Weather Shield Breda

Monika Stangenberg, Kiril Baharev, Stijn van der Pas, Bartosz Kudyba

Agenda

3 Introduction

The Weather Shield Breda Project enhances Breda city safety through predictive analytics.

Increase safety in Breda by predicting the severity of incidents based on weather conditions and incident history.

Using data-driven decision-making, the Weather Shield Breda Project aims to keep all residents of Breda safer.

Key Features:

- **Real-time Updates:** Provide real-time updates for citizens of Breda.
- **Danger Areas Monitoring:** Identify and monitor previously identified danger areas to determine if they are at risk on any given day based on current weather conditions.

Data Analysis and Preparation

5 Data Collection

We used two main sources to build our dataset for the predictive model:

- **Weather Data: Including temperature, relative humidity, precipitation, wind speed and sunshine duration.**
- **ANWB Data: Records of incidents in Breda, provided by ANWB.**

6 Data Preprocessing Steps

Loading Data:

- Established a connection to a PostgreSQL database.
- Retrieved column names and sample data from the weather table.

Handling Missing Values:

- Identified and removed columns with excessive missing values.
- Created a cleaned weather table without the visibility column.

Standardizing Data:

- Converted values in the `is_day` column to integers (0 and 1).

Detecting and Handling Outliers:

- Examined numerical columns for extreme values.
- Removed top 400 outliers in the ANWB data.

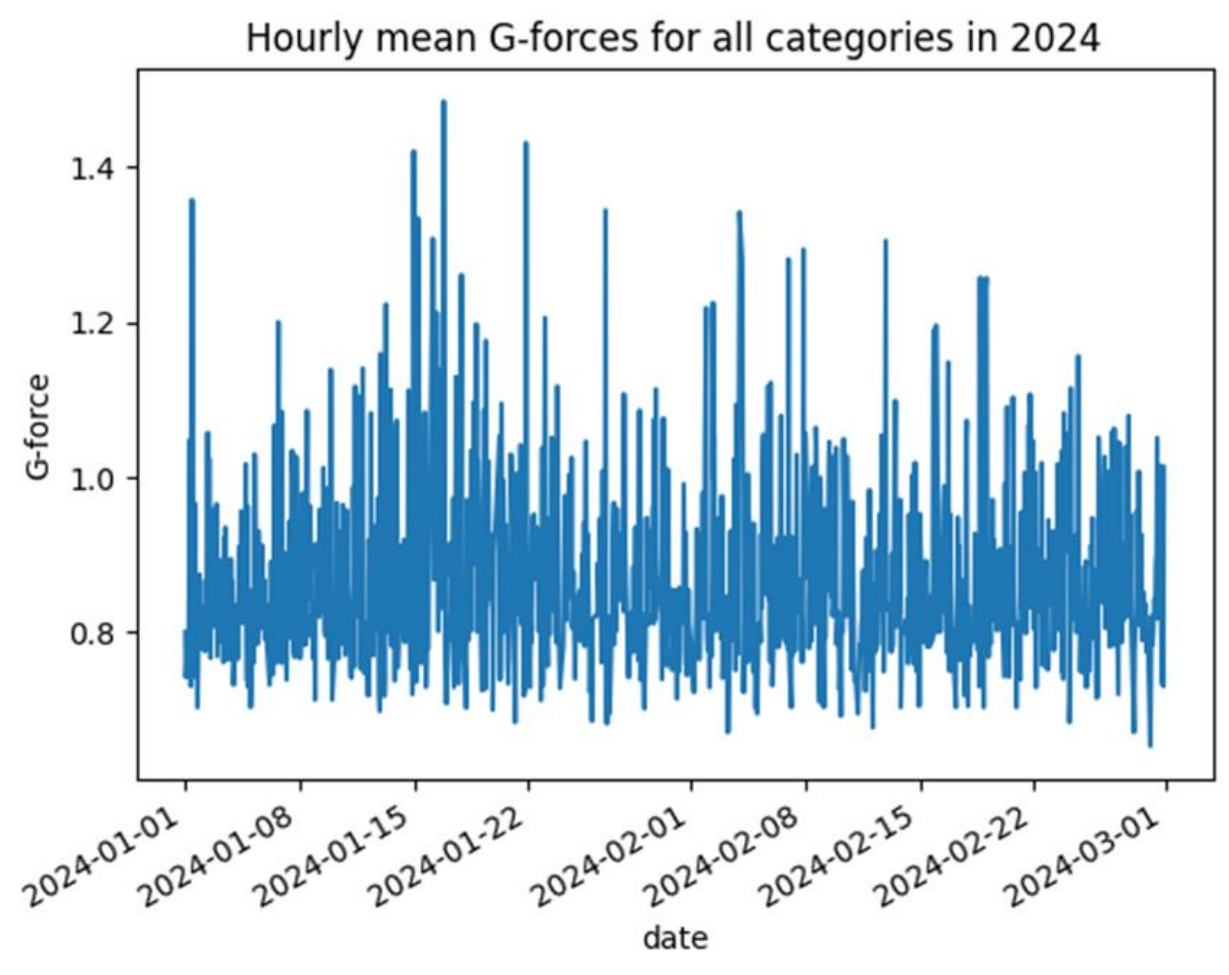
Loading and Cleaning ANWB Data:

- Loaded the ANWB incident data.
- Created a simplified view for easier analysis.

Data Integration:

- Joined weather and ANWB data to analyze relationships between weather conditions and driving incidents.

7 Visualizations and Insights



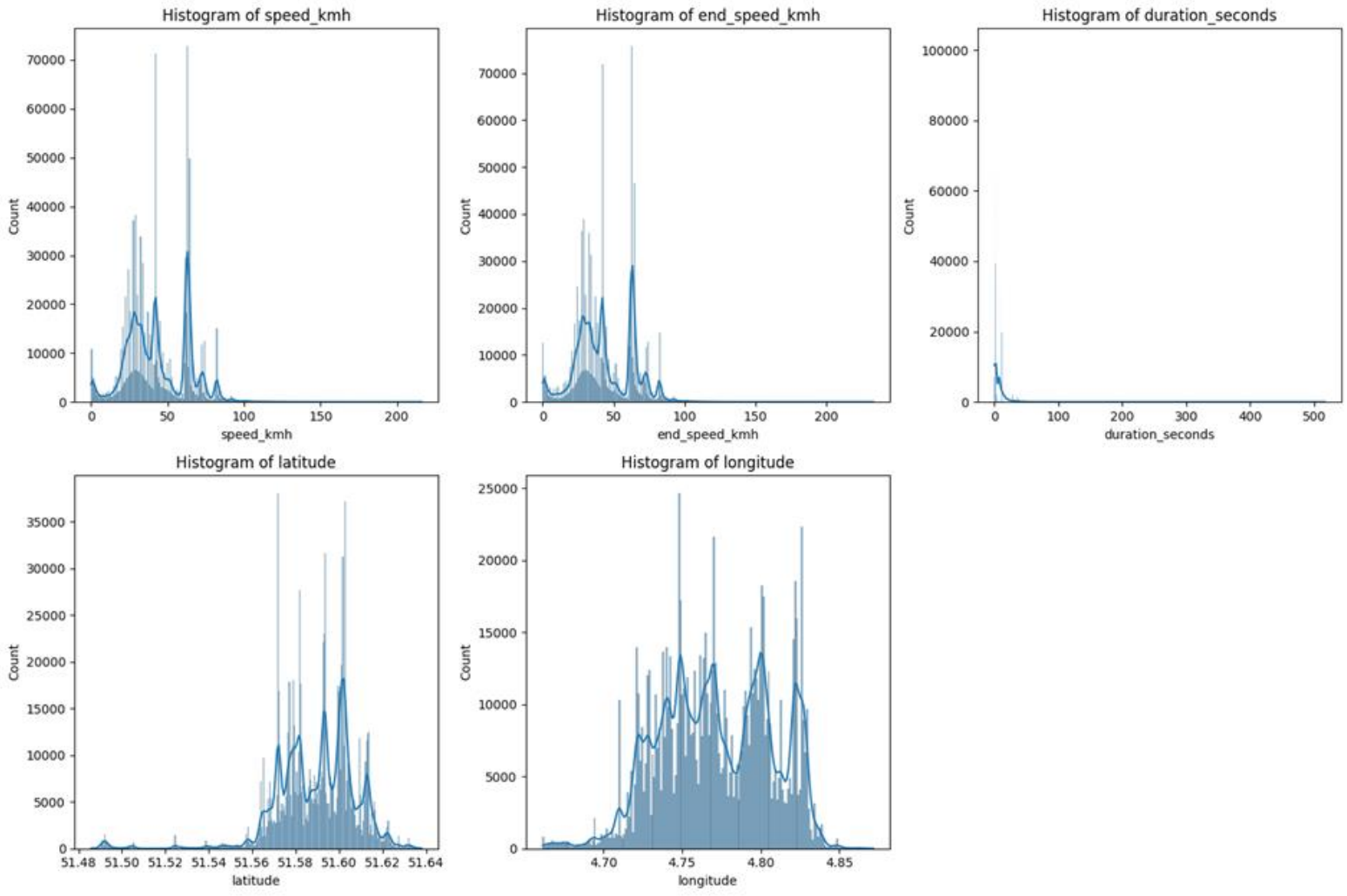
Hourly Mean G-Forces:

Plotted to visualize driving behavior across different times of the day.

8 Visualizations and Insights

Outlier Analysis:

Visualized distributions to detect and handle outliers effectively.



9 LSTM model

Our project uses an LSTM (Long Short-Term Memory) model to predict incident severity based on weather data.

Model: "sequential_21"

Layer (type)	Output Shape	Param #
conv1d_18 (Conv1D)	(None, 24, 64)	1,664
bidirectional_17 (Bidirectional)	(None, 24, 128)	66,048
bidirectional_18 (Bidirectional)	(None, 64)	41,216
dropout_1 (Dropout)	(None, 64)	0
dense_44 (Dense)	(None, 100)	6,500
dense_45 (Dense)	(None, 100)	10,100
dense_46 (Dense)	(None, 4)	404

Total params: 377,798 (1.44 MB)

Trainable params: 125,932 (491.92 KB)

Non-trainable params: 0 (0.00 B)

Optimizer params: 251,866 (983.86 KB)

10 LSTM model

- **Convolutional Layer (Conv1D): Extracts local features from the time-series data.**
- **Bidirectional LSTM Layers: These layers capture patterns in the data by processing it in both forward and backward directions.**
- **Dropout Layer: Applied to prevent overfitting.**
- **Dense Layers: Process the extracted features and make the final predictions.**

11 LSTM model

Results of the model classes(((

	precision	recall	f1-score	support
0	0.86	0.96	0.91	8412
1	0.96	0.90	0.93	28058
2	0.86	0.90	0.88	8464
3	0.83	0.96	0.89	2514
accuracy			0.92	47448
macro avg	0.88	0.93	0.90	47448
weighted avg	0.92	0.92	0.92	47448

12 Legal Compliance

Risk Management (Article 9)
We reduce potential risks our AI system might have, making sure it gives safe advice.

Data Quality (Article 10)
We use accurate and unbiased data to train our AI. This ensures our predictions are reliable.

Technical Documentation (Article 11)
We keep detailed records of how we built and tested the AI system. This helps us improve it over time.

Record-Keeping (Article 12)
We keep detailed records of the AI's design and updates to ensure compliance and easy tracking.

13 Legal Compliance

Transparency (Article 13)
We explain clearly to users what the AI can and cannot do, helping them use it correctly.

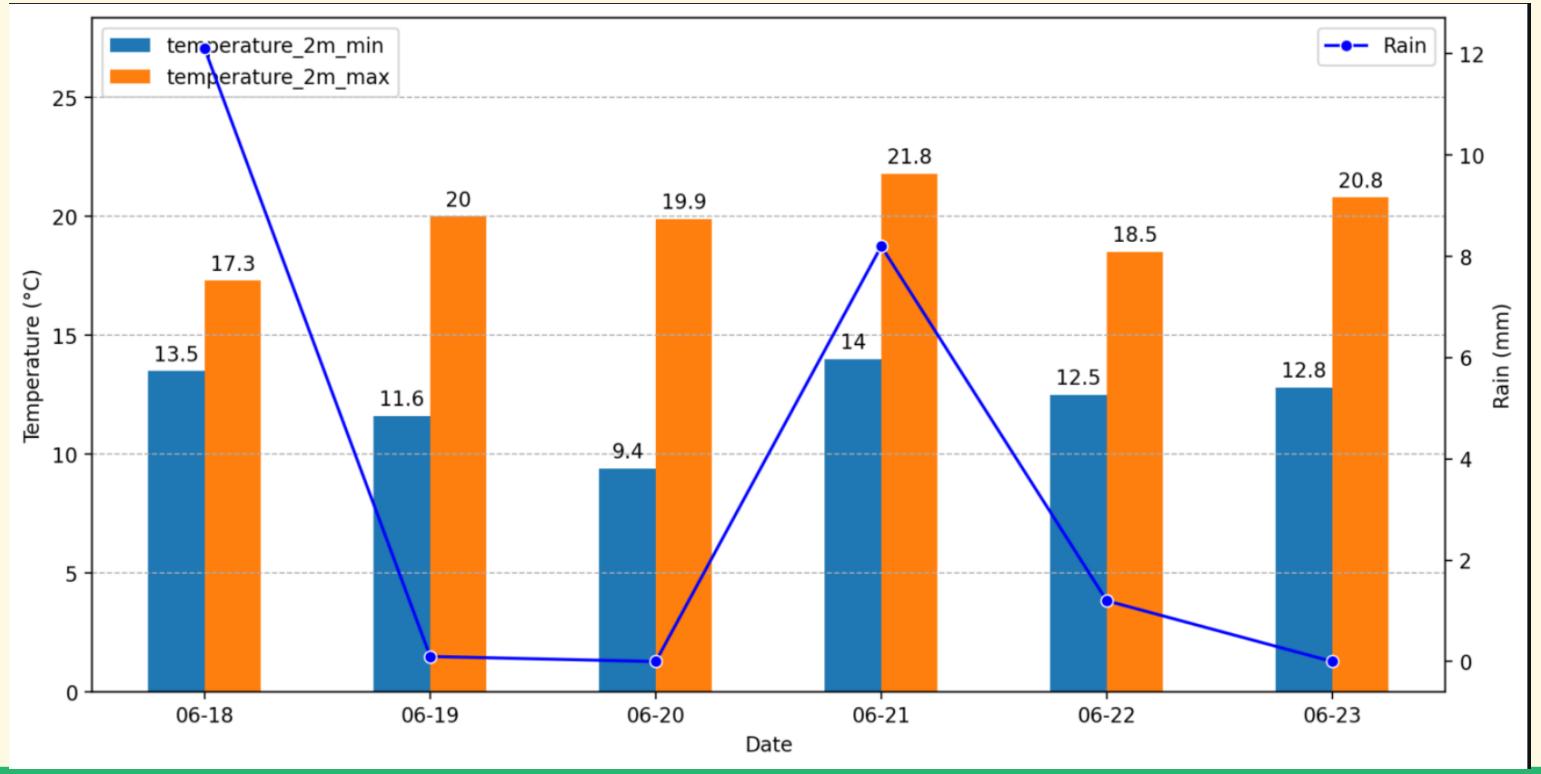
Human Oversight (Article 14)
Human intervention mechanisms correct any AI errors, ensuring safe and accurate advice.

By following these rules, we ensure the Weather Shield AI system is safe, reliable, and fair, meeting EU legal standards.

14 Deployment

15 Deployment

The website includes weather forecast for the next 5 days



16 Deployment

Filter Data

Select Date Range

2018/01/01 – 2024/02/29

Select Category

All

Showing data for category: All

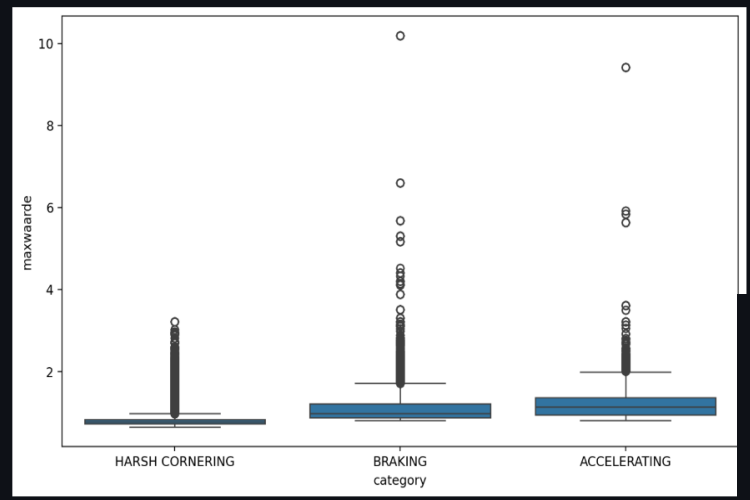
	maxwaarde	temperature_2m	relative_humidity_2m	precipitation	wind_speed_10m	sunshine_dura
count	1,947,628	1,947,628	1,947,628	1,947,628	1,947,628	1,947
mean	0.8653	12.4108	74.1551	0.1043	15.3518	1,749.7
std	0.2118	6.8626	16.4383	0.369	7.4799	1,716.6
min	0.6437	-9.563	19.1184	0	0	
25%	0.7339	7.387	62.5031	0	9.7267	
50%	0.8031	11.937	77.2955	0	14.3459	1,350.1
75%	0.9045	17.487	87.9628	0	19.8131	3
max	10.1823	37.437	100	15.7	59.509	3

Users can filter weather data.

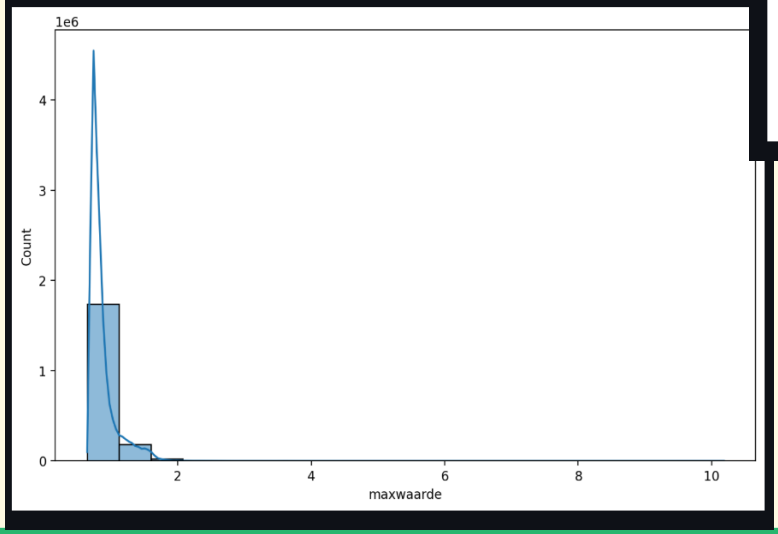
17 Deployment

Based on the selected time range of weather data, users can view relevant diagrams.

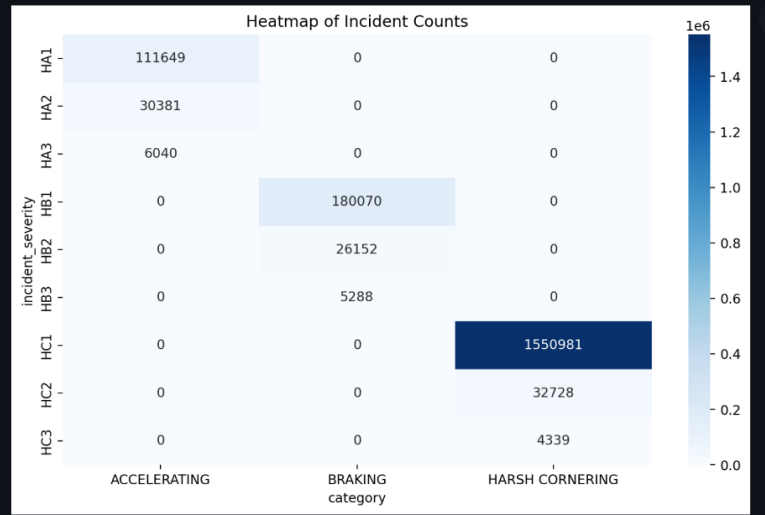
Box Plot for 'maxwaarde'



Histogram for 'maxwaarde'

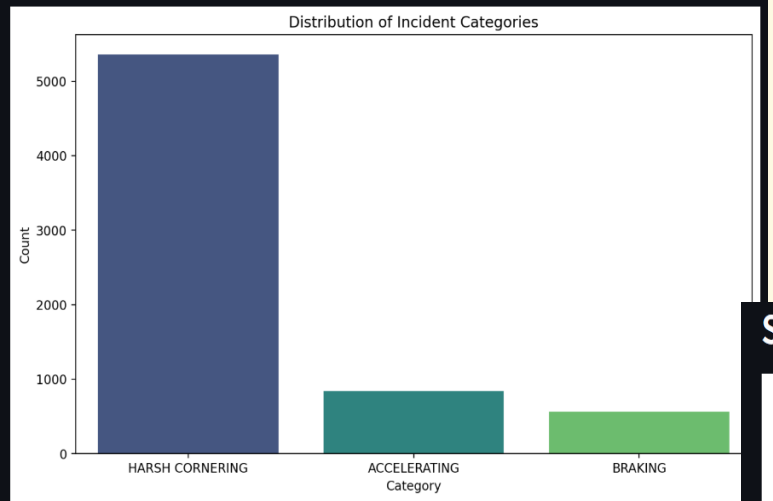


Heatmap for Incidents

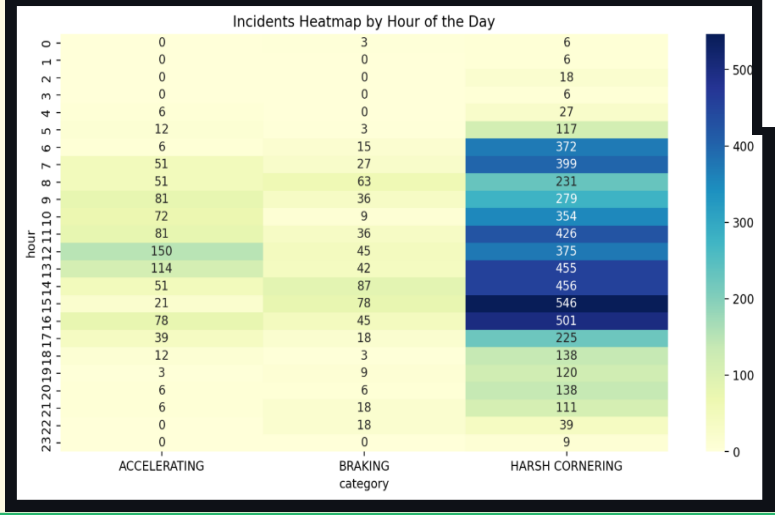


18 Deployment

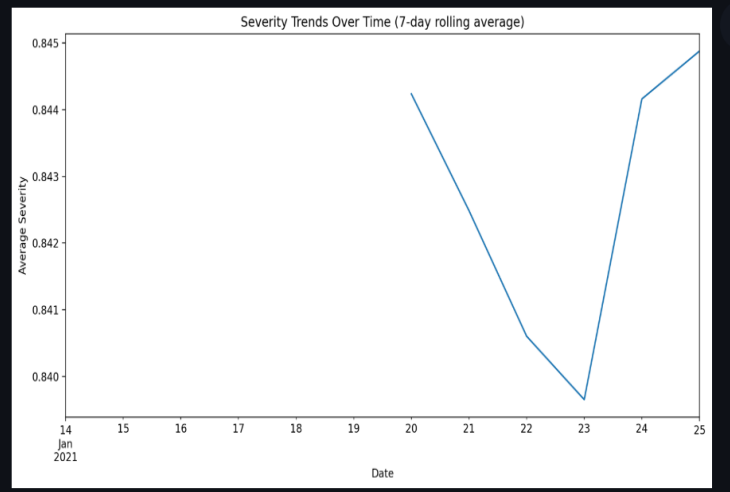
Category Distribution



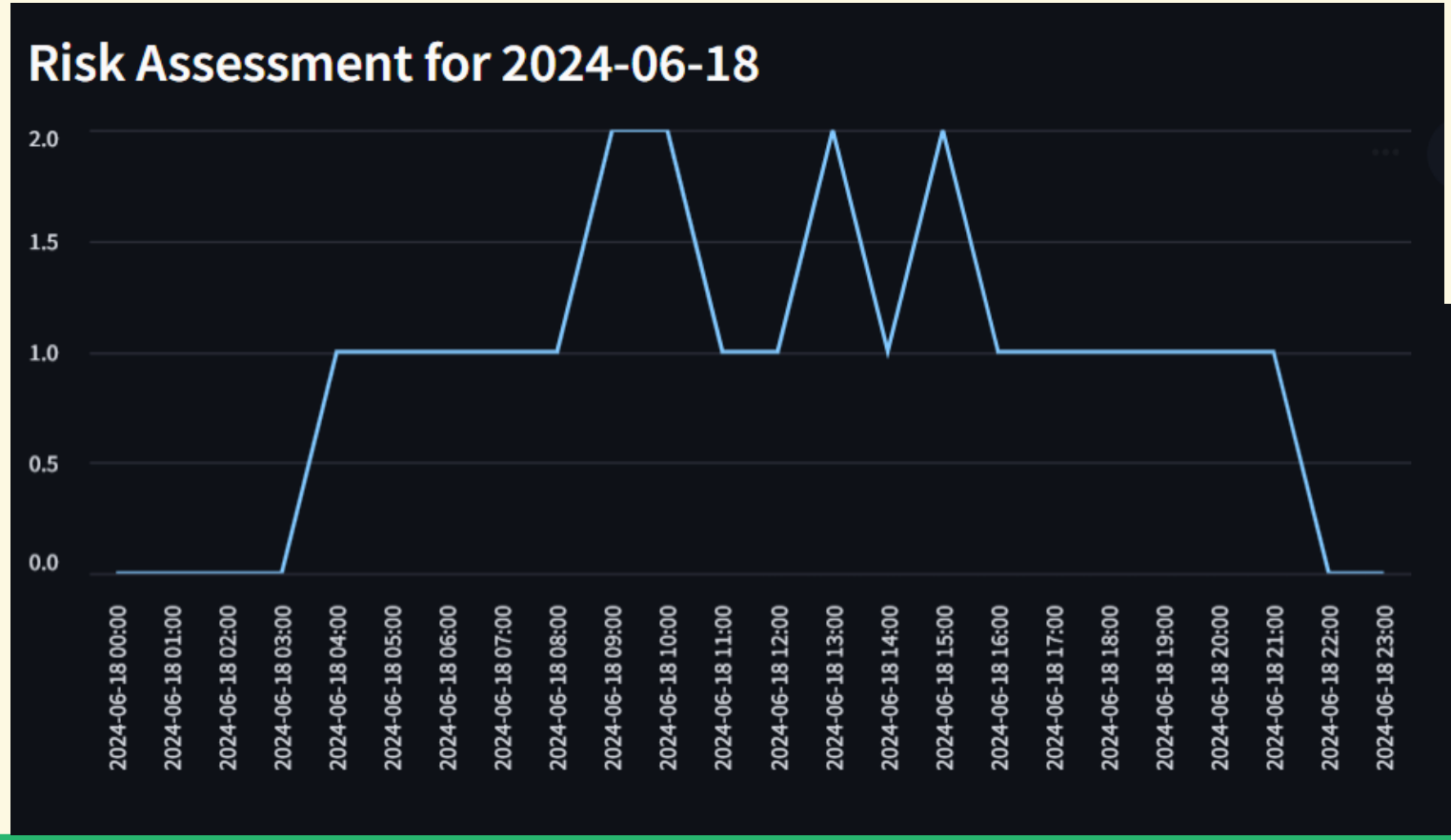
Incident Heatmap by Hour of the Day



Severity Trends Over Time



19 Deployment



The model predicts risk assessment for every hour.

Current Status for 2024-06-18 16:00: Normal level of Risk [0, 1]

Predictions DataFrame

	Time	Prediction
53	2024-06-20 05:00	0
54	2024-06-20 06:00	2
55	2024-06-20 07:00	2
56	2024-06-20 08:00	1
57	2024-06-20 09:00	1
58	2024-06-20 10:00	2
59	2024-06-20 11:00	2
60	2024-06-20 12:00	2
61	2024-06-20 13:00	1
62	2024-06-20 14:00	2
63	2024-06-20 15:00	1

Fetch New Data

20 Deployment

Current Status for 2024-06-18 15:00: Above average level of Risk [2]

At-risk locations with the most incidents and Severity above 2

Road name: Konijnenberg, **Incident count:** 3552, **Avg severity:** 2.13, Most common incident: Harsh Accelerating

Road name: Nieuwe Kadijk, **Incident count:** 2908, **Avg severity:** 2.02, Most common incident: Harsh Accelerating

Road name: Nieuwe Kadijk, **Incident count:** 2866, **Avg severity:** 2.02, Most common incident: Harsh Cornering

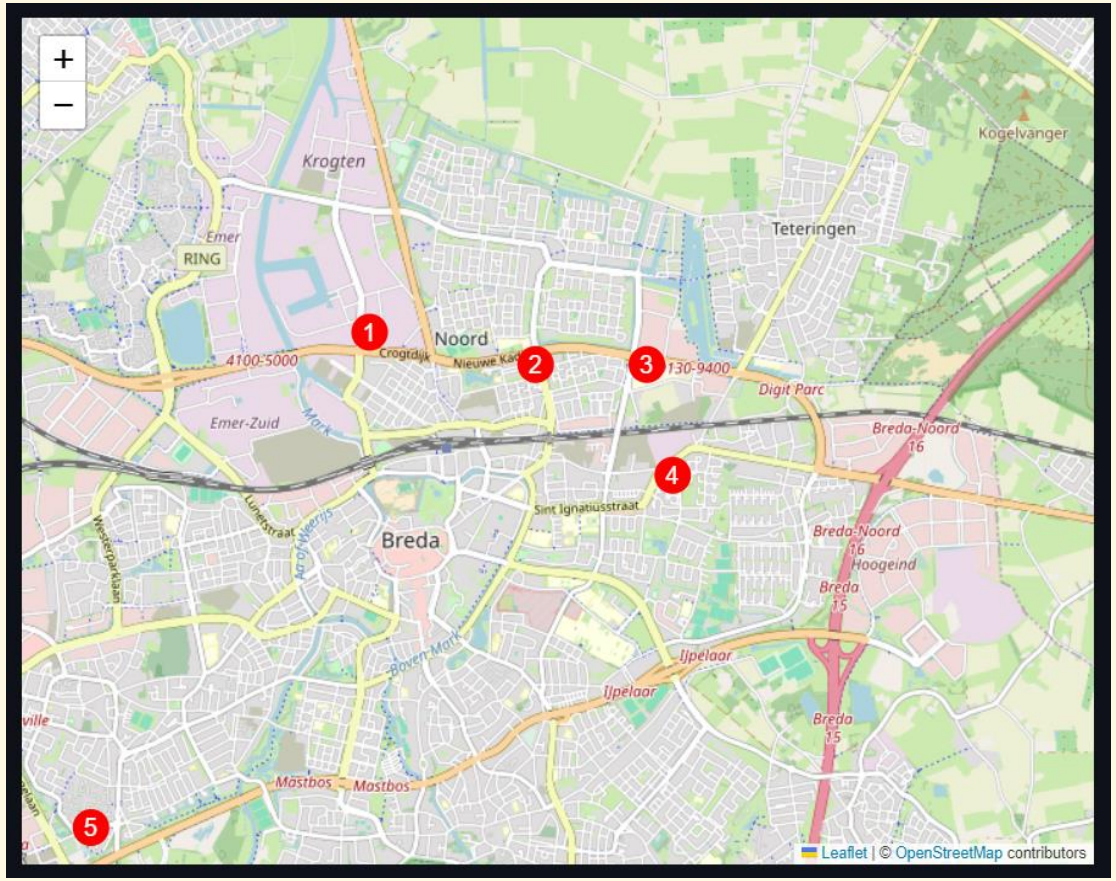
Road name: Teteringsedijk, **Incident count:** 2620, **Avg severity:** 2.00, Most common incident: Harsh Accelerating

Road name: Vogelschoot, **Incident count:** 2904, **Avg severity:** 2.00, Most common incident: Harsh Cornering

Road name: Kapelstraat, **Incident count:** 2996, **Avg severity:** 2.00, Most common incident: Harsh Cornering

When the model detects a high risk, the website highlights high-risk areas within the city of Breda.

21 Deployment



It also highlights this high risk areas on the map.

22 Our Team

Monika Stangenberg

Bartosz Kudyba

Kiril Baharev

Stijn van der Pas

Thank you