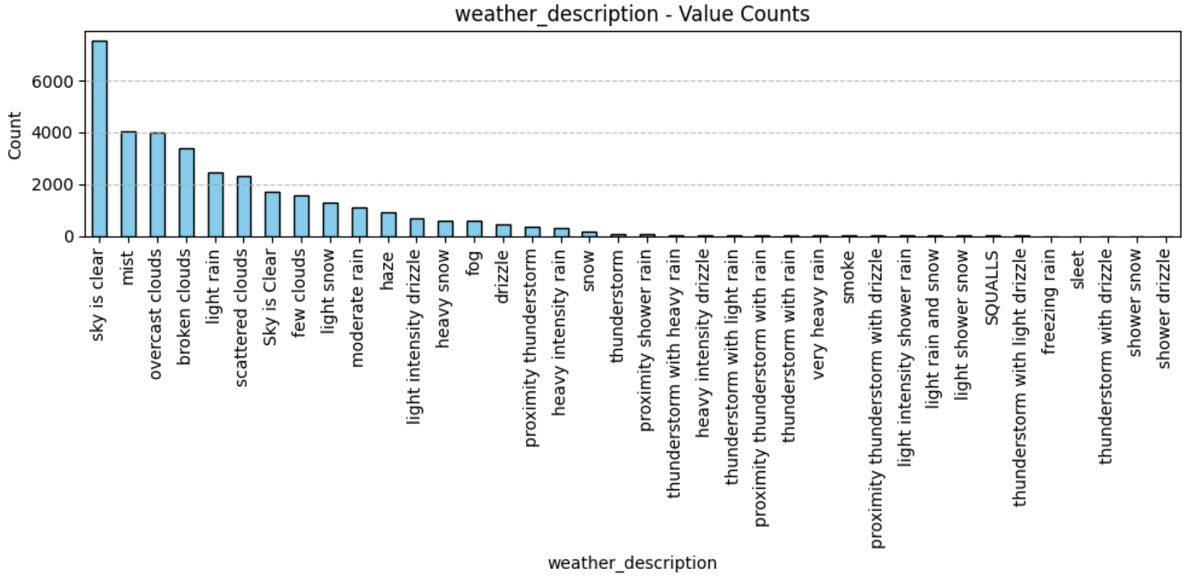




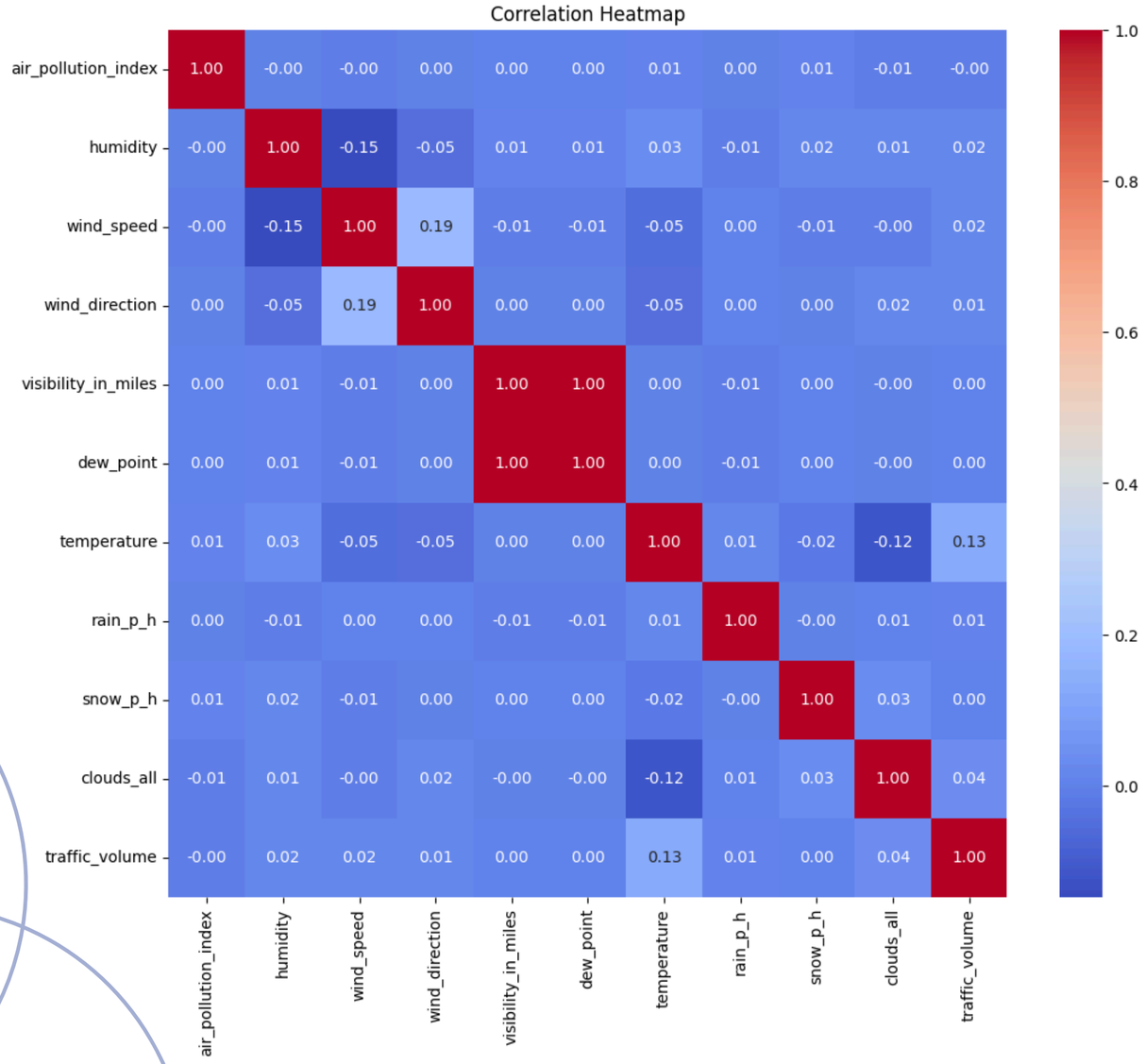
INDIVIDUAL FEATURE ANALYSIS - CATEGORICAL FEATURES





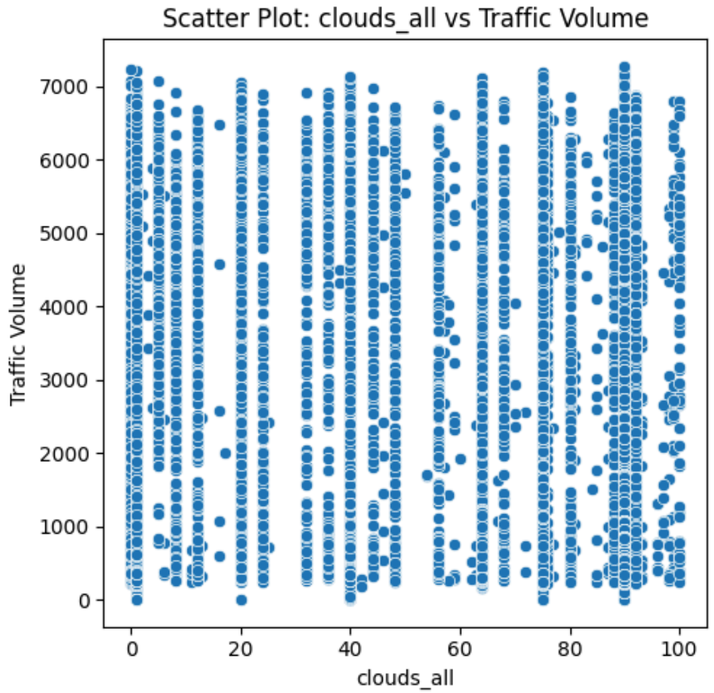


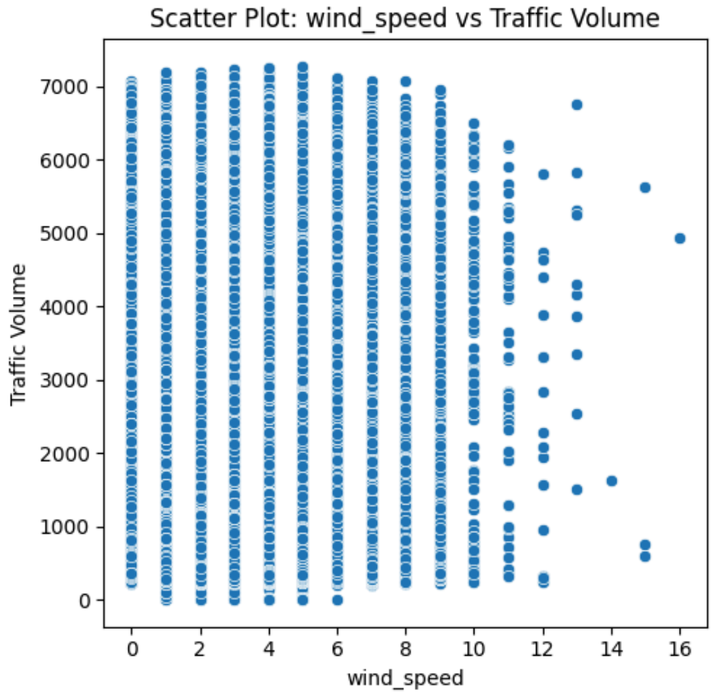
CORRELATION ANALYSIS

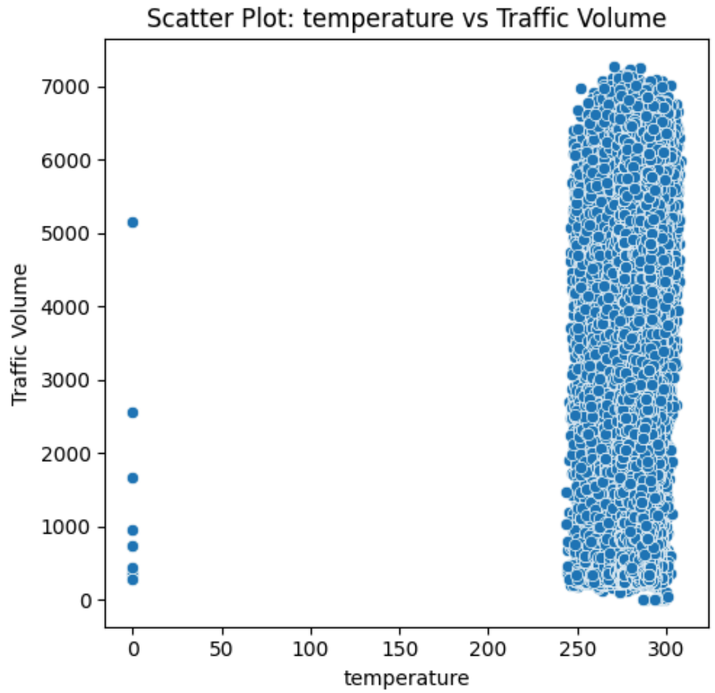


DEW_POINT WAS REMOVED FROM FURTHER ANALYSIS TO AVOID MULTICOLLINEARITY ISSUES

SCATTER PLOT ANALYSIS OF IMPORTANT FEATURES

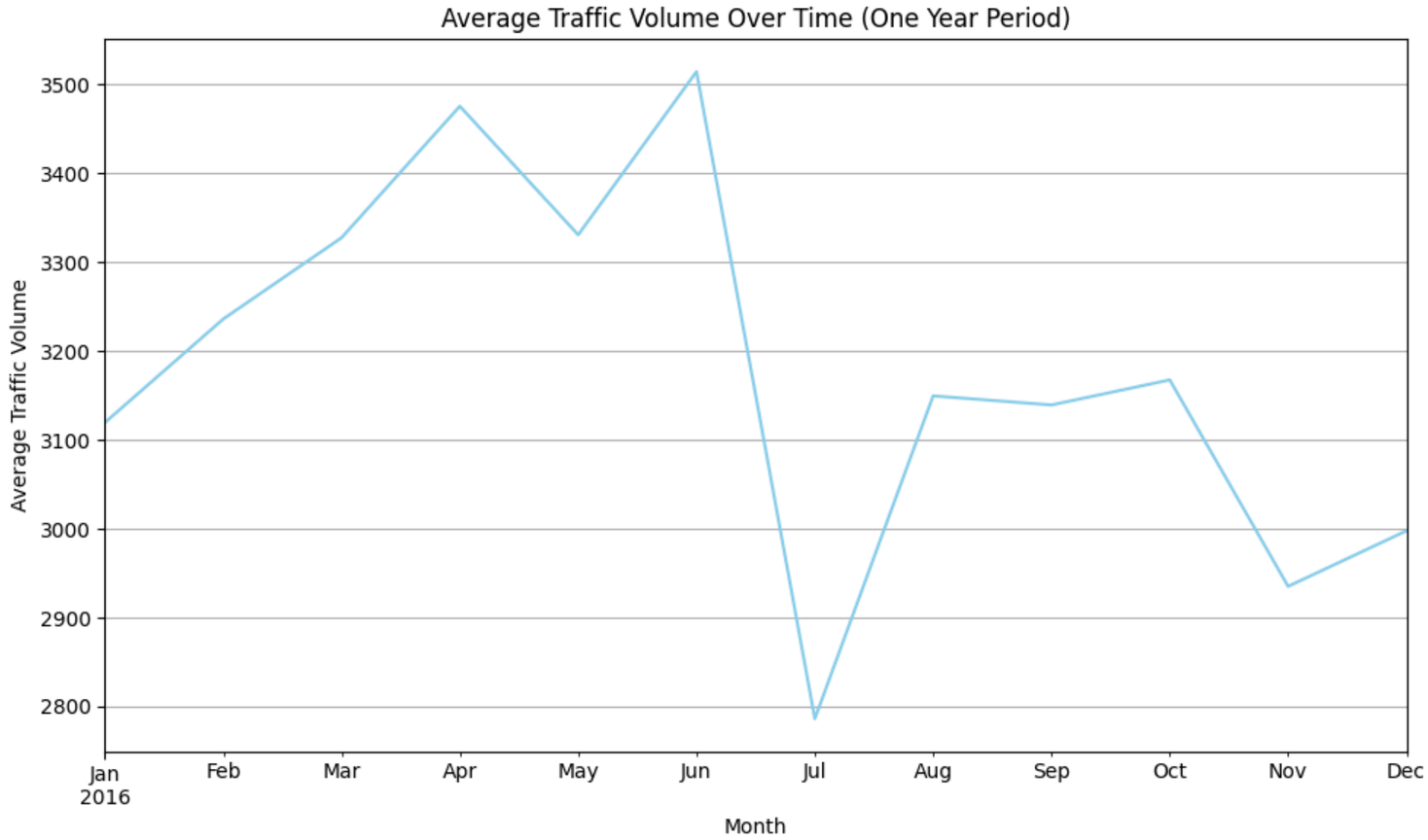






TIME SERIES ANALYSIS

The time series plot demonstrates that traffic volume fluctuates significantly, indicating periods of higher and lower traffic.



PREPROCESSING

Handling Missing Values

The analysis confirmed that the dataset did not contain any missing entries

Encoding Categorical Variables

This approach creates binary variables for each unique category

Filtering Data for a Consistent Time Period

This ensured that only continuous and reliable data was used for training and evaluation

Handling Duplicate and Incomplete Timestamps

The dataset was reindexed to enforce an hour frequency, ensuring that all hours within the recorded period were represented

Data Scaling and Splitting

To evaluate model performance effectively, the dataset was divided into training and testing subsets.

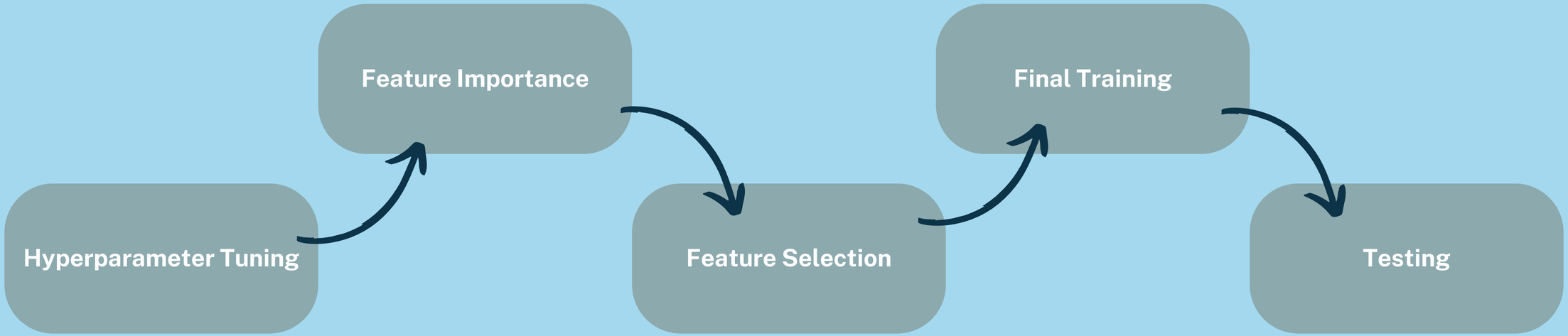
MODEL DEVELOPMENT

LSTM

GRU

TCN

XGBOOST



LONG SHORT-TERM MEMORY (LSTM)

Type of recurrent neural network (RNN) that are well-suited for modeling sequential data, particularly when long-term dependencies need to be captured

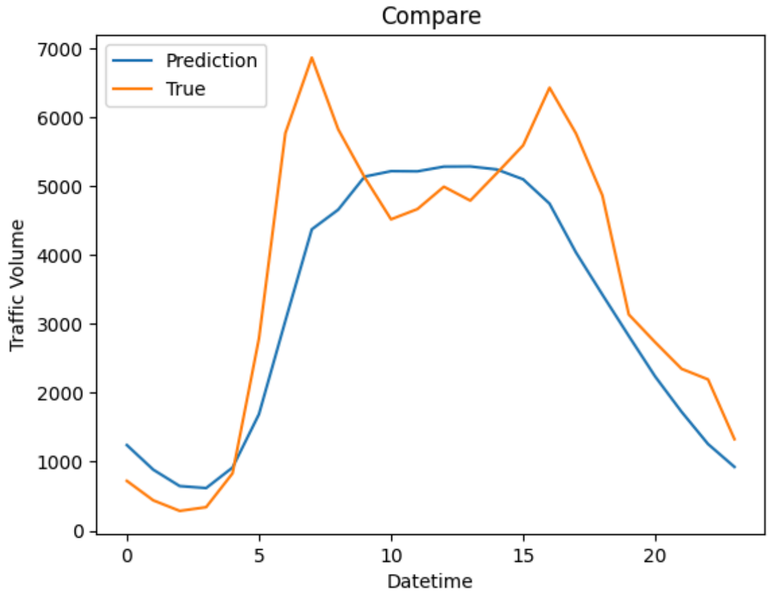
HYPERPARAMETER TUNING

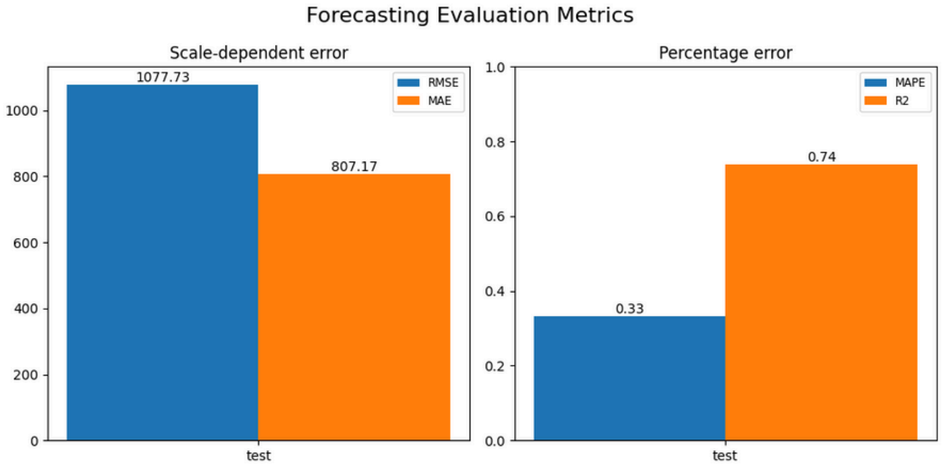
- **Number of LSTM Units: [32, 64]**
- **Learning Rates: [0.00001, 0.00005, 0.0001]**
- **Dropout Rates: [0.2, 0.5]**
- **Optimizers: Adam, RMSprop**
- **Activation Functions: 'relu', 'tanh'**
- **Number of Layers: [1, 2, 3]**

The **best-performing model configuration** was identified and saved with the following hyperparameters:

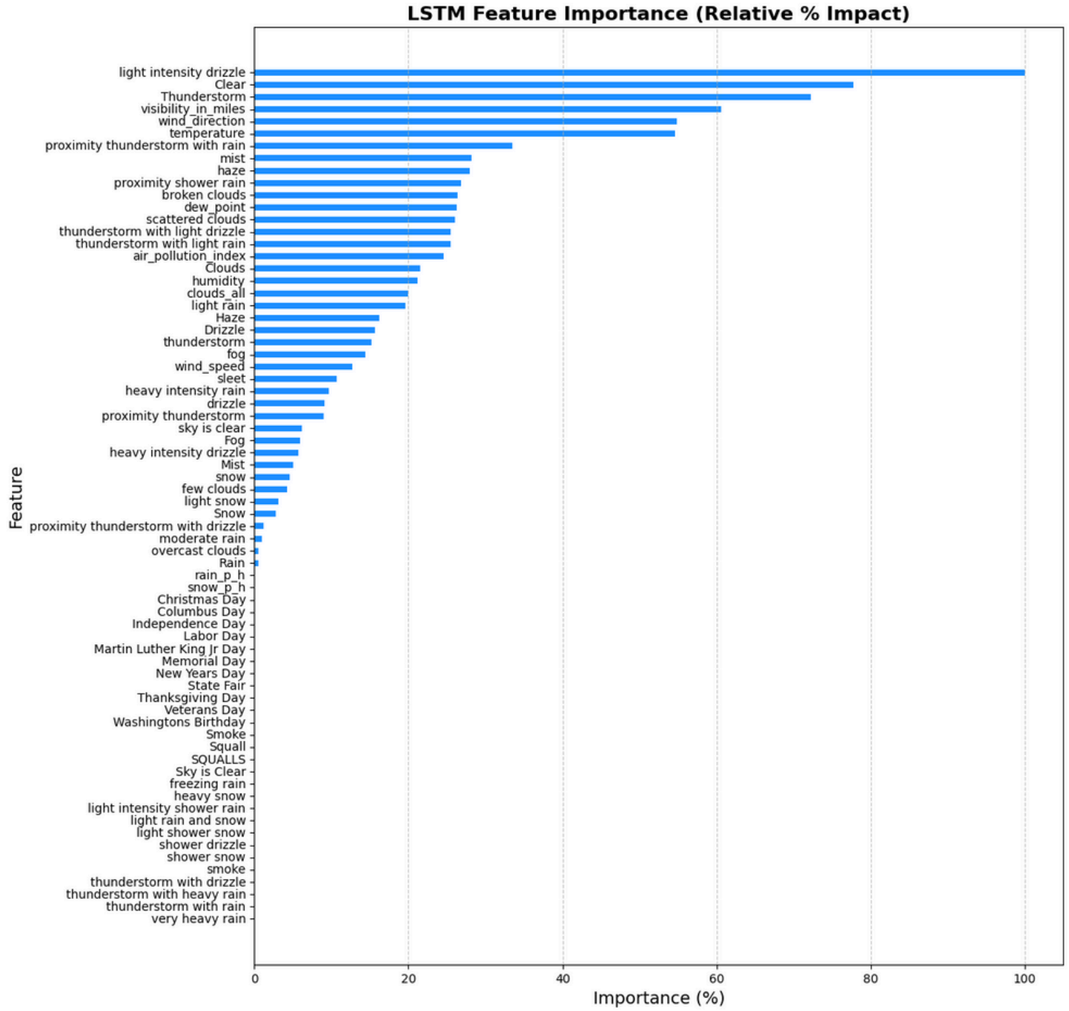
- **Number of Layers: 3**
- **LSTM Units per layer: [32, 32, 32]**
- **Optimizer: Adam**
- **Learning Rate: 0.0001**
- **Dropout Rate: 0.2**
- **Activation Function: 'tanh'**

MODEL EVALUATION WITH THE BEST CONFIGURATION

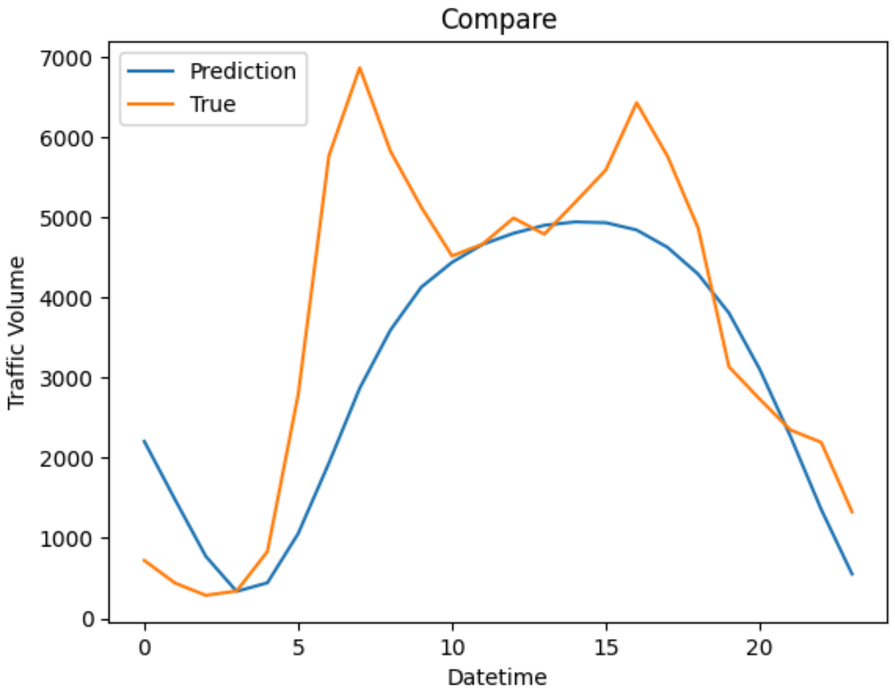


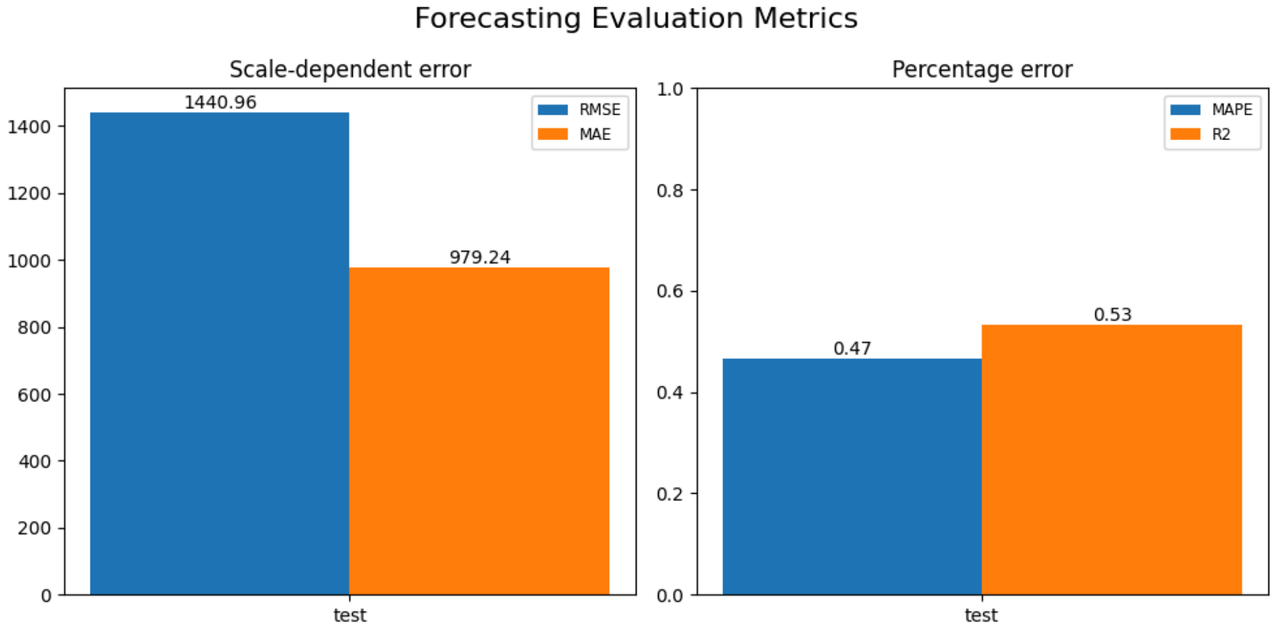


FEATURE IMPORTANCE AND SELECTION



FINAL MODEL TRAINING WITH SELECTED FEATURES





The performance did not improve as expected, indicating that feature selection might have removed useful information or that the model struggled to generalize with the reduced feature set

GATED RECURRENT UNIT (GRU)

Simplified variant of LSTMs that use gating mechanisms to capture dependencies in sequential data while reducing computational complexity

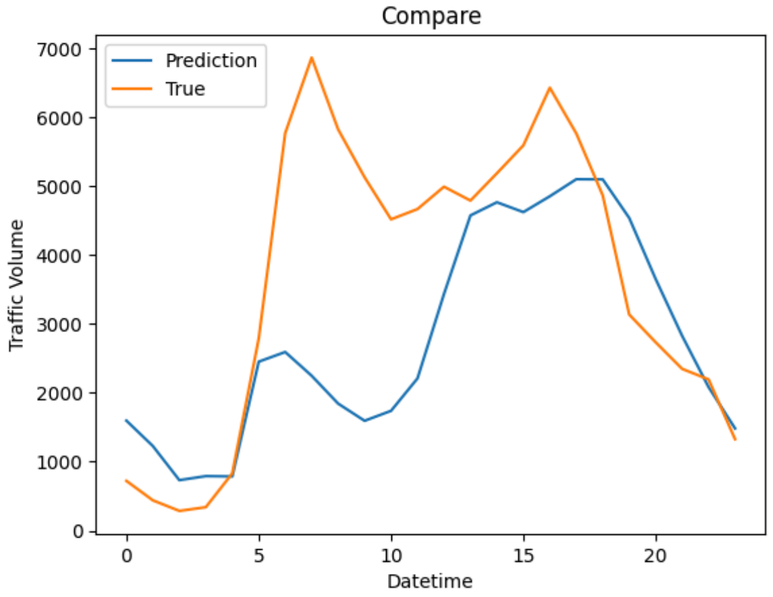
HYPERPARAMETER TUNING

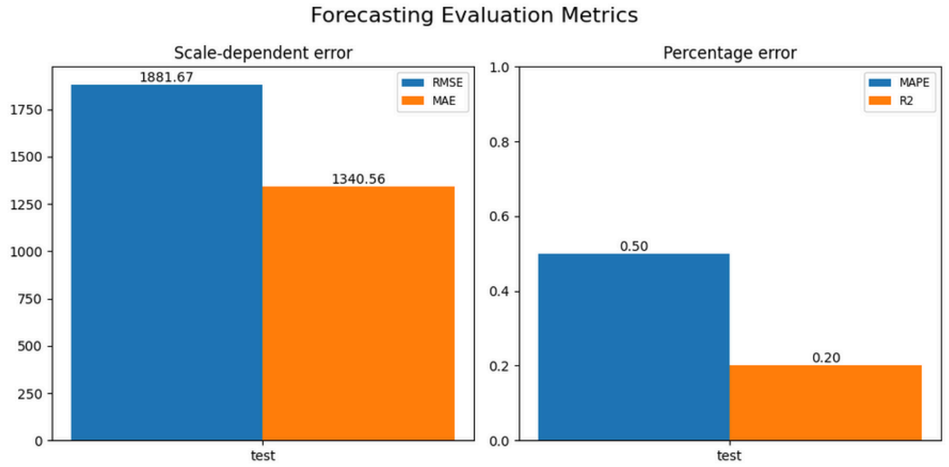
- **GRU Units: [32, 64, 96, 128, 160, 192, 224, 256]**
- **Extra Layer: True/False**
- **GRU Units for Second Layer: [16, 32, 48, 64, 80, 96, 112, 128]**
- **Learning Rates: [0.01, 0.001, 0.0001]**

The **best configuration** identified was:

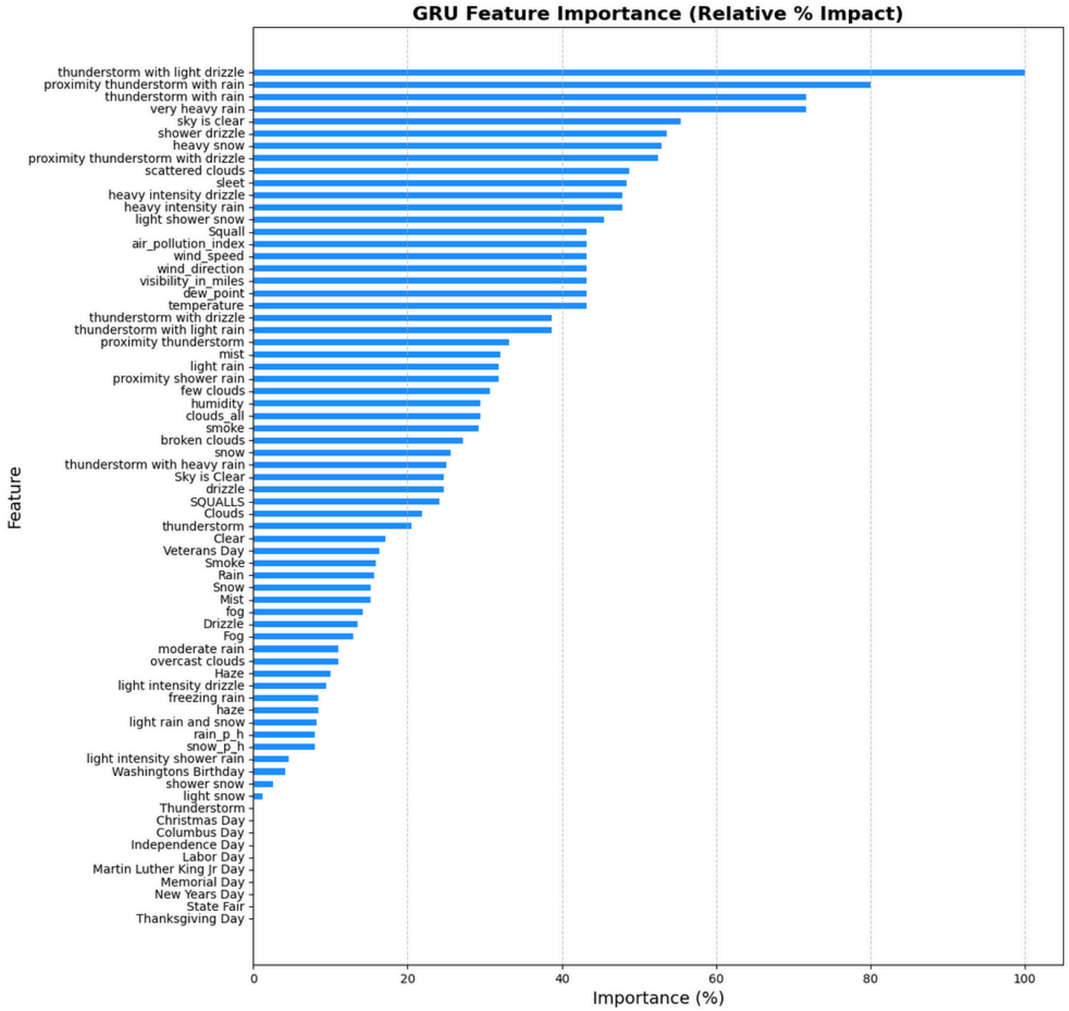
- **GRU Units: 192**
- **Extra Layer: True**
- **GRU Units for Second Layer: 16**
- **Learning Rates: 0.0001**

MODEL EVALUATION WITH THE BEST CONFIGURATION

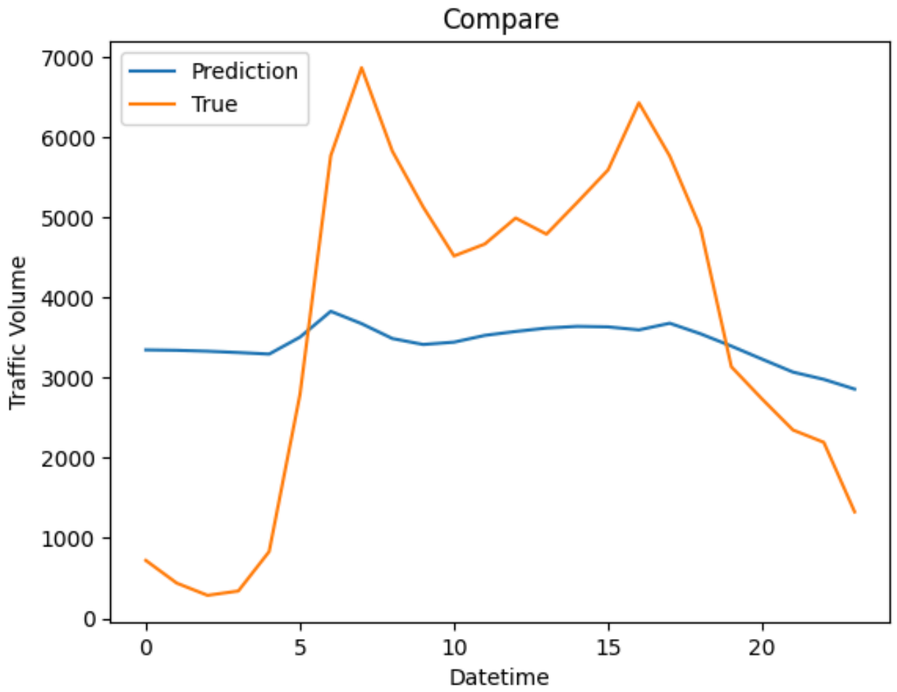


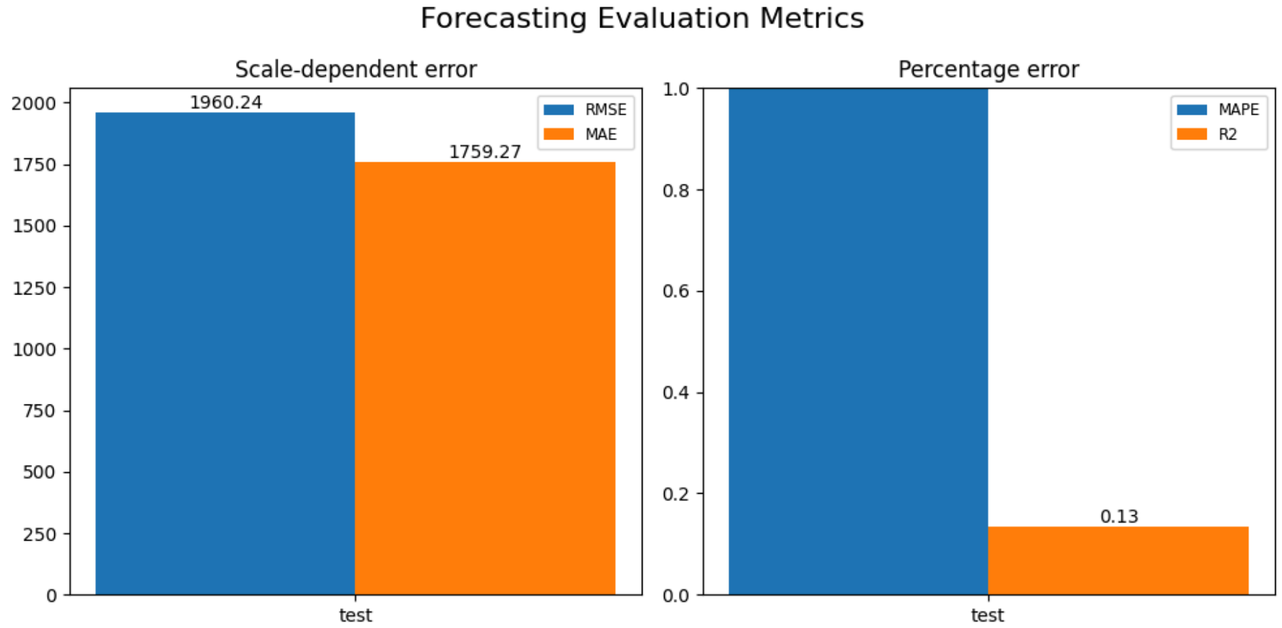


FEATURE IMPORTANCE AND SELECTION



FINAL MODEL TRAINING WITH SELECTED FEATURES





The final results indicated a decline in performance compared to the previous experiments, suggesting that the GRU model struggled to improve predictions despite feature selection efforts

TEMPORAL CONVOLUTIONAL NETWORK (TCN)

Type of deep learning model designed for sequential data processing. TCNs leverage causal convolutions with dilation, allowing them to capture long-range dependencies efficiently without suffering from vanishing gradient issues.

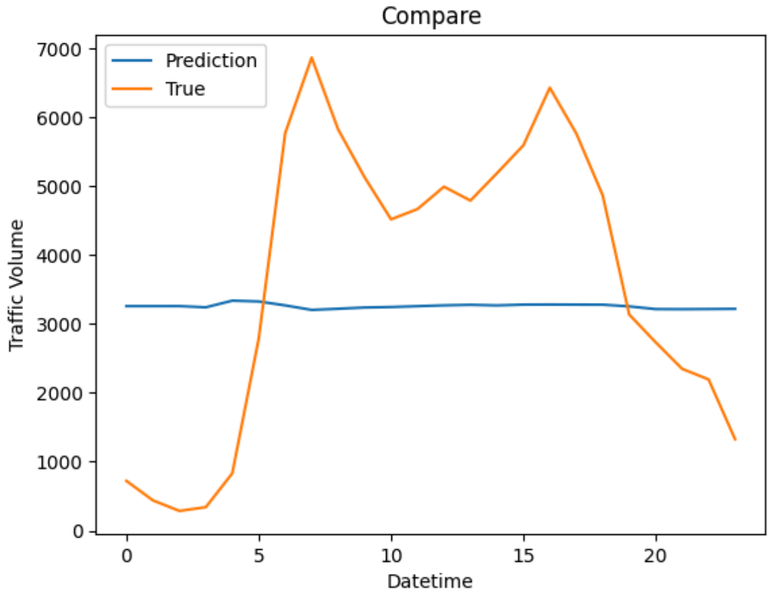
HYPERPARAMETER TUNING

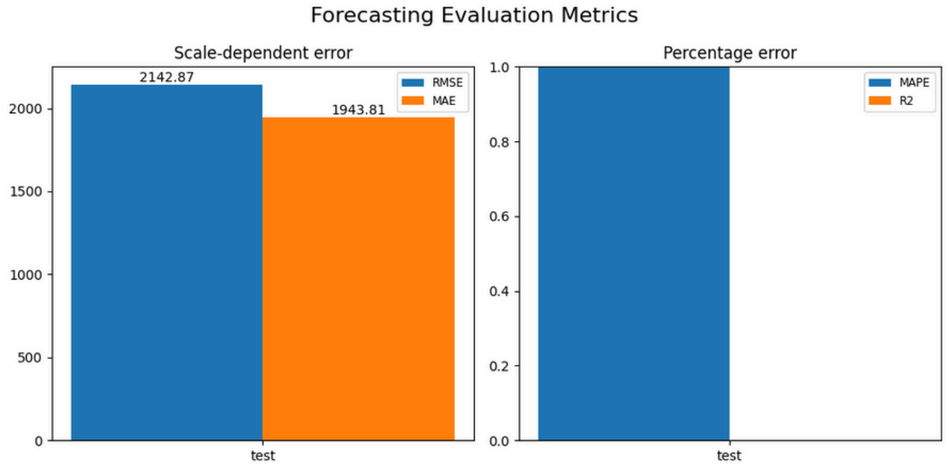
- **Number of Layers: [1, 2, 3, 4]**
- **Filters per Layer: [16, 32, 48, 64, 80, 96, 112, 128]**
- **Kernel Sizes: [2, 3, 5, 7]**
- **Dropout Rate: [0.2, 0.3, 0.4, 0.5]**
- **Activation Functions: ['relu', 'tanh', 'sigmoid', 'swish']**
- **Learning Rates: [1e-5, 1e-4, 1e-3, 1e-2]**

The **best configuration** identified was:

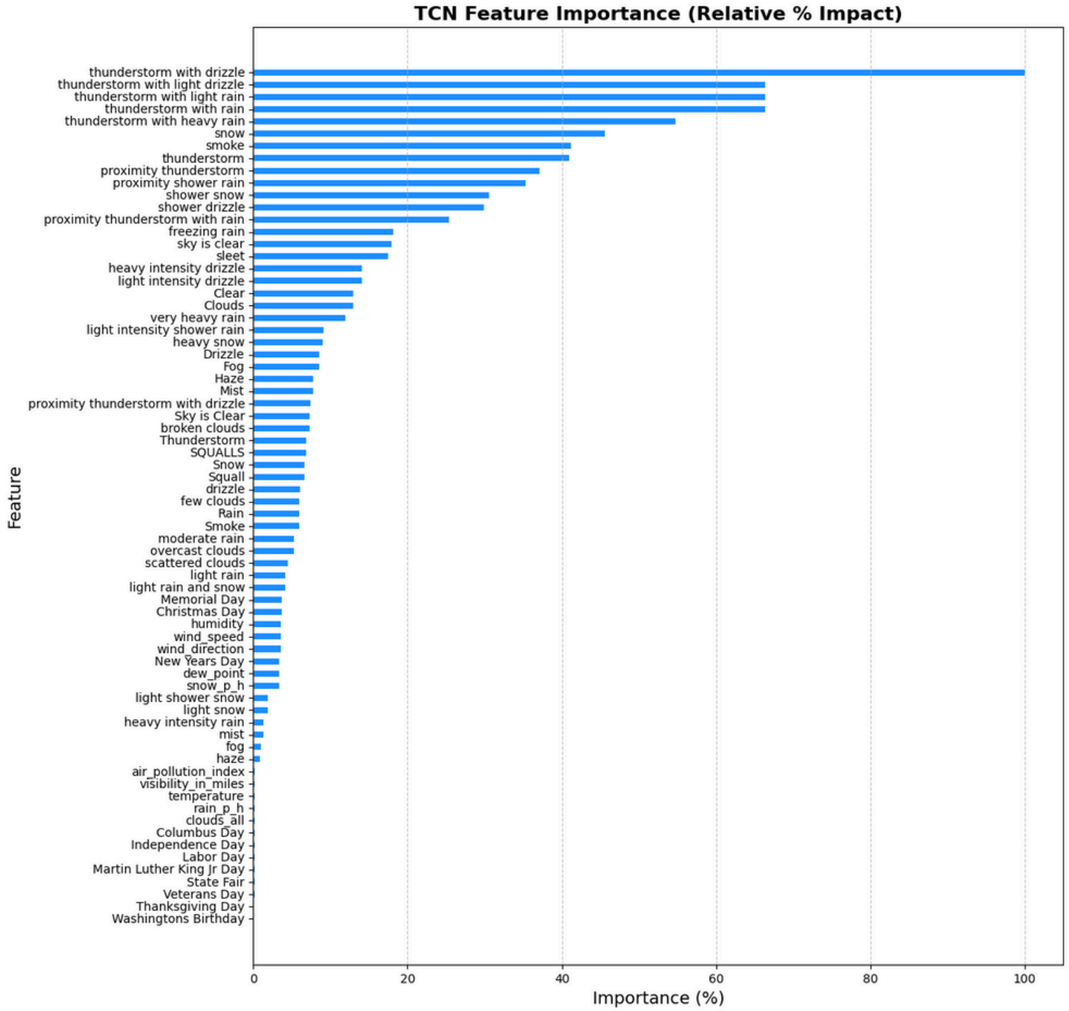
- **Number of Layers: 1**
- **Filters: 112**
- **Kernel Size: 2**
- **Dropout Rate: 0.3**
- **Activation Function: 'sigmoid'**
- **Learning Rate: 0.00086**

MODEL EVALUATION WITH THE BEST CONFIGURATION

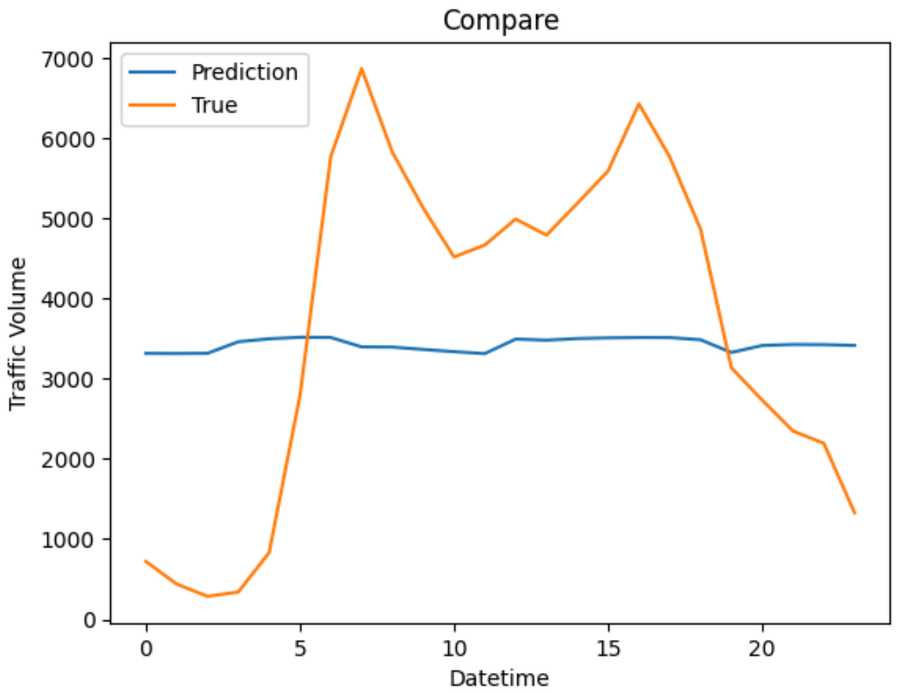


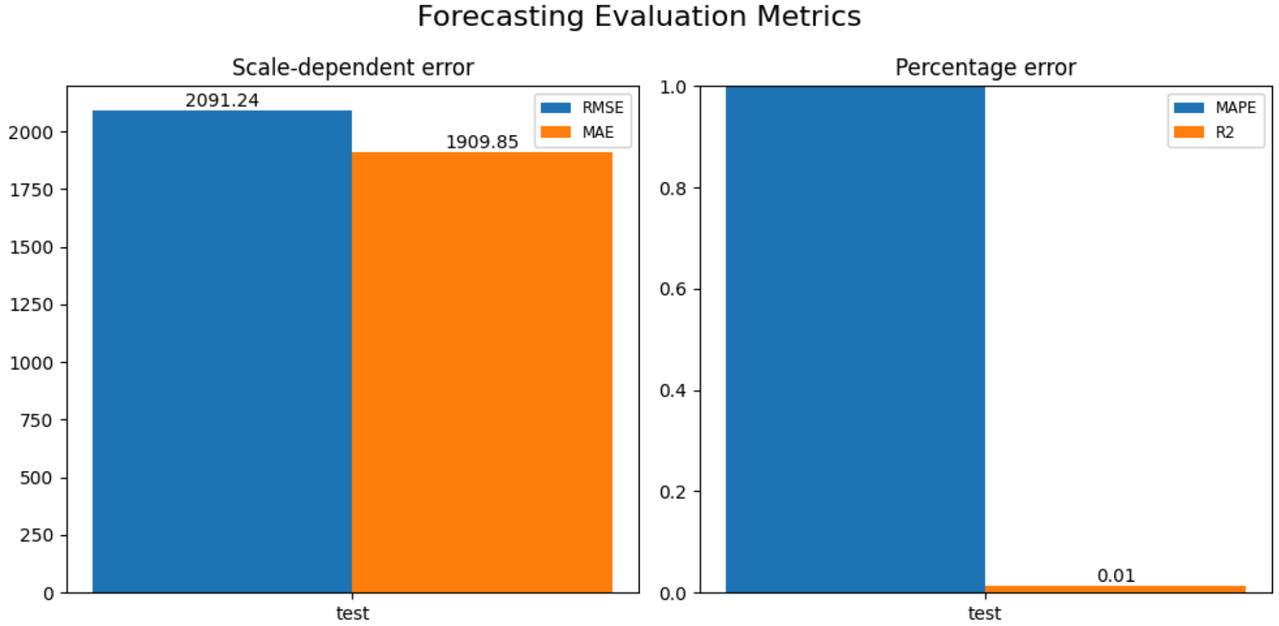


FEATURE IMPORTANCE AND SELECTION



FINAL MODEL TRAINING WITH SELECTED FEATURES





Upon evaluation, the R^2 value improved slightly, reaching 0.01. While this indicates a small gain in predictive performance, the model still struggles to capture significant variance in traffic volume.

EXTREME GRADIENT BOOSTING (XGBOOST)

Highly efficient and scalable machine learning algorithm based on gradient boosting. It has been widely used in various predictive modeling tasks due to its ability to handle missing values, prevent overfitting, and efficiently process large datasets

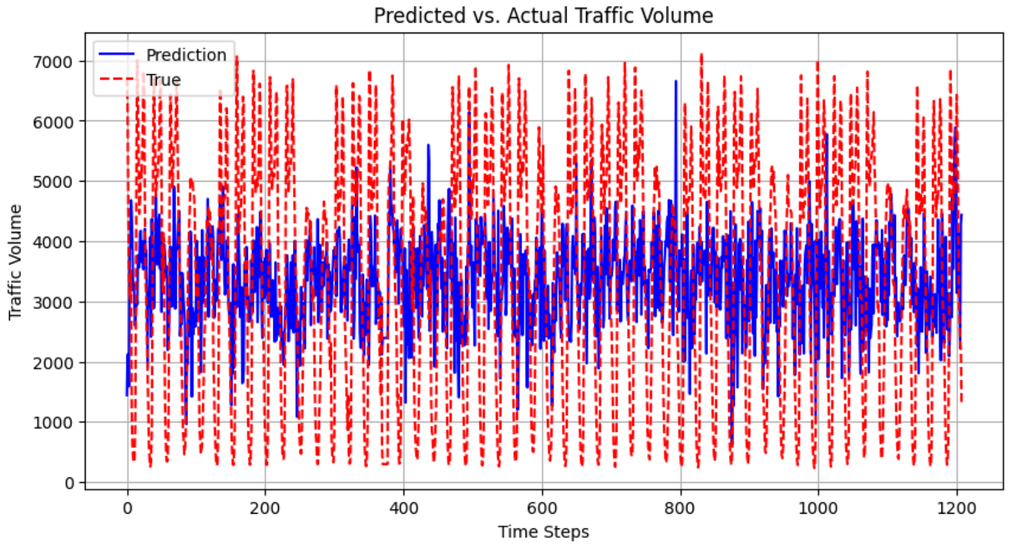
HYPERPARAMETER TUNING

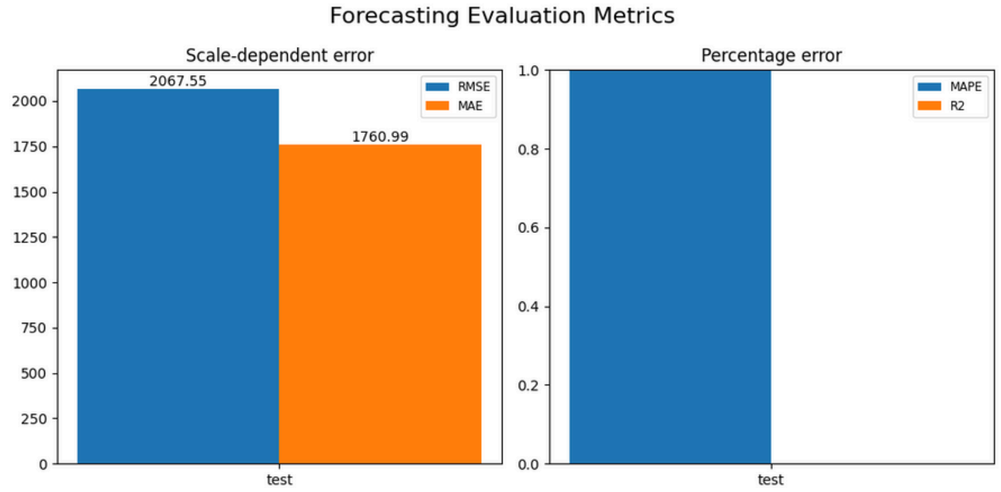
- **Number of Estimators: [50, 100, 200]**
- **Max Depth: [3, 6]**
- **Learning Rate: [0.01, 0.1, 0.3]**
- **Subsample: [0.7, 1.0]**
- **Colsample by Tree: [0.7, 1.0]**
- **Gamma: [0, 0.1]**
- **Min Child Weight: [1, 3]**

The **best configuration** identified was:

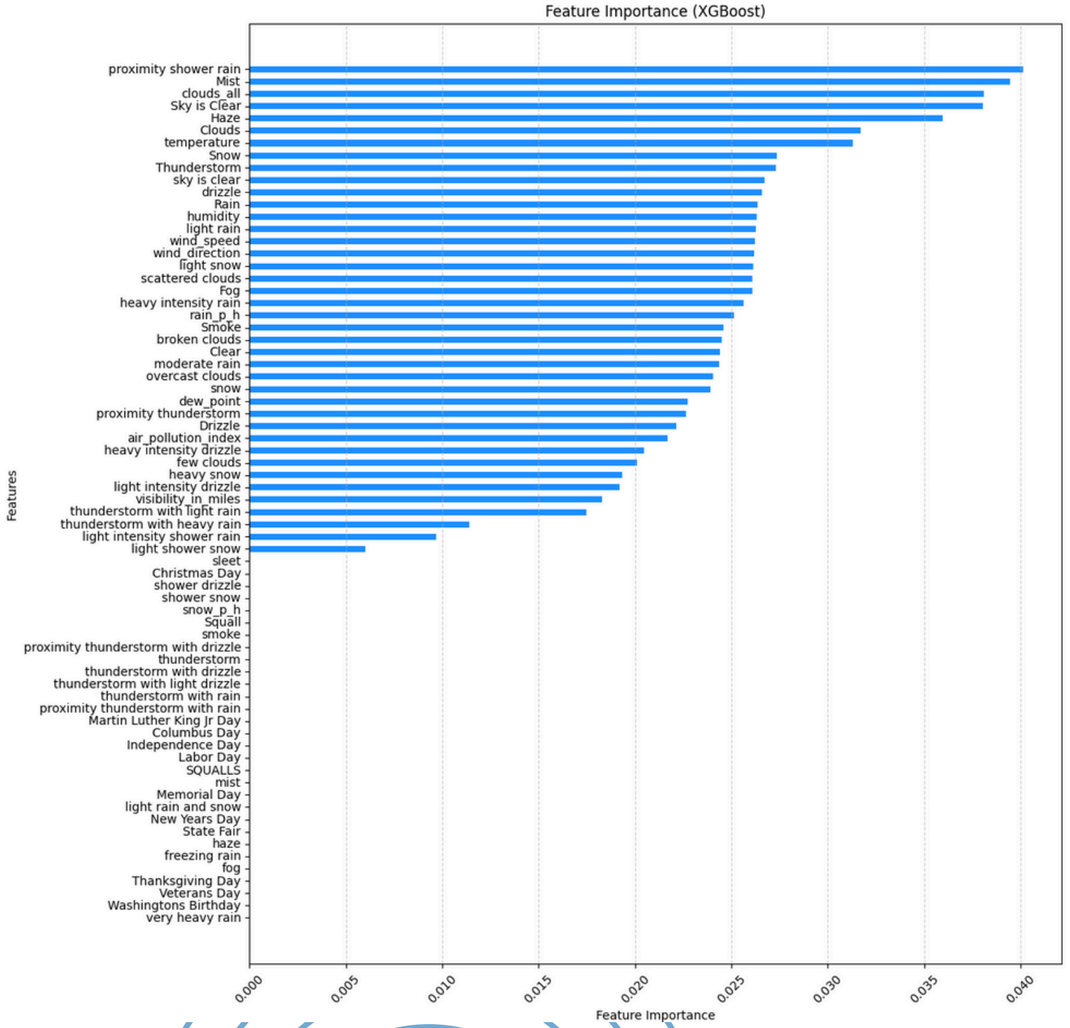
- **Number of Estimators: 200**
- **Max Depth: 6**
- **Learning Rate: 0.1**
- **Subsample: 0.7**
- **Colsample by Tree: 1.0**
- **Gamma: 0.1**
- **Min Child Weight: 3**

MODEL EVALUATION WITH THE BEST CONFIGURATION

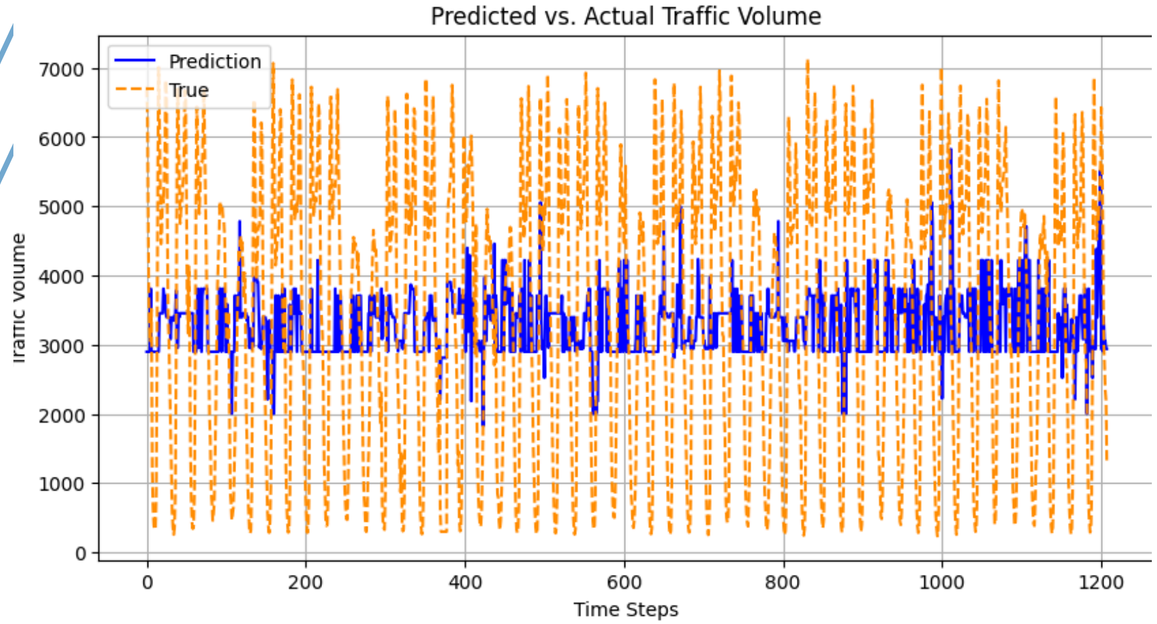


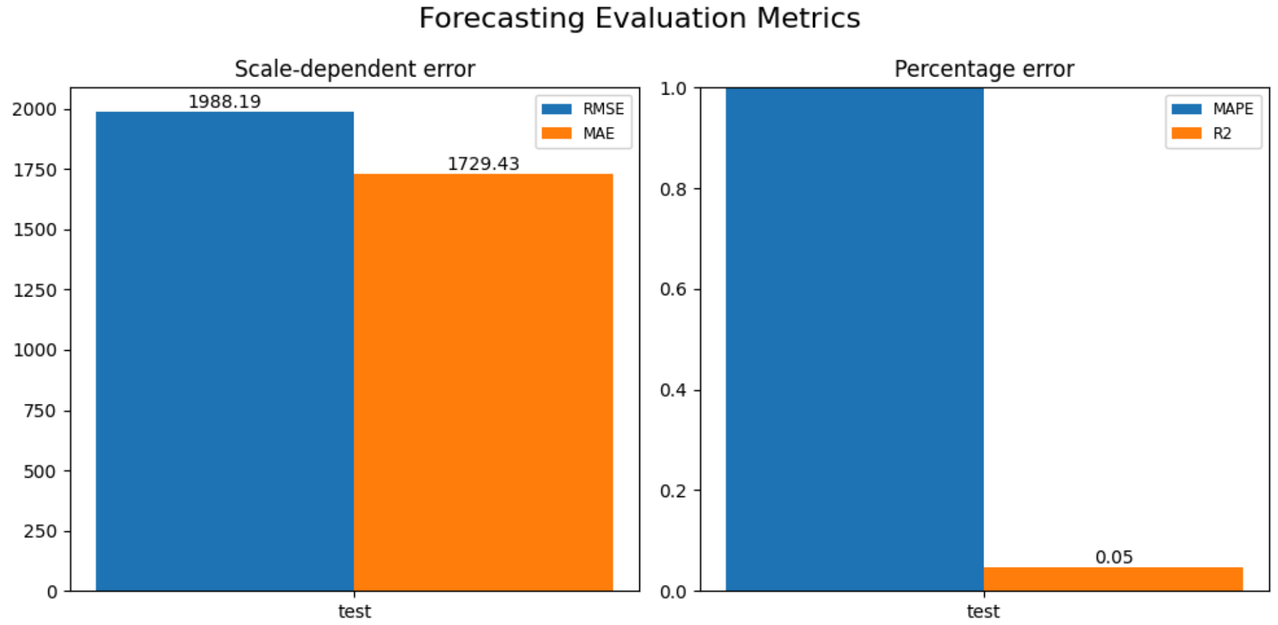


FEATURE IMPORTANCE AND SELECTION



FINAL MODEL TRAINING WITH SELECTED FEATURES





The performance metrics, including R^2, mean absolute error (MAE), and root mean squared error (RMSE), were computed

BEST PERFORMING MODEL SELECTION

Model	RMSE	MAE	MAPE	R^2
Long Short-Term Memory (LSTM)	1077.73	807.17	0.33	0.74
Gated Recurrent Uni (GRU)	1881.67	1340.56	0.50	0.20
Temporal Convolutional Network (TCN)	2091.24	1909.85	>1	0.01
Extreme Gradient Boosting (XGBosst)	1988.19	1729.43	>1	0.05

Based on these results, the **LSTM model emerged as the best-performing model**, achieving the highest R^2 value (0.74) and the lowest RMSE and MAE scores

ANALYSIS OF THE MODELS

Poorly Performing Models

- GRU
- TCN
- XGBOOST

Best Performing Model

- LSTM

THE ROLE OF THE DATASET IN POOR MODEL PERFORMANCE

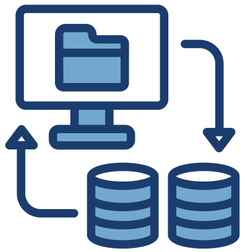
01. FEATURE QUALITY AND RELEVANCE

02. DATA NOISE AND INCONSISTENCIES

03. TEMPORAL RESOLUTION ISSUES

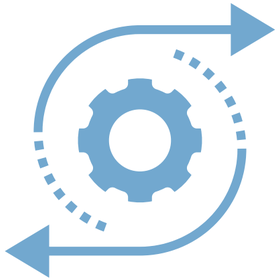
04. EXOGENOUS FACTORS NOT CONSIDERED

RESULTS DISCUSSION



Dataset enrichment

More refined feature selection process



Hybrid models

THANK YOU

Do you have any question?