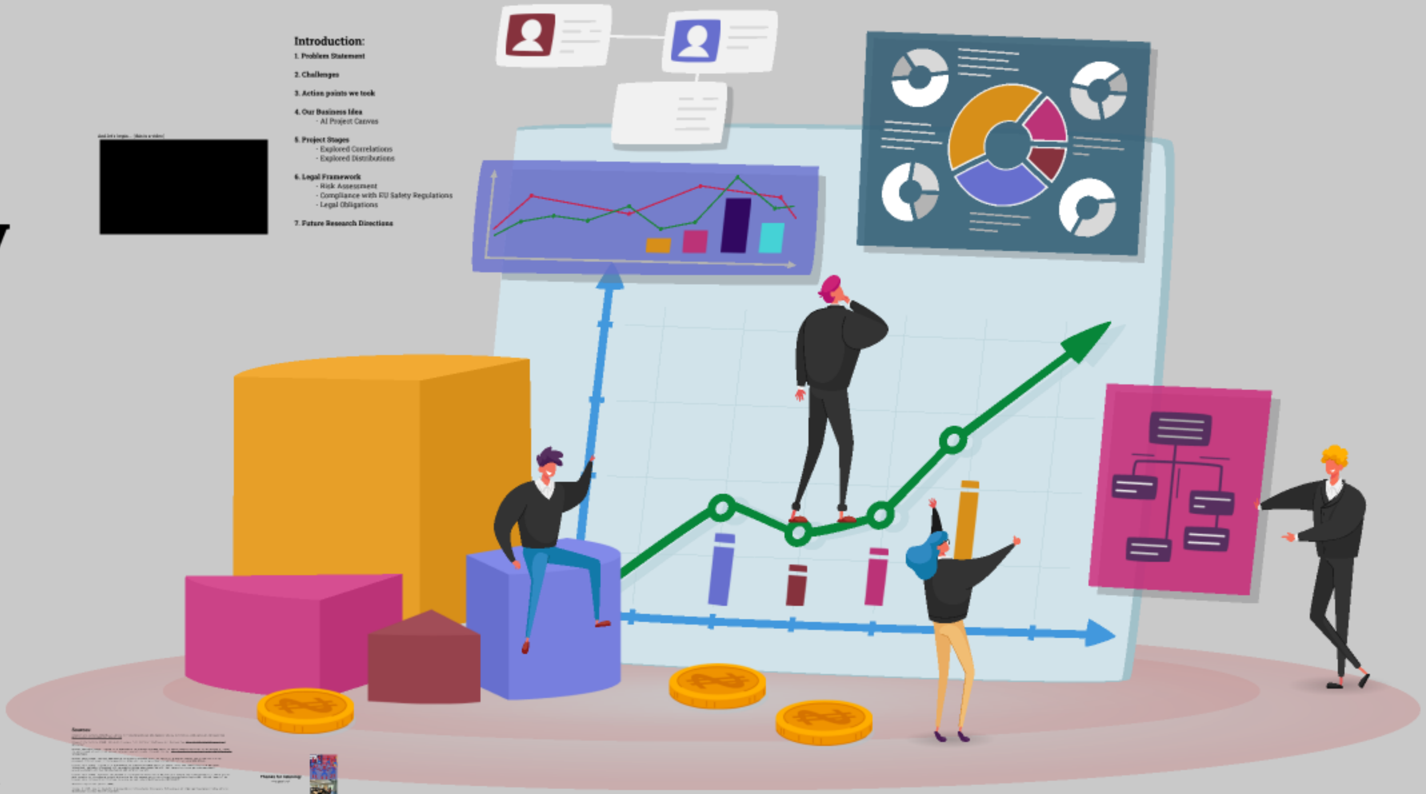
Contacts:
E-mail: 233182@buas.nl

Predictive Model for forecasting incident severity in Breda

Presented by Team 2

Introduction:

1. **Problem Statement**
2. **Challenges**
3. **Action points we took**
4. **Our Business Idea**
   - AI Project Canvas
5. **Project Stages**
   - Explored Correlations
   - Explored Distributions
6. **Legal Framework**
   - Risk Assessment
   - Compliance with EU Safety Regulations
   - Legal Obligations
7. **Future Research Directions**



Introduction:

1. Problem Statement

2. Challenges

3. Action points we took

4. Our Business Idea

- AI Project Canvas

5. Project Stages

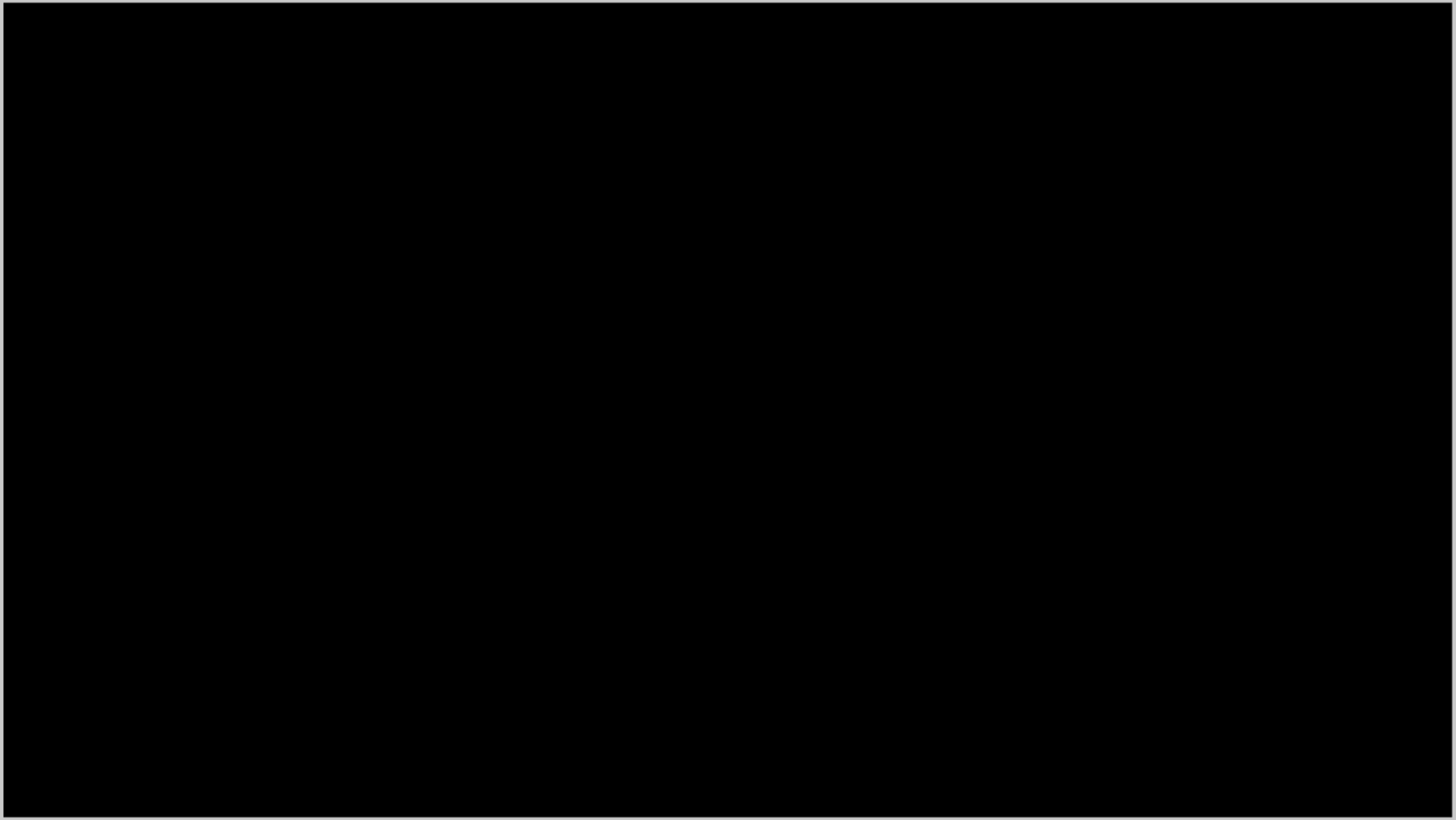
- Explored Correlations
- Explored Distributions

6. Legal Framework

- Risk Assessment
- Compliance with EU Safety Regulations
- Legal Obligations

7. Future Research Directions

And let's begin.... (this is a video)



Problem Statement: Incident Severity Prediction

Understanding the severity of road incidents is crucial for effective response and prevention strategies.

- Road incidents remain a critical concern globally and locally.
- Breda experiences frequent incidents with varying

  severity.

Source: https://nltimes.nl/2019/10/14/accidents-dutch-highways-doubled-5-years-report

Challenges of Incident Severity Prediction

Accurate prediction of incident severities helps in prioritizing emergency response and reducing traffic hazards.

- Current measures reduce accidents but struggle with incident severity.
- Vulnerable road users (pedestrians, cyclists, motorcyclists) are most affected.

What is our solution?

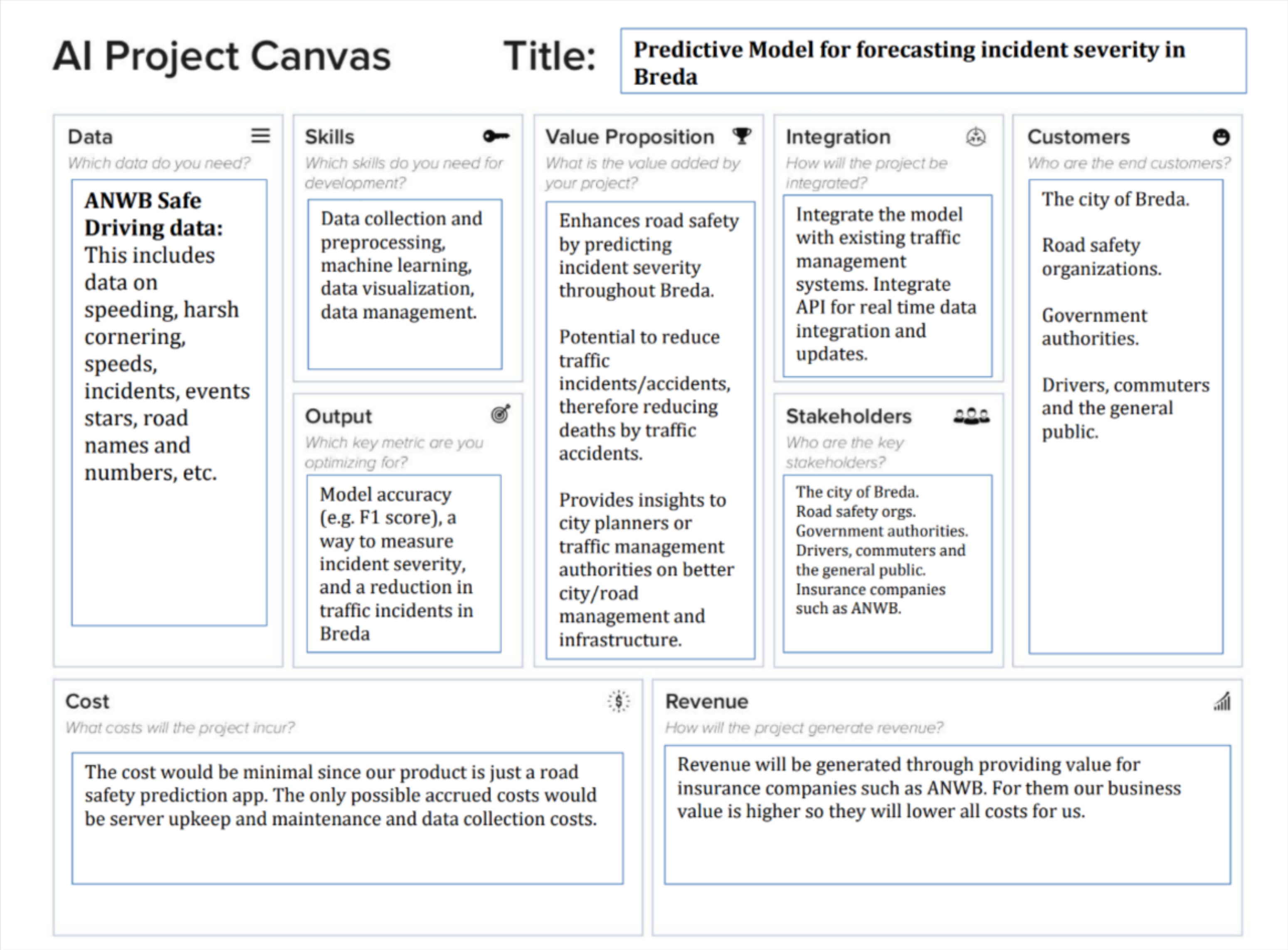
Analyzing historical data from the ANWB safe driving dataset provides insights into patterns and trends related to incident severity.

- We developed a Predictive Machine Learning model to forecast incident severity using data from the ANWB Safe Driving App.
- We won't solve all traffic problems; we just want to make it at least a little safer for everyone on the roads.

Source: https://static.vecteezy.com/system/resources/previews/000/400/556/original/creative-business-idea-concept-illustration-vector.jpg

Our Business Idea:

Incident Severity Bins



AI Project Canvas

Title: **Predictive Model for forecasting incident severity in Breda**

Data
Which data do you need?
ANWB Safe Driving data: This includes data on speeding, harsh cornering, speeds, incidents, events stars, road names and numbers, etc.

Skills
Which skills do you need for development?
Data collection and preprocessing, machine learning, data visualization, data management.

Output
Which key metric are you optimizing for?
Model accuracy (e.g. F1 score), a way to measure incident severity, and a reduction in traffic incidents in Breda

Value Proposition
What is the value added by your project?
Enhances road safety by predicting incident severity throughout Breda.

Potential to reduce traffic incidents/accidents, therefore reducing deaths by traffic accidents.

Provides insights to city planners or traffic management authorities on better city/road management and infrastructure.

Integration
How will the project be integrated?
Integrate the model with existing traffic management systems. Integrate API for real time data integration and updates.

Stakeholders
Who are the key stakeholders?
The city of Breda.
Road safety orgs.
Government authorities.
Drivers, commuters and the general public.
Insurance companies such as ANWB.

Customers
Who are the end customers?
The city of Breda.

Road safety organizations.

Government authorities.

Drivers, commuters and the general public.

Cost
What costs will the project incur?
The cost would be minimal since our product is just a road safety prediction app. The only possible accrued costs would be server upkeep and maintenance and data collection costs.

Revenue
How will the project generate revenue?
Revenue will be generated through providing value for insurance companies such as ANWB. For them our business value is higher so they will lower all costs for us.

Fig 1: AI project canvas

Project Stages:

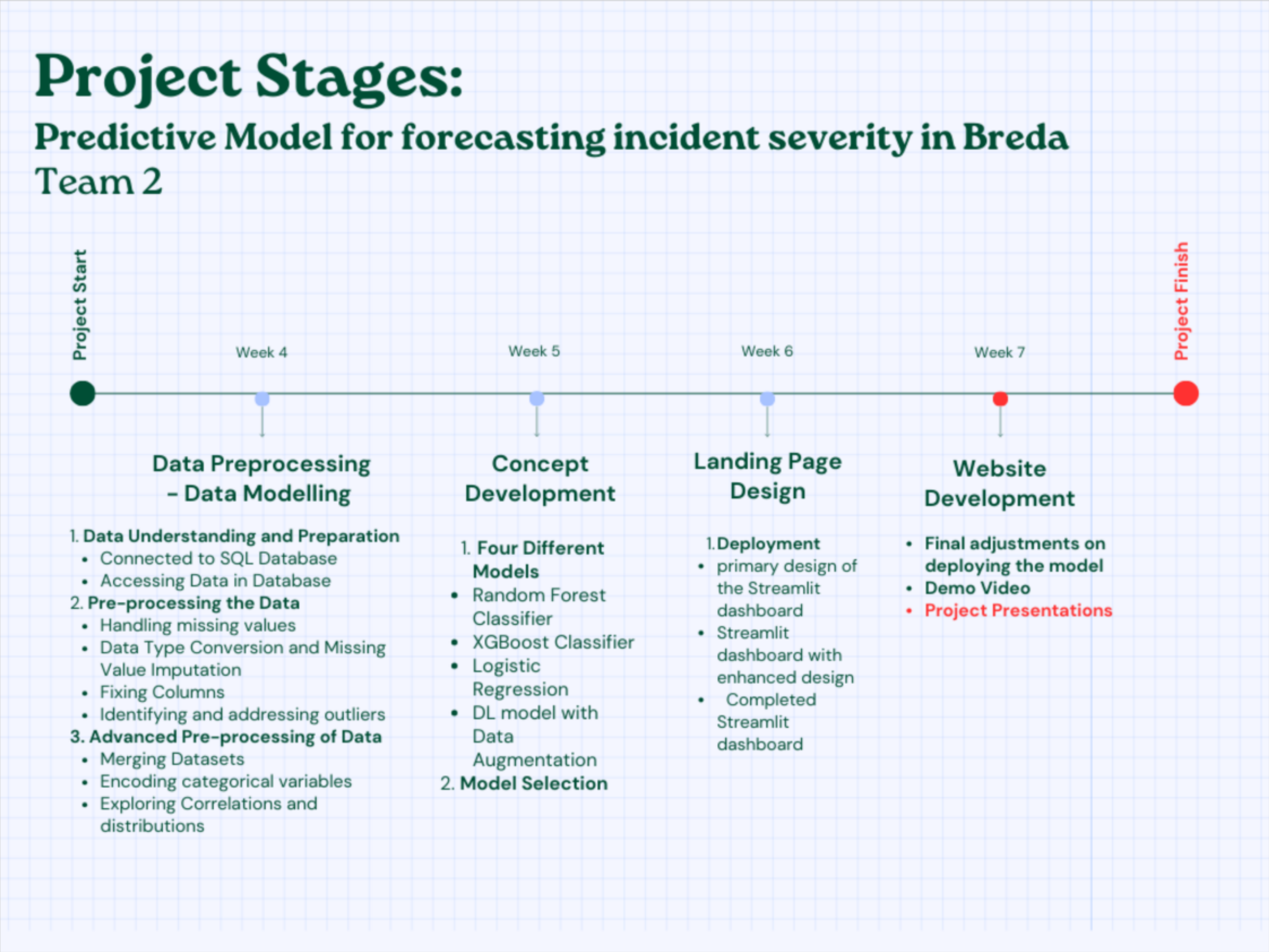


Fig 2: Project Stages Timeline

Exploring correlations and distributions:

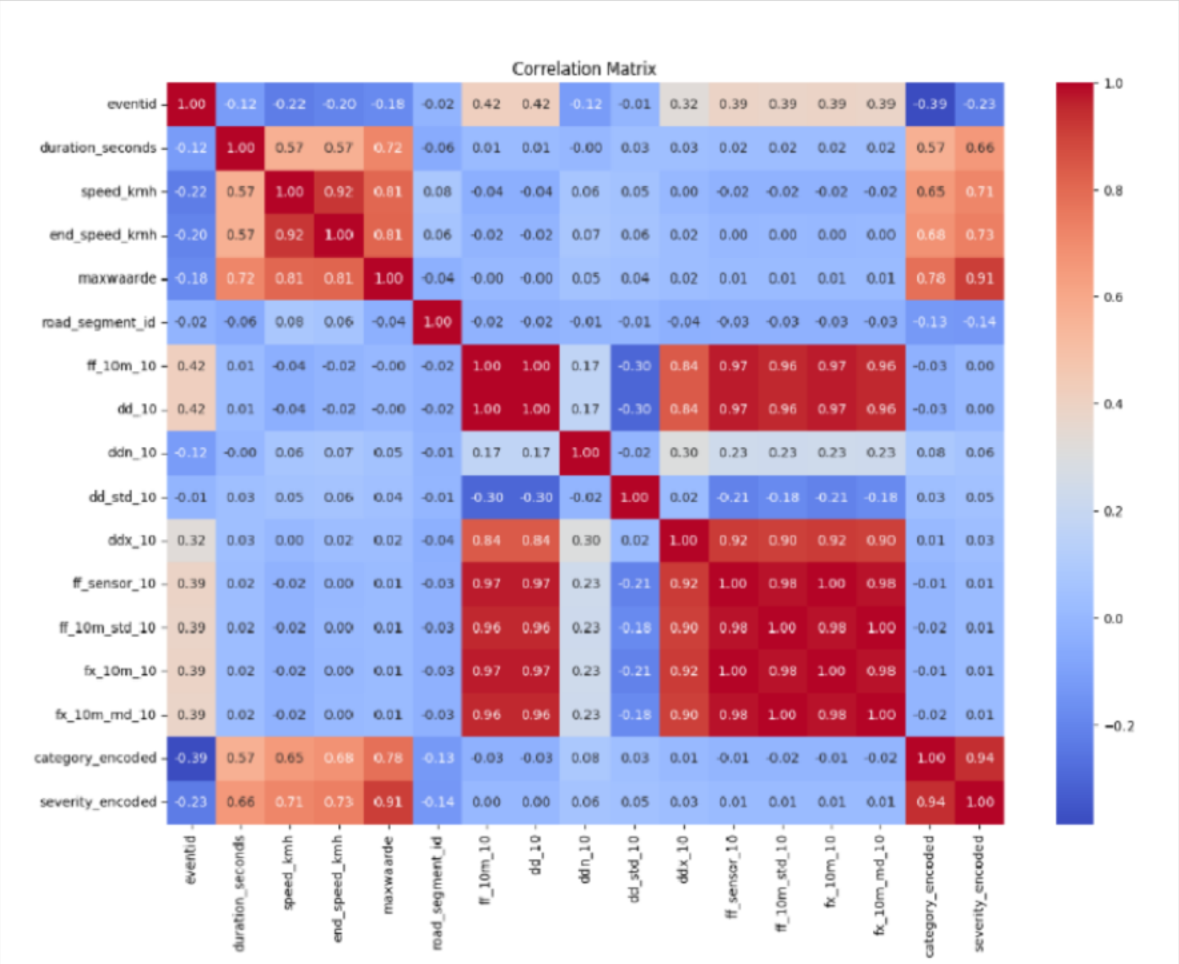


Fig 3: Correlation Matrix

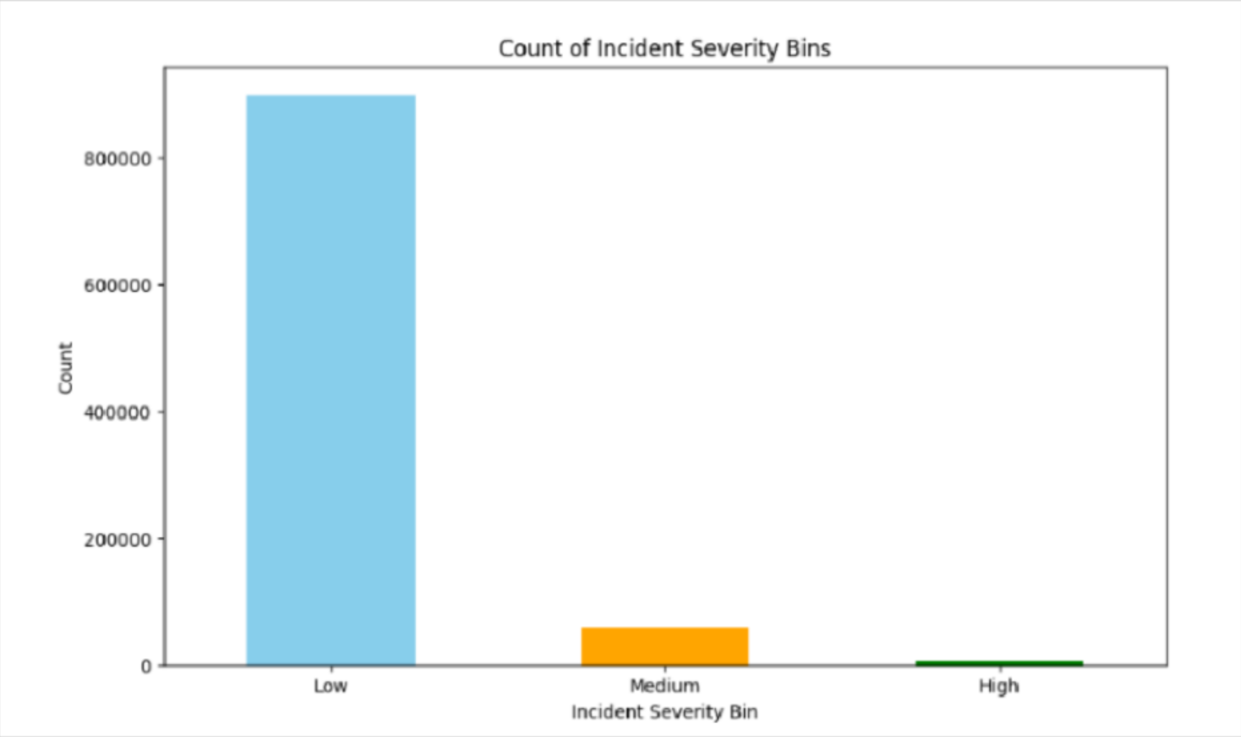


Fig 4: Count of Incident Severity Bins

Legal Framework :

Our AI system is of **High Risk**

Why is it high risk?

Safety Impact:

- Critical for road and human safety.
- Essential for emergency services and traffic maintenance.
- Prevents congestion and potential accidents.

Privacy Concerns:

- It handles sensitive data including geographic, weather, and personal driving information.
- Monitors routes and destinations, necessitating strict privacy measures.

Bias Mitigation:

- Ensure fairness and accuracy.
- Prevent biased predictions that could lead to unfair treatment of certain areas or demographics.

Compliance with EU Safety Regulations:

GDPR Compliance:

- **Key Principles** (Lawfulness, fairness, and transparency, Purpose limitation, Data minimization, Accuracy and Integrity and confidentiality):
- **High-Risk Classification:** It is Integrated into public infrastructure, impacting social safety.

Relevant EU Directives:

- **General Product Safety Directive** (GPSD): Ensures product safety in the market.
- **Directive on Road Infrastructure Safety Management (2008/96/EC):**
  - Requires risk assessments, safety audits, and inspections.
  - AI systems for road safety must align with this directive.

Regulation (EU) 2019/2144 on Type-Approval Requirement:

- Focuses on motor vehicle safety.
- Relevant for AI systems interfacing with vehicles

EDA, part 1

Overcoming challenges like data quality issues and model interpretability is crucial for enhancing the accuracy and reliability of incident severity predictions.

4.1 Data Cleaning

- Initial Inspection
  - Checked for duplicate and missing values.

  - Verified the integrity of key columns (e.g., event_id, event_start, etc.)
- Handling Missing Values
  - Identified columns with missing values.
- Fixing Columns
  - Visualized outliers using box plots.

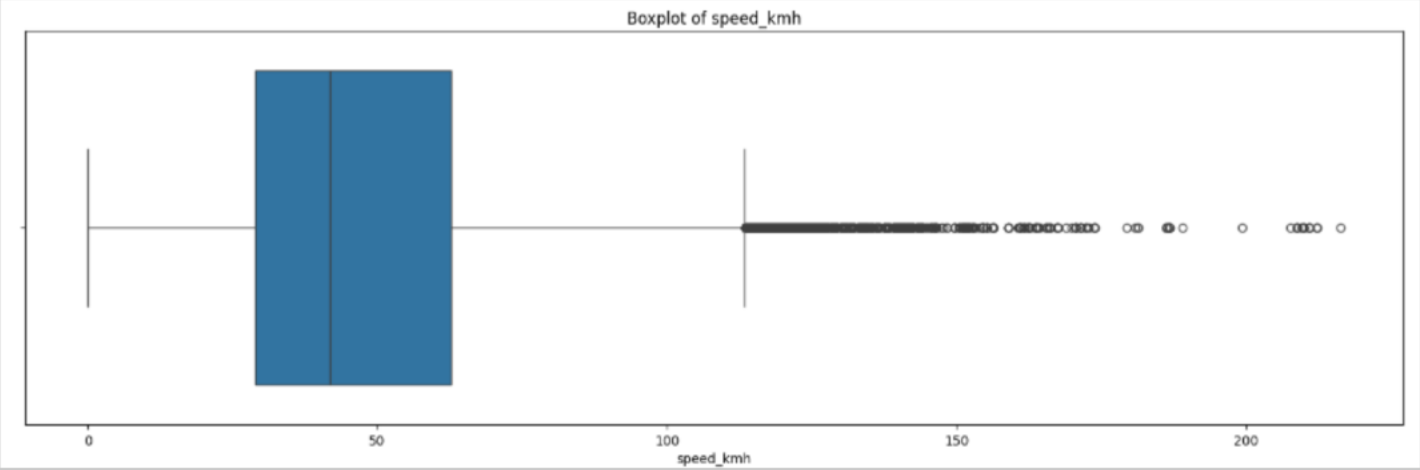


Fig 5: Boxplot of speed_kmh

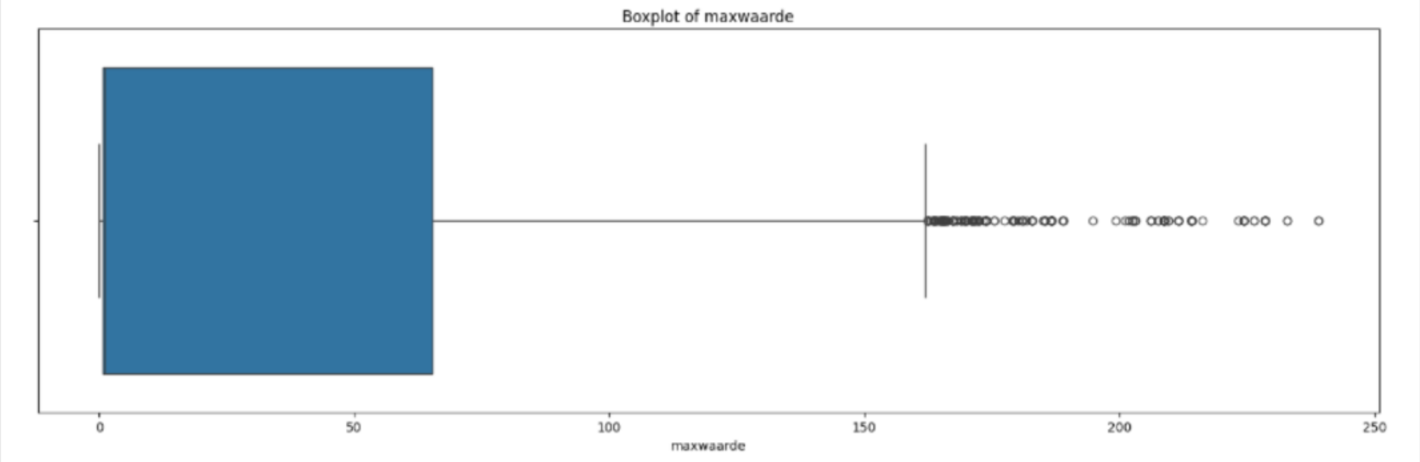


Fig 6: Boxplot of maxwaarde

EDA, part 2

- **Merging Datasets**

  - Loaded data from 'safe_driving' and 'wind' tables using SQL queries.
  - Ensured temporal consistency by converting datetime columns.
  - Merged datasets on the 'date' column.
  - Pruned unnecessary columns for a cleaner dataset.

- **Encoding Categorical Variables**

  - Used LabelEncoder to transform categorical variables into numerical formats.
  - Encoded 'category' and 'incident_severity' columns.

- **Correlation Analysis**

  - Generated a correlation matrix to explore relationships between variables.
  - Excluded wind-related data due to low correlation with incident severity.

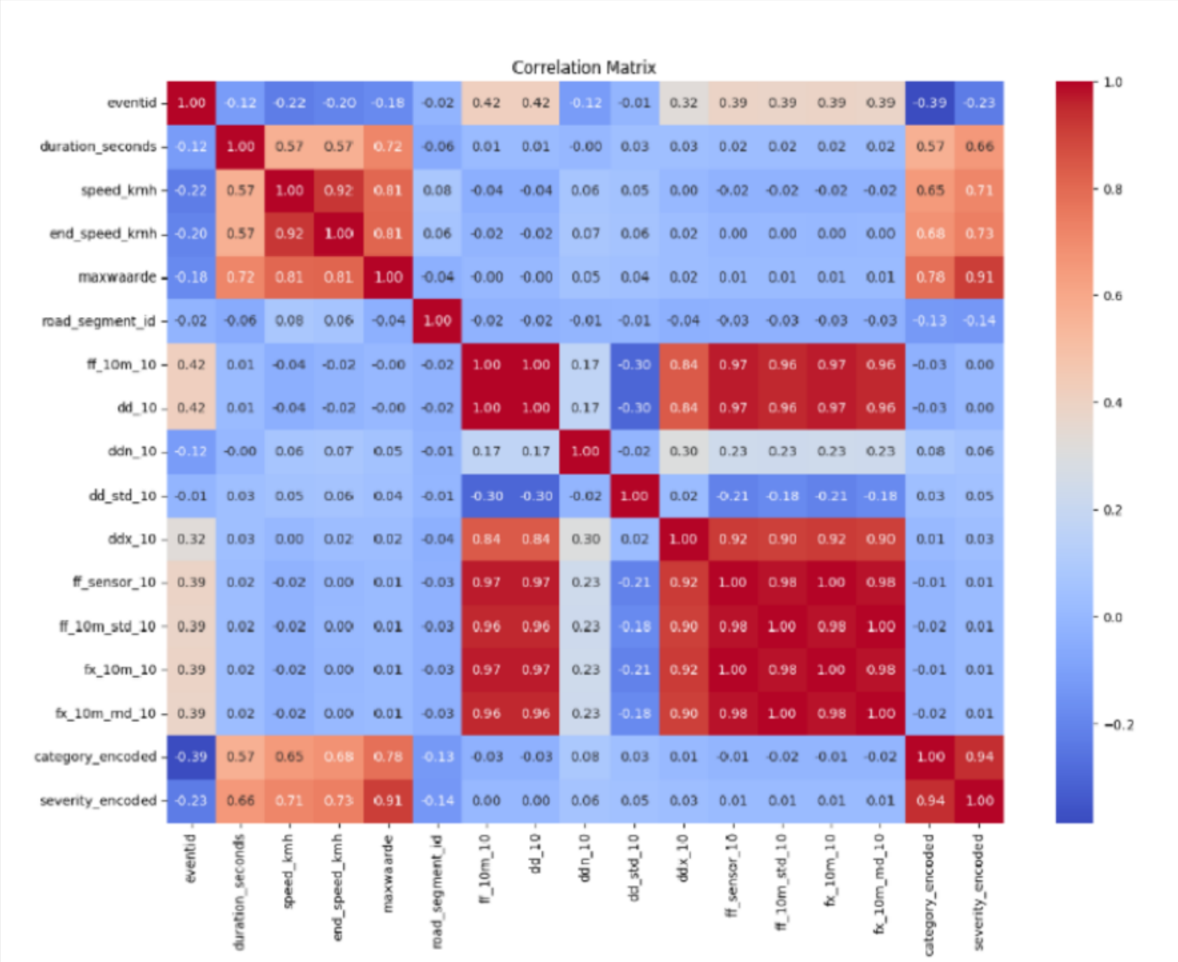


Fig 7: Correlation Matrix

Iteration 1

Logistic Regression

- More time to run
- Accuracy: 60%
- Other metrics are also with low scores

Classification Report:

	precision	recall	f1-score	support
Low	0.64	0.63	0.64	12607
Medium	0.59	0.72	0.65	12685
High	0.58	0.47	0.51	12828
accuracy			0.60	38120
macro avg	0.60	0.60	0.60	38120
weighted avg	0.60	0.60	0.60	38120

Fig 8: Classification Report

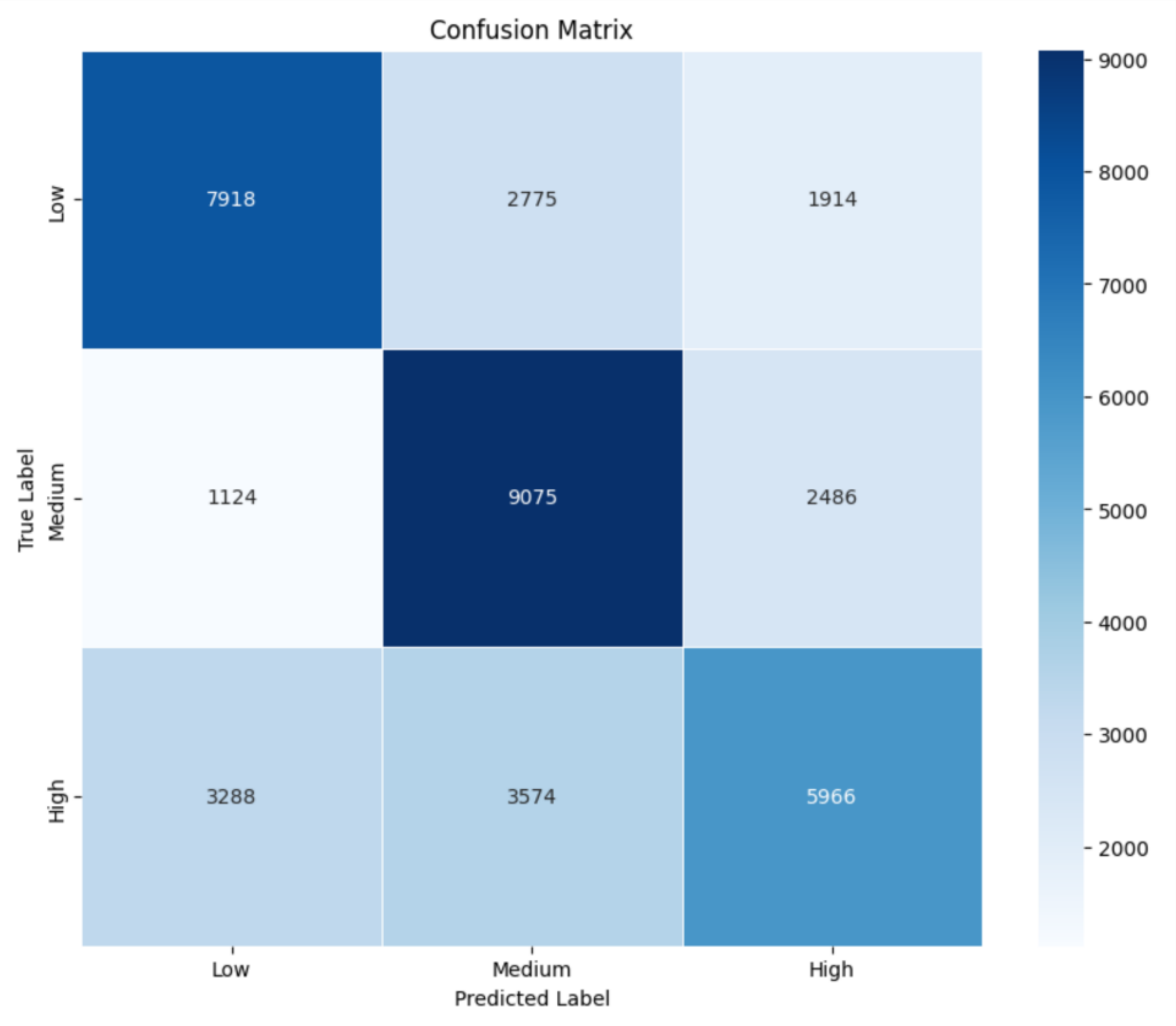


Fig 9: Confusion Matrix

Best Iteration:

Random Forest Classifier

- Around 20-30 seconds to run
- Accuracy: 95%
- Other metrics are also around 97-98

Classification report:

	precision	recall	f1-score	support
0	0.93	1.00	0.97	12607
1	0.98	0.92	0.95	12685
2	0.93	0.92	0.93	12828
accuracy			0.95	38120
macro avg	0.95	0.95	0.95	38120
weighted avg	0.95	0.95	0.95	38120

Fig 10: Classification Report Best Model

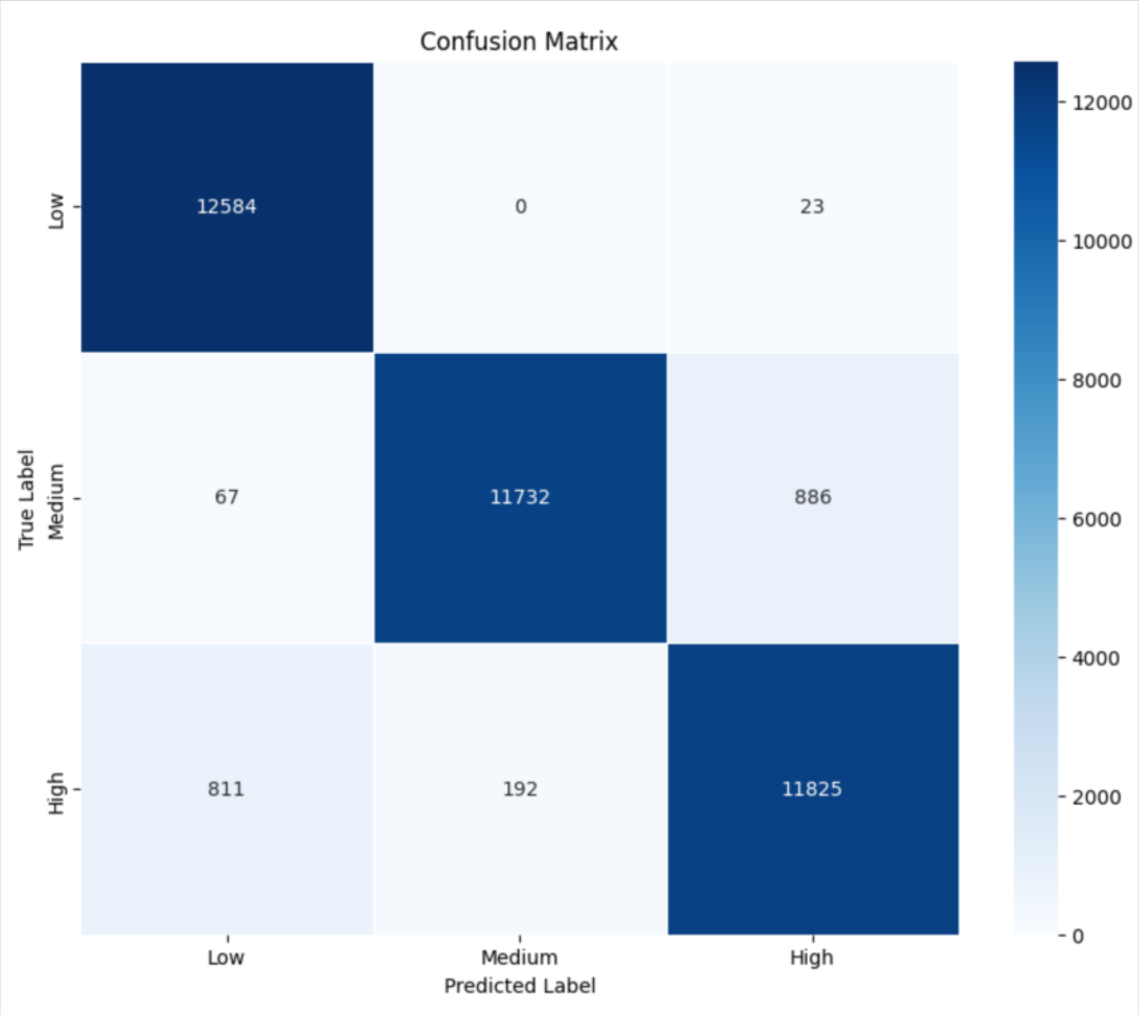


Fig 11: Confusion Matrix Best Model

DEMO:

http://194.171.191.226:3551/

Future Research Directions:

- **Evaluate Effectiveness of Interventions:**

- Assess various interventions based on insights from our tool.

- Example: Compare incident rates before and after installing speed bumps on identified hotspot roads.

- **Expand Model Scope:**

- Integrate more granular environmental data and additional behavioral metrics.

- Incorporate real-time data from emerging technologies like connected vehicle systems and smart city infrastructure.

- Enable dynamic, responsive model providing real-time alerts and recommendations.

Stakeholder Collaboration:

- Maintain a continuous feedback loop with ANWB, local authorities, and the public.

- Refine the model based on evolving community needs.

- Conduct longitudinal studies to track long-term impact of interventions on road safety.

Sources:

Future of Life Institute. (2024 May 2). Article 5: Prohibited artificial intelligence practices. EU Artificial Intelligence Act. Retrieved from https://artificialintelligenceact.eu/article/5/

Future of Life Institute. (2024). Article 52: Procedure. In EU Artificial Intelligence Act. Retrieved from https://artificialintelligenceact.eu/article/52/

European Commission. (2021). Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL amending Regulation (EU) No 2018/848 as regards the reinforcement of controls and measures and organic production rules [Proposal]. Eur-Lex. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206

European Union. (2008). Directive 2008/96/EC of the European Parliament and of the Council of 19 November 2008 on road infrastructure safety management. Official Journal of the European Union. https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex%3A32008L0096pub

European Union. (2021). Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS [PDF]. https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0001.02/DOC_1&format=PDF

European Union. (2016). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation). Official Journal of the European Union. Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32016R0679

Annex III: High-Risk AI Systems. (2024).

Tanjaya, A. (2024, June 6). How better AI documentation practices foster transparency. Partnership on AI. https://partnershiponai.org/how-better-ai-documentation-practices-foster-transparency/

Omroep Brabant. (n.d.). Slachtoffers fatbike ongeluk zijn meisjes van 8 en 13 jaar oud. Retrieved from https://www.omroepbrabant.nl/nieuws/4491341/slachtoffers-fatbike-ongeluk-zijn-meisjes-van-8-en-13-jaar-oud

NL Times. (2023, August 21). Romanian man identified as victim in fatal Breda drunk driving crash. Retrieved from https://nltimes.nl/2023/08/21/romanian-man-identified-victim-fatal-breda-drunk-driving-crash

The Brussels Times. (n.d.). E19 completely closed in direction to Breda following accident. Retrieved from https://www.brusselstimes.com/61498/e19-completely-closed-in-direction-to-breda-following-accident

NL Times. (2024, June 17). Police: Driver possibly under influence in crash that hurt children riding fat bike. Retrieved from https://nltimes.nl/2024/06/17/police-driver-possibly-influence-crash-hurt-children-riding-fat-bike

The Sun. (2019, February 25). Hollyoaks Spring trailer 2019 – Mercedes kills. Retrieved from https://www.thesun.co.uk/tvandshowbiz/8826713/hollyoaks-spring-trailer-2019-mercedes-kills/

Thanks for listening!

Any questions?