

<u>Traffic Intelligence: Advanced Traffic Volume Estimation</u> <u>with Machine Learning</u>

<mark>MENTOR</mark>: <mark>SRI L.LAKSHMI NARAYANA</mark>

TEAM ID: <u>LTVIP2025TMID40409</u>

TEAM MEMBER:

2) SAPPA KIRAN MAI



MAIL ID: kiransappa5@gmail.com

MY TASKS:

TSK-338839:

Splitting the Dataset into Dependent and Independent variable

```
[24]: # Target variable
y = df['traffic_volume']

# Feature set (drop traffic_volume)
x = df.drop(columns=['traffic_volume'], axis=1)
```

TSK-338840:

Feature Scaling

```
[24]: # Target variable
      y = df['traffic_volume']
      # Feature set (drop traffic_volume)
      x = df.drop(columns=['traffic_volume'], axis=1)
[25]: import pandas as pd
      from sklearn.preprocessing import scale
      # Step 1: Separate target variable
      y = df['traffic_volume']
      # Step 2: Drop target from features
      x = df.drop(columns=['traffic_volume'], axis=1)
      # Step 3: Save column names before scaling
      names = x.columns
      # Step 4: Apply scaling (Z-score normalization)
      x = scale(x) # This returns a NumPy array
      # Step 5: Convert scaled data back to DataFrame with original column names
      x = pd.DataFrame(x, columns=names)
      # Step 6: Preview the result
      x.head()
       holiday
[25]:
                      temp
                                rain
                                           snow weather
                                                               year
                                                                        month \
          0.0 0.530485 -0.007463 -0.027235 -0.566452 -1.855294 1.02758
            0.0 0.611467 -0.007463 -0.027235 -0.566452 -1.855294 1.02758
            0.0 0.627964 -0.007463 -0.027235 -0.566452 -1.855294 1.02758
            0.0 0.669205 -0.007463 -0.027235 -0.566452 -1.855294
            0.0 0.744939 -0.007463 -0.027235 -0.566452 -1.855294 1.02758
                     hours minutes seconds
             day
      0 -1.574903 -0.345548 0.0
                                      0.0
     1 -1.574903 -0.201459 0.0
2 -1.574903 -0.057371 0.0
3 -1.574903 0.086718 0.0
4 -1.574903 0.230807 0.0
                                         0.0
                                         0.0
```

TSK-338841:

Splitting the data into Train and Test

```
[26]: from sklearn.model_selection import train_test_split
      # Split the dataset into training and testing sets
      x_train, x_test, y_train, y_test = train_test_split(
         x, y, test_size=0.2, random_state=0
```

0.0 0.0