

NAME : MUHAMMAD RAEES (19122041)

SECTION : 6<sup>TH</sup> A

NAME : MUHAMMAD HUSSAIN (19122010)

SUBJECT : DATA MINING

GITHUB LINK : <https://github.com/Raees161/TERM-PROJECT>

## FINAL REPORT

### INSTALLATION OF TRAFFIC LIGHTS TO PREVENT ACCIDENT

#### OBJECTIVE:

The primary objective of this analysis is to predict the possibility of traffic light installations based on historical data. Leveraging the Naive Bayes classification algorithm, the analysis aims to categorize locations as either "possible" or "not possible" for traffic light installations. The focus is on achieving an accurate prediction to inform decision-making in urban planning and traffic management.

#### INTRODUCTION AND BACKGROUND OF THE PROBLEM:

Traffic light installations are critical for managing traffic flow and ensuring road safety. Understanding historical trends in traffic light installations provides valuable insights into the evolving needs of urban and suburban areas. This analysis aims to contribute to efficient traffic management by identifying patterns in installations and providing recommendations for strategic planning.

#### DATA COLLECTION:

The dataset used for this analysis encompasses historical records of traffic light installations, including information on location, Traffic flow, and relevant features. Data was sourced from municipal records, traffic management authorities, and historical infrastructure databases, ensuring a comprehensive representation of urban and suburban areas. For Example, we take data from Google, Google dataset search, Kaggle datasets, and construct by self.

#### DATA PREPROCESSING:

##### 1. DATA CLEANING:

Duplicate and irrelevant entries were removed to enhance data accuracy.

##### 2. HANDLING MISSING VALUES:

Missing data was addressed using appropriate methods such as imputation or removal.

##### 3. FEATURE SELECTION:

Key features, including location, date, and contextual factors influencing traffic light installations, were identified.

## **MODELING AND EVALUATION:**

### **1. NAIVE BAYES OPERATOR:**

Utilized the Naive Bayes classification algorithm in RapidMiner to categorize locations as "possible" or "not possible" for traffic light installations.

### **2. VISUALIZATION:**

Utilized RapidMiner's visualization tools to illustrate the distribution of predicted and actual outcomes, offering a visual representation of the model's performance.

### **3. RESULT:**

Accuracy of 77.99% in predicting the possibility of traffic light installations.

Successful categorization of **ACCIDENT** as "possible" (124 rows) and "not possible" (35 rows) based on historical data.

## **CONCLUSION:**

### **1. PREDICTIVE POWER:**

The Naive Bayes model exhibits reasonable predictive power, achieving an accuracy of 77.99%. This suggests its effectiveness in categorizing locations based on historical patterns.

### **2. TARGETED DECISION MAKING:**

The model's ability to distinguish between "possible" and "not possible" scenarios can inform targeted decision-making for future traffic light installations.

This report emphasizes the successful application of the Naive Bayes algorithm in predicting the possibility of traffic light installations based on historical data. The achieved accuracy and categorization provide valuable insights for decision-makers, enabling strategic planning and enhancing traffic management efficiency.