Enhancing road safety with AI driven traffic accident analysis and prediction

Student name: PRIYADHARSHINI R
Register number: 732323106040
Institution: SSM college of Engineering
Department: BE(ECE)-2Year
Date of submission: 23.04.2025

1) Problem Statement:

Road accidents remain a significant public safety concern worldwide, leading to loss of life, injuries, and economic damage. Traditional methods of traffic accident analysis often lack real-time capabilities and fail to leverage the vast amounts of data generated by modern transportation systems. There is a critical need for intelligent systems that can not only analyze traffic accident data but also predict potential accident hotspots to implement proactive safety measures.

This project aims to enhance road safety by utilizing Artificial Intelligence (AI) for traffic accident analysis and prediction. By leveraging machine learning and data analytics, the system will identify patterns and factors contributing to accidents, enabling timely interventions and informed decision-making by authorities and drivers alike.

2) Objectives of the project:

1. To collect and preprocess traffic accident data
Gather relevant historical traffic accident datasets from reliable sources and clean the data for further analysis.

2. To identify key factors contributing to traffic accidents
Use data analytics to determine patterns, trends, and common causes of road accidents.

3. To develop AI models for accident prediction
Apply machine learning algorithms to predict the likelihood of accidents based on factors like time, location, weather, and traffic conditions.
4. To visualize accident-prone zones
Create interactive visualizations or heatmaps to highlight high-risk areas for better decision-making by authorities and the public.
5. To evaluate the performance of the AI system

Assess the accuracy and effectiveness of the AI models in predicting accidents using suitable evaluation metrics.
6. To propose data-driven recommendations for road safety improvements
Suggest strategies for minimizing accident risks based on AI-driven insights.

3) scope of the project:

1. Data Collection and Processing:
The project will focus on gathering historical traffic accident data from open datasets, government records, or traffic monitoring systems.

2. AI Model Development:
Implementation of machine learning algorithms to analyze historical data and predict the probability of traffic accidents based on various input factors.
3. Geospatial and Temporal Analysis:
Identification of accident hotspots by analyzing spatial (location-based) and temporal (time-based) patterns.
4. Visualization Tools:
Development of dashboards or heatmaps to represent accident-prone areas and prediction outcomes in an intuitive and accessible way.
5. Predictive Insights for Authorities:
Providing insights to traffic management authorities and city planners to implement proactive safety measures.
6. Limitations:
The project will not involve real-time traffic monitoring or deployment on physical infrastructure. It is focused on data-driven prediction and analysis.

4) data source:

1. Kaggle Datasets
https://www.kaggle.com
Examples:
US Accident Data (2016–Present)
Road Accident Data Analysis
Traffic Violations and Weather-related accidents
2. Open Government Data Portals
India: https://data.gov.in
UK: https://data.gov.uk
USA: https://www.transportation.gov/data
Provide accident records, vehicle data, and road conditions.
3. City and Regional Transport Departments

Many cities provide accident reports and traffic logs through their local transport authority websites.
4. World Health Organization (WHO)
https://www.who.int
Offers global traffic accident statistics and safety reports.
5. National Highway Traffic Safety Administration (NHTSA)
https://www.nhtsa.gov/research-data
Offers in-depth crash data for analysis in the U.S.
6. OpenStreetMap (OSM) and Google Maps APIs
For geolocation data and mapping accident-prone areas.

5)High - level methodology:

1. Data Collection
Gather traffic accident data from open datasets, government portals, or APIs.
Collect supporting data like weather conditions, traffic volume, road types, and time stamps.
2. Data Preprocessing
Clean and normalize the data (handle missing values, remove duplicates).
Perform feature engineering (e.g., converting time into day/night, categorizing weather types, etc.).
3. Exploratory Data Analysis (EDA)
Analyze patterns and correlations between accident causes and influencing factors.
Visualize accident trends using charts and heatmaps.
4. Model Building
Train machine learning models such as Logistic Regression, Random Forest, or Neural Networks to predict accident likelihood.
Use classification or regression based on the problem framing (e.g., accident risk score or binary accident risk).
5. Model Evaluation
Evaluate models using accuracy, precision, recall, F1-score, or ROC-AUC.
Perform cross-validation to improve model robustness.
6. Visualization and Reporting
Generate accident heatmaps and dashboards to display predictions and patterns.
Provide insights to stakeholders for decision-making.
7. Recommendations
Suggest road safety improvements based on insights (e.g., enhanced signage, speed limits in high-risk areas).

6) Tools and technologies:

1. Programming Languages
Python – For data analysis, machine learning, and visualization.
R (optional) – For statistical modeling and data visualization.
2. Data Analysis & Machine Learning Libraries
Pandas – Data manipulation and analysis.
NumPy – Numerical computations.
Scikit-learn – Machine learning models and evaluation.
TensorFlow / Keras / PyTorch – For deep learning models (if needed).
XGBoost / LightGBM – For efficient gradient boosting.
3. Data Visualization Tools
Matplotlib / Seaborn – Plotting and statistical visualization.
Plotly / Dash – Interactive dashboards.
Folium / Geopandas – Mapping and geospatial data visualization.
4. Development Environment
Jupyter Notebook / Google Colab – For interactive coding and documentation.
VS Code / PyCharm – Full-featured IDEs.
5. Data Sources & APIs
Kaggle / Open Government Data – For accident datasets.
Google Maps API / OpenStreetMap – For location-based data.
6. Database (if applicable)
SQLite / MySQL / MongoDB – For storing and querying large datasets.
7. Version Control
Git / GitHub – For version tracking and collaboration.

7) Team members and Roles:

1. Data Collection and Preprocessing [PRAKASH P]
2. AI Model Development [PAVAN V]
3.System Integration and Visualization:[PERMALINGAM P]
4. Evaluation and Real-world Application[PRIYADHARSHINI R]