

# UPSKILLS DATA SCIENCE AND MACHINELEARNING INTERNSHIP

## WEEK - 3

### **Team member :-**

Harshwardhansinh Chauhan

Pranav Khot

Pathan MoinKhan

Kshitij Solanki

**SIGMA INSTITUTE OF ENGINEERING**

**UPSKILLS DATA SCIENCE AND MACHINE LEARNING INTERNSHIP****WEEK - 3**

I might want to provide you with an advancement report for my third week in the Upskills UCT Machine Learning and Data Science Internship. The accompanying focuses feature the critical parts of my exercises and encounters

**Project Overview:**

The smart City Traffic Example ML project plans to break down and foresee traffic designs in a shrewd city climate utilizing AI strategies. By getting it and foreseeing traffic designs, we can advance traffic stream, further develop transportation effectiveness, and improve in general metropolitan versatility. This report gives an outline of the issue explanation and talks about potential calculations that can be utilized in the undertaking

**Problem Statement:**

You are working with the public authority to change your city into a savvy city. The vision is to change over it into a computerized and smart city to work on the proficiency of administrations for the residents. One of the issues looked by the public authority is traffic. You are an information researcher attempting to deal with the traffic of the city better and to give input on foundation wanting to what's to come. The public authority needs to execute a powerful traffic framework for the city by being ready for traffic tops. They need to comprehend the traffic examples of the four intersections of the city.

Traffic designs on vacations, as well as on different events during the year, vary from ordinary working days. This means a lot to consider for your estimating. Presently we done the fundamental investigation of the PS and our dataset and assessed the accompanying realities about the given dataset and what we need to submit in the last undertaking report. Along these lines, we will work likewise.

## Data Dictionary

Variable	Description
ID	Unique ID
DateTime	Hourly Datetime Variable
Junction	Junction Type
Vehicles	Number of Vehicles (Target)

### sample\_submission.csv

Column Name	Description
ID	Unique ID
Vehicles	Number of Vehicles (Target)

### Progress Of The Week:

The objective of this venture is to foster a brilliant city traffic expectation framework utilizing AI procedures. In this report, we will cover the headway made during the third seven day stretch of the venture, which incorporates information stacking, exploratory information examination (EDA), and highlight designing

## **2. Data Loading**

During the third week, we zeroed in on stacking the vital information for our shrewd city traffic forecast framework. The information comprises of verifiable traffic records gathered from different sensors and sources in the city. We got the information from the city's transportation division, which gave us an organized dataset in CSV design.

## **3. Exploratory Data Analysis (EDA):**

In the wake of stacking the information, we performed exploratory information examination to acquire experiences into the dataset and grasp its qualities. The EDA cycle included the accompanying advances:

**a) Data Summary :-** We produced rundown insights of the dataset, including measures of focal inclination, scattering, and dispersions of various highlights. This permitted us to grasp the reach and changeability of the information

**b) Data Visualizaation :-** We made different representations, for example, histograms, box plots, disperse plots, and time series plots to investigate the connections between various factors. These perceptions assisted us with recognizing examples, exceptions, and possible relationships inside the information

**c) Correlation Analysis :-** We determined the connection framework to evaluate the connections between various highlights. This examination assisted us with recognizing exceptionally corresponded factors, which might actually be utilized as instructive elements for our traffic expectation model.

**d) Data Quality Check :-** We performed information quality checks to recognize missing qualities, exceptions, and irregularities in the dataset. We utilized methods, for example, checking for invalid qualities, information attribution, and eliminating or revising exceptions to guarantee the information's uprightness and dependability

**4. Feature Engineering :-** Highlight designing is an essential move toward AI projects as it includes changing crude information into significant highlights that can work on the prescient execution of the models. In this stage, we played out the accompanying component designing undertakings

a) **Temporal Features:-** We extricated helpful transient highlights from the timestamp section, like hour of the day, day of the week, month, and year. These highlights can catch the periodicity and time-subordinate examples in rush hour gridlock information.

b) **Categorical Encoding :-** We encoded clear cut factors utilizing methods like onehot encoding or name encoding, contingent upon the idea of the factors. This transformation permitted us to address clear cut information mathematically, making it appropriate for AI calculations.

c) **Feature Scaling:** We applied include scaling procedures like normalization or standardization to guarantee that every one of the highlights were on a comparative scale. This step is significant on the grounds that many AI calculations perform better when the elements are standardized

d) **Feature Selection:-** We utilized different component choice strategies, like relationship investigation, measurable tests, and space information, to recognize the most significant highlights for our traffic expectation model. This cycle assisted us with lessening the dimensionality of the dataset and center around the most educational factors

## **5. Conclusion:**

During the third week stretch of our smart city traffic forecast project, we effectively stacked the information, performed exploratory information investigation (EDA), and directed include designing. These means permitted us to acquire experiences into the dataset, distinguish examples, and designer significant elements for our MI models. In the following stage, we will continue with model choice and preparing in view of this current week progress.