FINAL PROJECT PROPOSAL

Team No.22

Vladislav Zamuruev
Yorbe Heeren
Tudor Pitulice
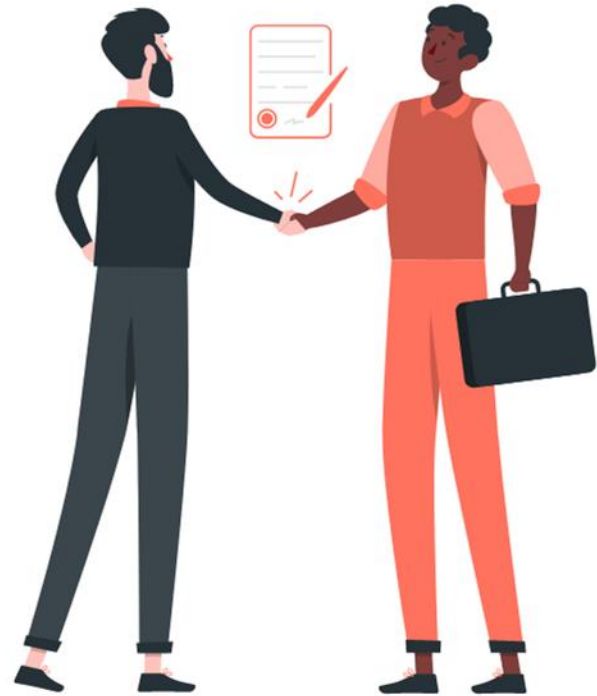*Vladislav Stoimenov*

TABLE OF **CONTENTS**

01

About the Project Idea

02

Project Stages

03

Legal & Ethics Considerations

04

ML iterations

01

About the Project

ABOUT THE PROJECT

- Long emergency response times.

- Increased severity of injuries and fatalities during traffic incidents.

- Higher economic costs due to extended incident durations.

WHAT
SOLUTION
DID WE
THINK OF?

OUR
STAKEHOLDERS

OUR AI CANVAS

AI Project Canvas

Title: Smart Accident Severity Prediction

Data
Which data do you need?

- Historical accident data (frequency, geographical coordinates, causes such as speeding, braking, cornering)
- Weather data (level of precipitations, snowing etc.)
- Real-time traffic data (vehicle speed, braking intensity, cornering angles)

Skills
Which skills do you need for development?

- Data Science and Machine Learning
- Data Engineering
- Data Translator
- Data Analyst
- Software Developer

Output
Which key metric are you optimizing for?

Accident severity prediction

Value Proposition
What is the value added by your project?

- Safety: improving road safety in Breda
- Efficiency: optimizing traffic flow in the city

Integration
How will the project be integrated?

- Traffic Management Infrastructure
- In-Vehicle Integration; Car Application

Stakeholders
Who are the key stakeholders?

- ANWB
- Municipality
- Citizens of Breda
- Authorities

Customers
Who are the end customers?

- ANWB customers
- Drivers in the municipality of Breda
- Citizens in Breda
- Tourists in Breda
- Authorities
- Hospitals
- Law enforcement forces

Cost
What costs will the project incur?

- Initial Development (Data Collection, Model Development, System Integration, Dashboard Creation)
- Marketing Campaign (User Trainings, System promotion and Learning, Advertising)

Revenue
How will the project generate revenue?

- Subscription Fees (Charging users a fee to use the system's dashboard)
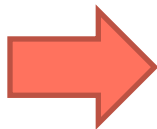- Service Contracts
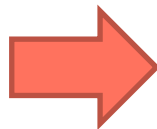- Partnerships

02

Project stages

Development Stages

PROJECT STAGES

03

Legal & Ethical considerations

You could enter a subtitle here if you need it

MAJOR REQUIREMENTS

Obligations

- Monitoring the behavior of the AI
- Documentation of the project
- Transparency
- Regular testing
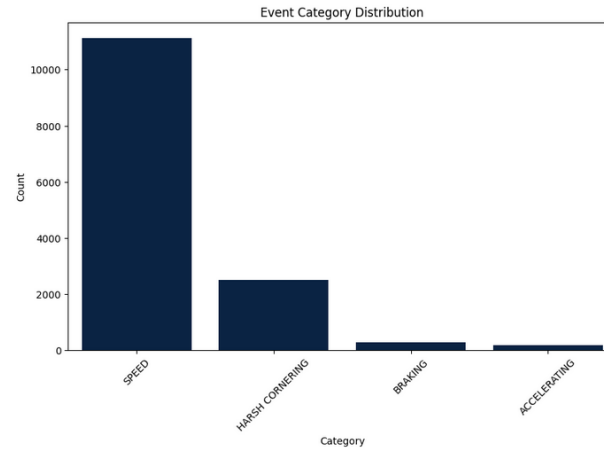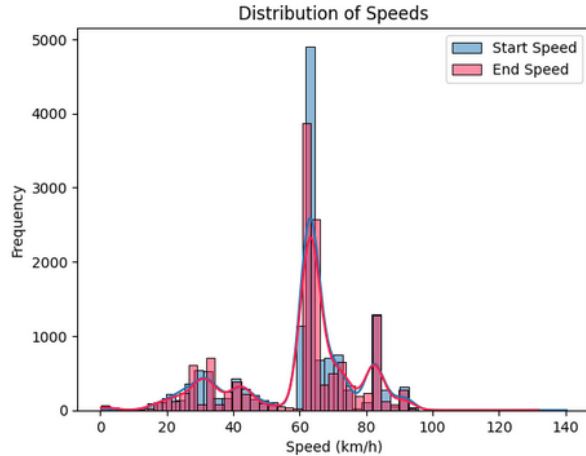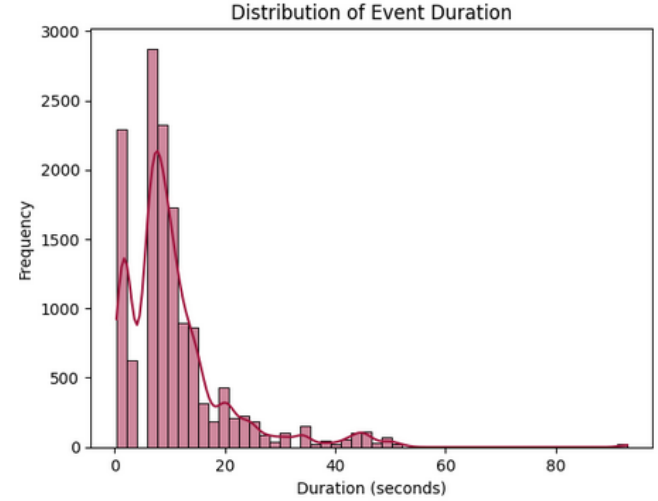- Compliance with the Data Protection Regulations
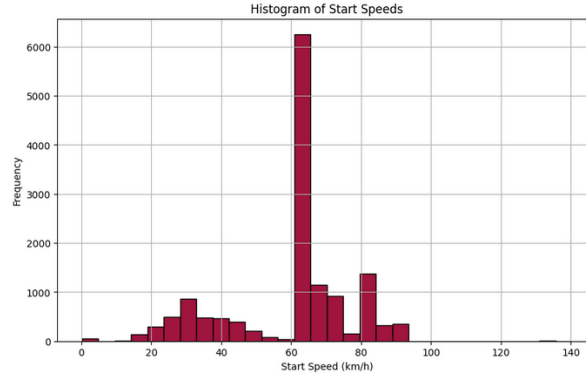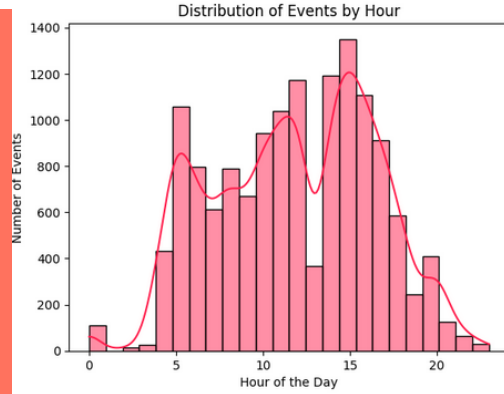
Compliances

- GDPR
- EU AI act
- Company internal policy
- Company ethics

04

ML iterations

EDA

Iteration 1

```python
categorical_transformer = Pipeline(steps=[
    ('imputer', SimpleImputer(strategy='most_frequent')),
    ('onehot', OneHotEncoder(handle_unknown='ignore'))
])
# Create transformers
numerical_transformer = Pipeline(steps=[
    ('imputer', SimpleImputer(strategy='median')),
    ('scaler', RobustScaler())
])
```

```
Random Forest Accuracy: 0.9968
              precision    recall  f1-score   support

           1       0.98      0.97      0.97       178
           2       1.00      1.00      1.00        27
           3       1.00      1.00      1.00        25
           4       1.00      1.00      1.00      2581

    accuracy                           1.00      2811
   macro avg       0.99      0.99      0.99      2811
weighted avg       1.00      1.00      1.00      2811

-------------------------------------------------
Gradient Boosting Accuracy: 0.9982
              precision    recall  f1-score   support

           1       1.00      0.97      0.99       178
           2       1.00      1.00      1.00        27
           3       1.00      1.00      1.00        25
           4       1.00      1.00      1.00      2581

    accuracy                           1.00      2811
   macro avg       1.00      0.99      1.00      2811
weighted avg       1.00      1.00      1.00      2811

-------------------------------------------------
XGBoost Accuracy: 0.9968
              precision    recall  f1-score   support

           1       0.98      0.97      0.97       178
           2       1.00      1.00      1.00        27
           3       1.00      1.00      1.00        25
           4       1.00      1.00      1.00      2581

    accuracy                           1.00      2811
   macro avg       0.99      0.99      0.99      2811
weighted avg       1.00      1.00      1.00      2811

-------------------------------------------------
DecisionTree Accuracy: 0.9968
              precision    recall  f1-score   support

           1       0.97      0.98      0.97       178
           2       1.00      1.00      1.00        27
           3       1.00      1.00      1.00        25
           4       1.00      1.00      1.00      2581

    accuracy                           1.00      2811
   macro avg       0.99      0.99      0.99      2811
weighted avg       1.00      1.00      1.00      2811

-------------------------------------------------
```
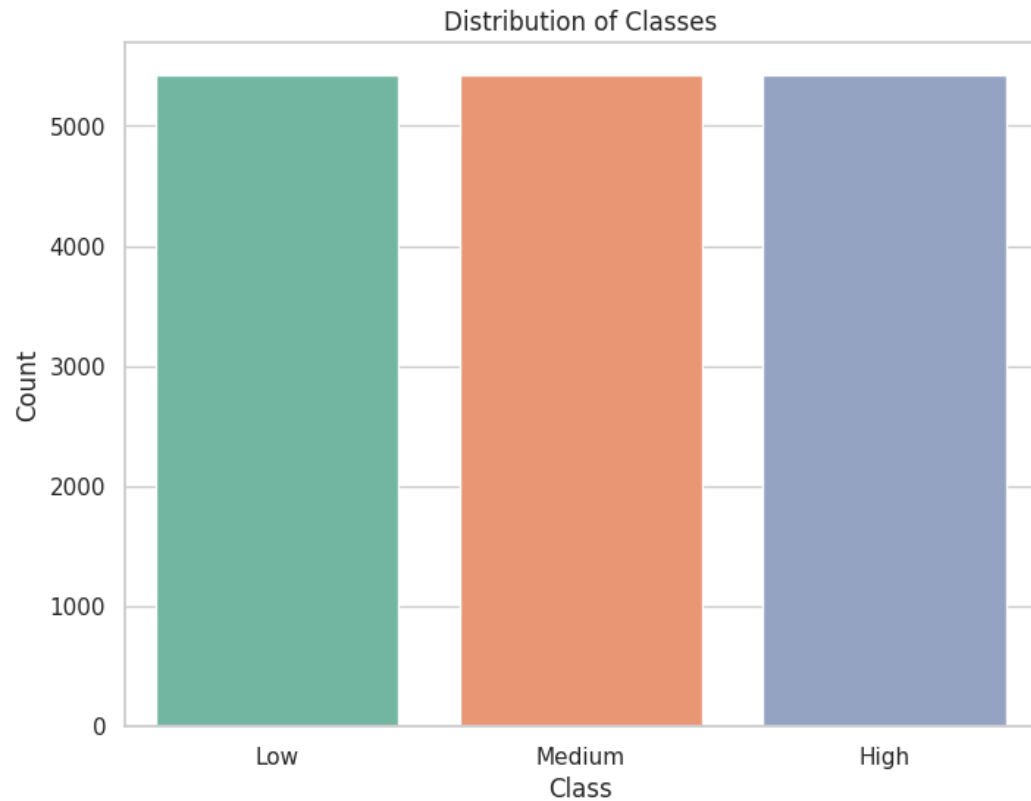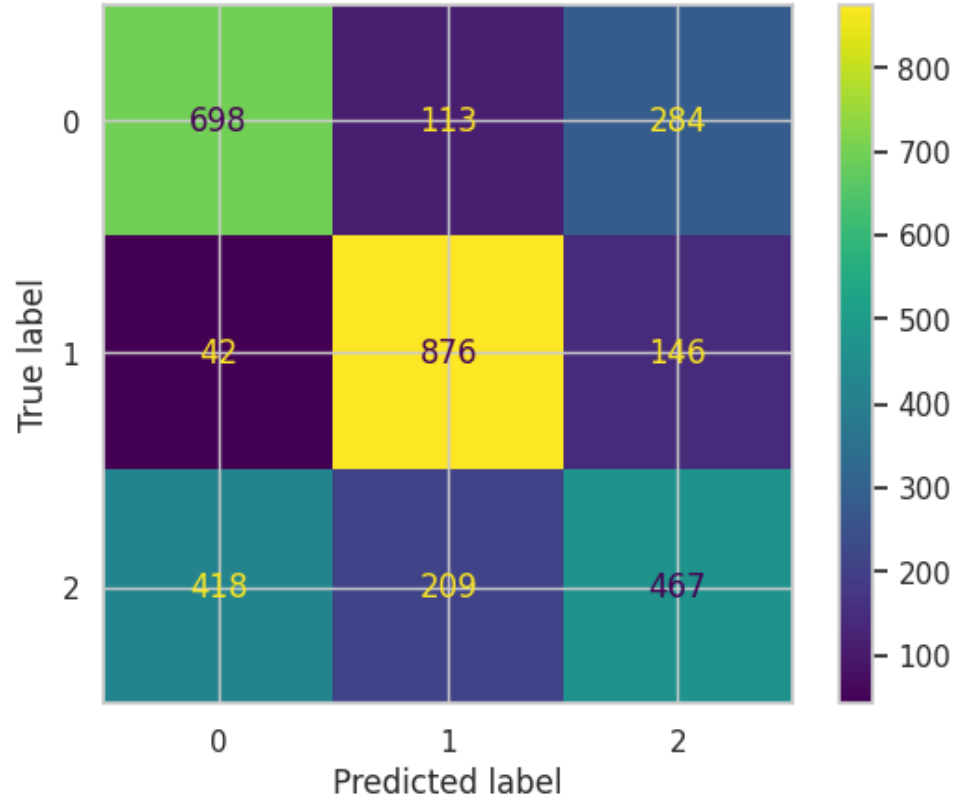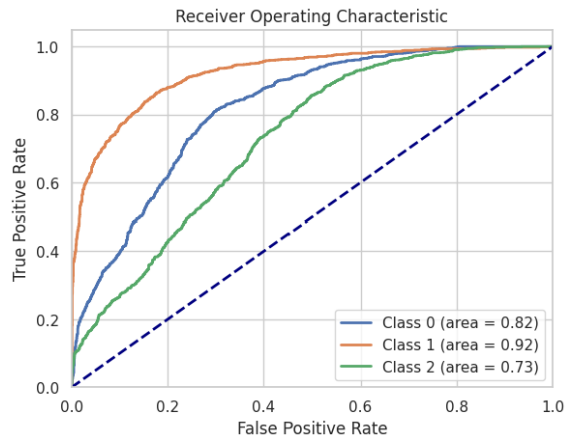

Distribution of Classes

Learning Curve
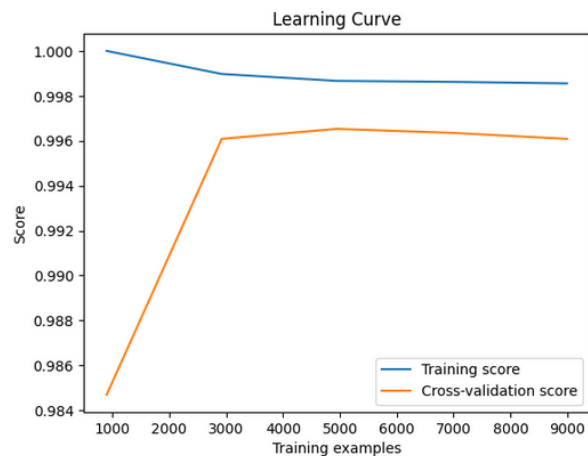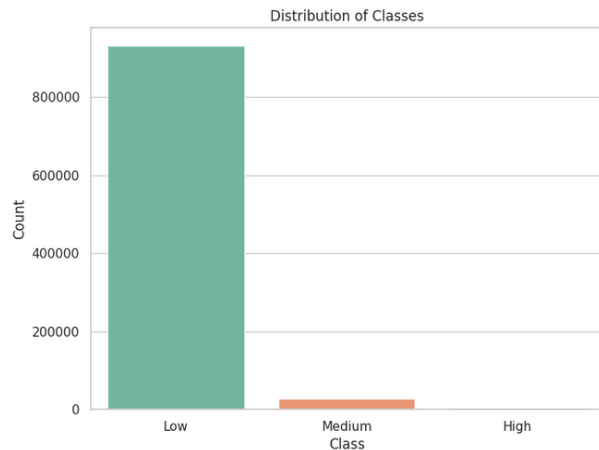
Best Iteration

Balanced classes!!!!

Distribution of Classes

Best Iteration
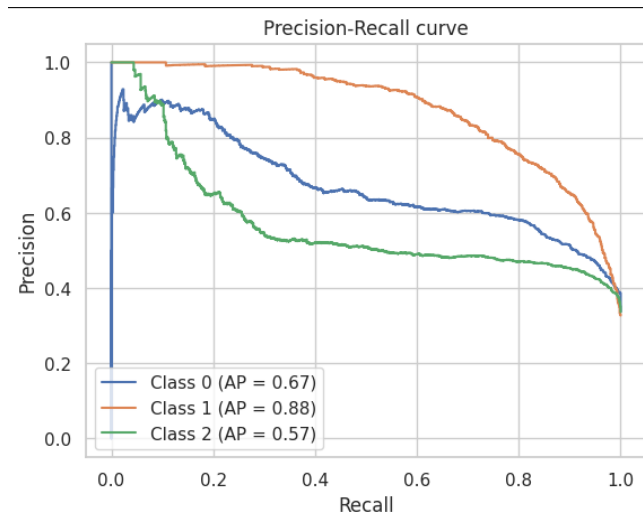
Confusion matrix of the best model:

Best Iteration

Roc-AUC curve

Precision Recall curve

Best Iteration

Gradient Boosting Classifier
before Hyperparameter tuning

```
Gradient Boosting

Accuracy: 0.6302
Macro Average Precision: 0.62
Macro Average Recall: 0.63
Macro Average F1-Score: 0.63
```

Gradient Boosting Classifier
after Hyperparameter tuning

```
Accuracy: 0.64

Macro Average:
    Precision: 0.63
    Recall: 0.64
    F1-Score: 0.64
Weighted Average:
    Precision: 0.63
    Recall: 0.64
    F1-Score: 0.64
```

OUR TEAM

Pitulice Tudor

Project Leader/ Data Engineer

Stoimenov Vladislav

Data Scientist

Heeren Yorbe

Analytics Translator

Zamuruev Vladislav

Data Scientist

THANK YOU!!!!!