Student Name: A.DHANISH RAZA

Register Number: 410623104002

Institution: DHAANISH AHMED COLLEGE OF ENGINEERING

Department: COMPUTERSCIENCEANDENGINEERING

Date of Submission: 30/04/2025

Project Title:

Enhancing Road Safety with AI-Driven Traffic Accident Analysis and Prediction

1. Problem Statement

Road accidents are one of the leading causes of death and injury worldwide. Traditional accident prevention methods rely heavily on human judgment and delayed response. By leveraging Artificial Intelligence (AI) to analyze accident data and predict high-risk zones or times, authorities can take preventive measures in advance. This project uses AI and machine learning to analyze traffic accident datasets and predict accident probability based on historical patterns.

---

2. Objectives of the Project

- Analyze historical traffic accident data to find patterns and trends.
- Develop a predictive model to identify accident-prone areas or conditions.
- Provide actionable insights for road safety improvements.
- Visualize accident hotspots and prediction outcomes via dashboards.

3. Scope of the Project

Use of machine learning for predicting accident likelihood.

- Focus on structured accident datasets (date, time, weather, location, etc.).
- Visual mapping of accident zones (using libraries like Folium/Plotly).
- Limitations: No real-time GPS or sensor data integration (static analysis only).

---

4. Data Sources

- Dataset: Traffic Accident Data (e.g., UK, US, or Indian datasets)
- Source:
- Kaggle (e.g., US Accidents)
- Government open data portals (e.g., data.gov.in)
- Type: Public, Static dataset

---

5. High-Level Methodology

Data Collection

- Download accident data from open government or Kaggle sources.
- Load into environment using pandas.

Data Cleaning

- Handle missing values, incorrect timestamps, location mismatches.
- Normalize features (weather conditions, time of day, etc.).

Exploratory Data Analysis (EDA)

- Analyze accident frequency by time, location, weather, day of week.
- Visualize accident trends using bar charts, heat maps, and maps.

Feature Engineering

Convert date/time into useful features (e.g., peak hour, weekday/weekend).

- Create binary labels (accident/no accident).

Model Building

- Use classification models:
- Logistic Regression
- Decision Trees
- Random Forest
- Gradient Boosting

Model Evaluation

- Metrics: Accuracy, Precision, Recall, F1-Score, ROC-AUC.
- Use k-fold cross-validation for model stability.

Visualization & Interpretation

- Map accident-prone zones using Folium/Plotly.
- Display high-risk times/conditions.

Deployment

- Deploy using Streamlit: enter time, weather, location to get accident risk prediction.

---

6. Tools and Technologies

Language: Python

IDE/Notebook: Google Colab / Jupyter

Libraries: pandas, matplotlib, seaborn, scikit-learn, xgboost, folium, plotly

- NLP (if analyzing accident descriptions): nltk, spaCy

Deployment Tools: Streamlit or Flask

7. Team Members and Roles

Member 1: A.Dhanish Raza

Role: Data Collection and Preprocessing

Member 2: Luthfi Bassam

Role: Exploratory Data Analysis and Visualization

Member 3: G.Chandru

Role: Feature Engineering and Model Building

Member 4: Anudarsh Sunil

Role: Model Evaluation and Optimization

Member 5: P.D.Balaji

Role: Dashboard, Documentation, and Deployment