**Phase-1**

**Student Name:** DINESH R

**Register Number:** 212923104026

**Institution:** ST.Joseph College Of Engineering

**Department:** BE - Computer Science and Engineering

**Date of Submission:** 26.04.2025

## **PROJECT TITLE:**

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

## **1. Problem Statement**

Road accidents pose a serious threat to public safety, resulting in significant loss of life and economic damage. Traditional methods of traffic monitoring and accident prevention are often reactive rather than proactive. There is a critical need for intelligent systems capable of analyzing traffic patterns, predicting accident risks, and providing actionable insights to prevent accidents before they occur. This project aims to leverage AI-driven data analysis and machine learning models to forecast potential accident hotspots and contribute to safer transportation environments.

## **2. Objectives of the Project**

* Analyze historical traffic accident data to identify patterns and risk factors.
* Build predictive models capable of forecasting accident-prone zones.
* Develop a real-time alert system to inform authorities and drivers about potential risks.
* Provide actionable recommendations to improve traffic safety measures.
* Build a user-friendly dashboard for visualizing accident trends and predictions.

## **3. Scope of the Project**

**Features:**

* In-depth analysis of traffic-related datasets (weather, time, road conditions, vehicle types).
* Application of AI techniques like classification, clustering, and time-series prediction.
* Identification of high-risk areas and accident hotspots.

**Limitations:**

* Predictions rely heavily on the availability and quality of traffic and accident datasets.
* The model's performance may vary across different geographic regions.

**Constraints:**

* Only publicly available or government-published traffic accident datasets will be used.
* Focus will be on prediction and analysis; implementation of physical interventions (e.g., road repairs) is outside the project's scope.

## **4. Data Sources**

* **Dataset**: Road Accident Data (e.g., National Highway Traffic Safety Administration, Kaggle public datasets)
* **Sources**:
  + Kaggle - US Accidents (3.0 million records)
  + Government traffic accident reports and open datasets.
* **Type**: Public, time-series, and geo-spatial data.

## **5. High-Level Methodology**

**Data Collection:**

* Download accident datasets from public sources.

**Data Cleaning:**

* Handle missing or inconsistent data entries.
* Normalize weather and location features.

**Exploratory Data Analysis (EDA):**

* Visualize accident frequency based on time, location, weather, and road conditions.
* Identify correlations between factors and accident occurrences.

**Feature Engineering:**

* Create new features like "peak traffic hours", "adverse weather indicator", etc.
* Use geospatial features (latitude, longitude clustering).

**Model Building:**

* Models: Random Forest, XGBoost, Decision Trees, LSTM for time-series accident prediction.
* Justification: Ensemble models and sequence models help capture complex patterns and trends.

**Model Evaluation:**

* Metrics: Accuracy, Precision, Recall, F1-Score, AUC-ROC for classification tasks.
* Validation Strategy: Stratified K-Fold Cross Validation.

**Visualization & Interpretation:**

* Accident heatmaps.
* Risk-level classification maps.

**Deployment:**

* Build a dashboard using Streamlit to visualize accident hotspots and risk predictions in real time.

## **6. Tools and Technologies**

* **Programming Language**: Python
* **Notebook/IDE**: Jupyter Notebook, Google Colab
* **Libraries**: pandas, numpy, scikit-learn, matplotlib, seaborn, xgboost, folium (for maps), streamlit
* **Optional Deployment Tools**: Streamlit or Flask for web deployment

**7. Team Members and Roles**

1. DEEPAK S – DATA ENGINEER

2. HARISH R – MACHINE LEARNING ENGINEER

3.DINESH R – FULL STACK DEVELOPER