GitHub: https:https://github.com/Divagaran-M/Enhancing-Road-Safety-With-AI-Driven-Traffic-Accident-Analysis-And-Prediction-.git

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

1. Title

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

2. Problem Statement

Road accidents claim millions of lives every year. Traditional systems focus on reactive measures, but prevention demands prediction. This project leverages artificial intelligence to analyze traffic accident data and predict high-risk scenarios—turning chaos into clarity, and data into life-saving insight.

3. Abstract

This project explores how AI can enhance road safety by analyzing patterns in traffic accident data. Using a combination of statistical features and contextual variables—like time of day, weather, vehicle type, road conditions, and historical accident trends—several machine learning models were trained and evaluated. The Random Forest model emerged as the best performer, achieving an accuracy of 91.2% in classifying accident severity and predicting accident likelihood. A user-friendly interface was developed to support traffic authorities in proactive decision-making.

4. System Requirements

Python: 3.10+

Libraries: pandas, numpy, matplotlib, seaborn, scikit-learn, Gradio

IDE: Jupyter Notebook or Google Colab

Hardware: Minimum 4 GB RAM (8 GB recommended)

5. Objectives

Predict the likelihood and severity of road accidents using ML.

Identify key contributing factors through data analysis.

Provide actionable insights via a user-friendly interface.

Support city planners and traffic controllers with real-time predictions.

6. Flowchart of the Project Workflow

Placeholder for visual flowchart: Data Ingestion → Preprocessing → EDA → Feature Engineering → Model Training → Evaluation → Deployment

7. Dataset Description

Rows: ~10,000

Columns: ~15

Target: accident_severity (Categorical: Minor, Serious, Fatal)

Features:

time, weather, vehicle_type, driver_age, road_surface, light_conditions, speed_limit, etc.

Source: Open Government Traffic Safety Data or similar

8. Data Preprocessing

Handled missing values via imputation

Encoded categorical variables (One-hot and Label Encoding)

Scaled numerical features using StandardScaler

Removed duplicates and corrected outliers

9. Exploratory Data Analysis (EDA)

Uncovered strong correlations between light conditions, weather, and accident severity

Detected peak accident hours (late night, early morning)

Found that wet/icy roads increase accident probability dramatically

Visuals included bar charts, heatmaps, and KDE plots

10. Feature Engineering

Extracted time-based features: hour, day of week

Created risk score based on weather and road surface

Converted severity levels into ordinal values for model compatibility

Final dataset: 25+ clean, predictive features

11. Model Building

Logistic Regression

Random Forest Classifier

SVM

K-Nearest Neighbors (KNN)

XGBoost

Decision Tree

12. Model Evaluation

Random Forest Classifier:

Accuracy: 91.2%

Precision (Fatal): 95.6%

SVM:

Accuracy: 89.5%

Precision (Fatal): 92.1% Logistic Regression:

Accuracy: 80.3%

Precision: 85.7%

KNN:

Accuracy: 84.7%

Precision: 83.2%

Decision Tree:

Accuracy: 86.5%

Precision: 87.9%

13. Deployment

Built with Gradio for real-time predictions

Users input conditions like time, road surface, weather

Outputs predicted accident risk and severity

Can be embedded on government dashboards or used internally by traffic control centers

14. Future Scope

Use LSTM or Transformers to analyze sequential accident data

Integrate with live traffic APIs (Google Maps, GPS systems)

Develop mobile alert apps for drivers entering high-risk zones

Collaborate with traffic police to deploy prediction zones and signage

15. Team Members and Roles

Data Collection & Preprocessing: [G.Ashwin]

EDA & Feature Engineering: [M.Divagaran]

Model Building & Evaluation: [N.Srinivasan]

Interface & Deployment: [D.Vicknesh]