**Road Traffic Accident Analysis**

**Overview**

This project aims to analyze road traffic accidents using machine learning classification algorithms to identify major causes and predict accident severity based on various factors.

**Dataset**

The dataset used in this project contains records of road traffic accidents from 2017 to 2020. The original dataset was pre-processed and cleaned to exclude sensitive information. It consists of 32 features and 12,316 instances of accidents.

**Data Columns**

Age\_band\_of\_driver

Sex\_of\_driver

Educational\_level

Vehicle\_driver\_relation

Driving\_experience

Lanes\_or\_Medians

Types\_of\_Junction

Road\_surface\_type

Light\_conditions

Weather\_conditions

Type\_of\_collision

Vehicle\_movement

Pedestrian\_movement

Cause\_of\_accident

Accident\_severity

**Models**

Three machine learning models were employed for analysis:

Decision Tree Classifier

Random Forest Classifier

Gradient Boosting Classifier

**Evaluation Metrics**

The models were evaluated using various metrics including accuracy, precision, recall, and F1-score.

**Results**

Decision Tree: Accuracy - 73.6%, Precision - 73%, Recall - 74%, F1-score - 73%

Random Forest: Accuracy - 83.6%, Precision - 75%, Recall - 84%, F1-score - 77%

Gradient Boosting: Accuracy - 83.6%, Precision - 75%, Recall - 84%, F1-score - 77%

**Conclusion**

Based on the evaluation metrics, both Random Forest and Gradient Boosting models outperformed the Decision Tree model. Further analysis might be required to select the best model based on specific requirements.