Hammad and Haseeb

70131434 ans 70135821

AL NOOR COMPUTER

12/31/2025

[Type the abstract of the document here. The abstract is typically a short summary of the contents of the document. Type the abstract of the document here. The abstract is typically a short summary of the contents of the document.]

Project

Road Accident Severity Prediction

1. Introduction

Road accidents are one of the major causes of injuries and deaths across the world. The severity of a road accident depends on various factors such as weather conditions, road type, and driver condition. Predicting the severity of road accidents in advance can help traffic authorities and emergency services to take quick and effective actions.

This project uses machine learning techniques to predict the severity level of road accidents. The system analyzes different accident-related factors and classifies the accident severity into low, medium, or high levels.

2. Problem Statement

Road accident severity is usually identified after the accident occurs, which delays emergency response and increases damage. There is a need for an automated system that can predict accident severity using available accident data.

Problem:
To predict the severity of road accidents based on weather conditions, road type, and driver condition using machine learning algorithms.

3. Objectives

The objectives of this project are:

- To analyze factors affecting road accident severity
- To develop a machine learning-based prediction system
- To train and compare multiple classification algorithms
- To identify the best-performing model for accident severity prediction

4. Dataset Description

A structured dataset in Excel/CSV format is used for this project. The dataset contains 45 records and the following attributes:

4.1 Input Features

Attribute	Description
Weather	Weather condition during the accident (Clear, Rainy, Foggy)
Road_Type	Type of road (Highway, Urban, Rural)
Driver_Condition	Driver condition (Normal, Fatigued, Drunk)

4.2 Target Variable

Severity Value	Description
0	Low Severity
1	Medium Severity
2	High Severity

5. Tools and Technologies

- Python
- Google Colab
- Pandas
- Scikit-learn
- XGBoost
- Excel / CSV dataset

6. Methodology

The project follows the following steps:

6.1 Data Loading

The dataset is uploaded and loaded into Google Colab using Pandas.

6.2 Data Preprocessing

- Categorical features are converted into numerical values using Label Encoding

- Dataset is divided into features and target variable

6.3 Train–Test Split

The dataset is split into training and testing sets using an 75:25 ratio.

6.4 Model Training

Three machine learning models are trained:

- Decision Tree
- Random Forest
- XGBoost

6.5 Model Evaluation

Models are evaluated using:

- Accuracy score
- Confusion matrix
- Classification report

7. Algorithms Used

7.1 Decision Tree

Decision Tree works by splitting data based on conditions. It is simple and easy to interpret but may overfit on small datasets.

7.2 Random Forest

Random Forest uses multiple decision trees and combines their predictions. It improves accuracy and reduces overfitting.

7.3 XGBoost

XGBoost is a boosting algorithm that improves model performance by correcting previous prediction errors. It provides high accuracy and efficiency.

8. Model Training and Results

All three models were successfully trained in Google Colab. Accuracy scores were calculated for each model. Due to the structured dataset, Random Forest and XGBoost showed better performance compared to Decision Tree.

8.1 Model Comparison

Model	Performance
Decision Tree	Average
Random Forest	High
XGBoost	Very High

9. Confusion Matrix

A confusion matrix was generated to evaluate the prediction performance of the Random Forest model. It compares actual and predicted severity classes and helps identify classification errors.

10. Conclusion

This project successfully demonstrates the use of machine learning for predicting road accident severity. The trained models were able to classify accident severity based on weather, road type, and driver condition. Among the tested models, Random Forest and XGBoost performed better than Decision Tree. With a larger real-world dataset, the system's accuracy can be further improved.

11. Future Scope

- Use real-world traffic accident datasets
- Include additional features such as speed, time, and traffic density
- Deploy the system as a web or mobile application
- Integrate real-time accident prediction

12. Applications

- Traffic safety analysis
- Emergency response planning
- Road safety policy making

- Intelligent transportation systems