Phase-1 Submission

Student Name: Divagaran M

Register Number: 511523205013

Institution: P. T. Lee Chengalvaraya Naicker College of Engineering and Technology

Department: B.Tech(Information Technology)

Date of Submission: 2025-04-26

1. Problem Statement

Traffic accidents pose a significant threat to public safety, leading to loss of lives, property damage, and economic burdens. The lack of effective data-driven methods to anticipate and mitigate these events calls for innovative approaches. This project aims to enhance road safety by leveraging AI for comprehensive traffic accident analysis and predictive modeling to identify potential high-risk scenarios and prevent accidents.

2. Objectives of the Project

- Analyze historical traffic accident data to discover contributing factors.
- Build predictive models to forecast accident-prone areas and timings.
- Develop visualizations for actionable insights.
- Suggest safety interventions based on predictive outcomes.

3. Scope of the Project

The project focuses on analyzing datasets containing traffic accident records. It includes data preprocessing, EDA, and predictive modeling. Limitations include dependence on data availability and not deploying a real-time alert system in this phase.

4. Data Sources

Datasets will be obtained from public sources such as Kaggle, UCI Machine Learning Repository, or government open data portals. These datasets are static and contain past accident records.

5. High-Level Methodology

- ***Data Collection***: Download static datasets from public sources.
- ***Data Cleaning***: Handle missing values, remove duplicates, and format standardization.
- ***Exploratory Data Analysis***: Use visualizations like heatmaps, bar plots, and line charts to find patterns.
- ***Feature Engineering***: Create features like time of day, weather conditions, and accident density zones.
- ***Model Building***: Use models like Random Forest, Logistic Regression, and XGBoost.
- ***Model Evaluation***: Evaluate with accuracy, precision, recall, and F1-score.
- ***Visualization & Interpretation***: Present insights via graphs and dashboards.
- ***Deployment***: Develop a simple dashboard using Streamlit.

6. Tools and Technologies

- ***Programming Language***: Python
- ***Notebook/IDE***: Jupyter Notebook, Google Colab
- ***Libraries***: pandas, numpy, matplotlib, seaborn, scikit-learn, xgboost
- ***Optional Tools for Deployment***: Streamlit

7. Team Members and Roles

-***Vicknesh V***: Data Collection and Cleaning
- ***Srinivasan N***: EDA and Feature Engineering
- ***Ashwin G***: Model Building and Evaluation
- ***Divagaran M***: Report Writing and Presentation