# Traffic Accident Prediction In Hamburg

## Accident Prophet

Andreas Kunzendorf, Felix Geyer, Carolina Körber, Till Meineke

16.10.2023

# OUR **TEAM**



**Felix Geyer**

Mathematician

https://www.linkedin.com/in/felix-geyer-ds/



**Andreas Kunzendorf**

Biochemist

www.linkedin.com/in/andreaskunzendorf/



**Carolina Körber**

Chemical Engineer

www.linkedin.com/in/carolina-mergulhao-körber



**Till Meineke**

Biochemist & Data Analyst

https://github.com/TillMeineke

"1.35 million individuals lose their lives in traffic accidents every year."

—**World Health Organization (WHO)**

"Traffic crashes totalled EUR 31.5 billion for German society in 2020."

—**ROAD SAFETY REPORT 2021**

Motivation | Data | EDA | Modeling | Evaluation | Dashboard | Summary

# MOTIVATION &
# **PRODUCT**

- Predict traffic accidents in Hamburg

- Resource planning for police and hospitals

- Reduce time for police to arrive on accident
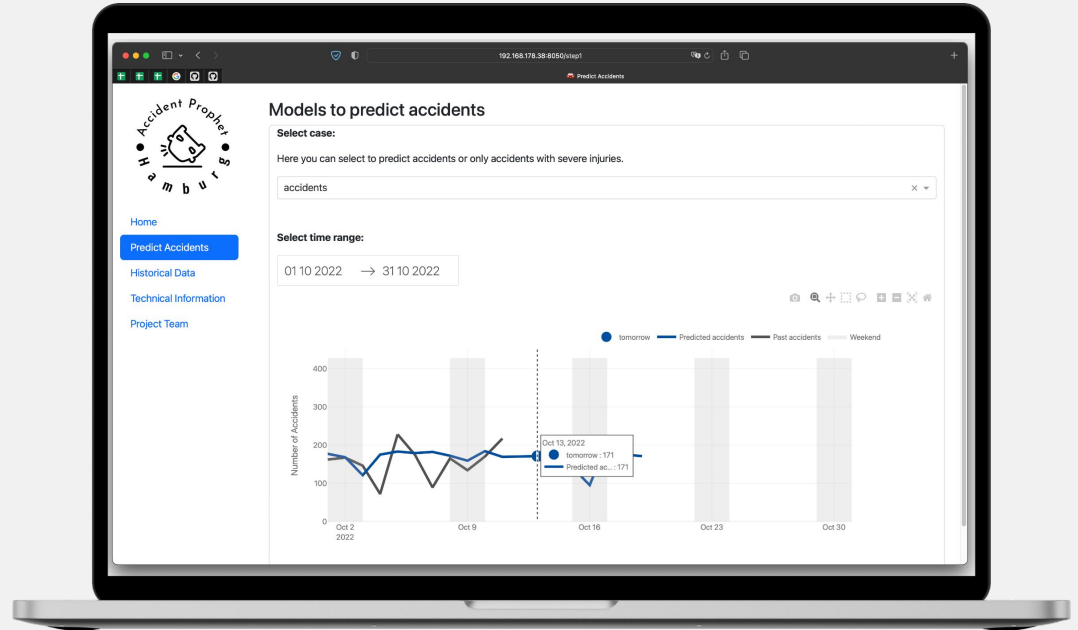
# DATA **SOURCES**

## Accidents Data

- 2009 - 2022
- Accidents: 909273
- Features : 47
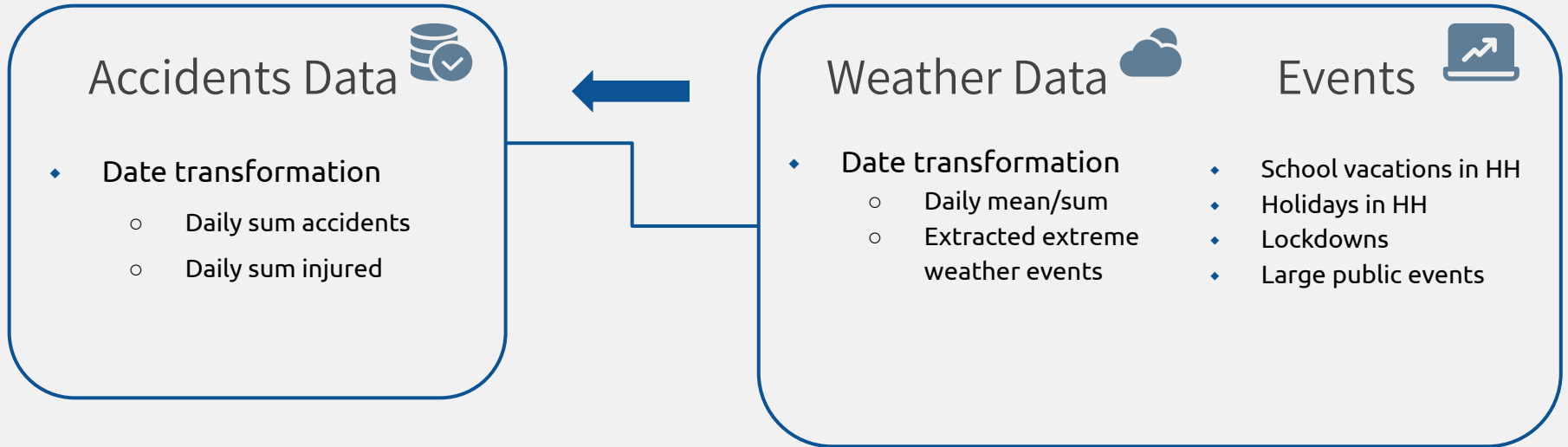- Date, time, location, age, no. injured

## Weather Data

- Numerical weather data
- Weather phenomena

## Events

- School vacations in HH
- Holidays in HH
- Large public events

# DATA **ENGINEERING**

## Accidents Data

- Date transformation
  - Daily sum accidents
  - Daily sum injured

## Weather Data

- Date transformation
  - Daily mean/sum
  - Extracted extreme weather events

## Events

- School vacations in HH
- Holidays in HH
- Lockdowns
- Large public events

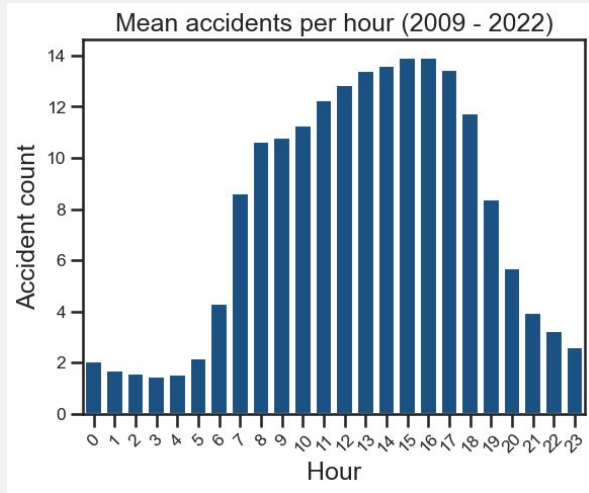# TRAFFIC ACCIDENTS **PER YEAR**

**+65000**

ACCIDENTS

in Hamburg
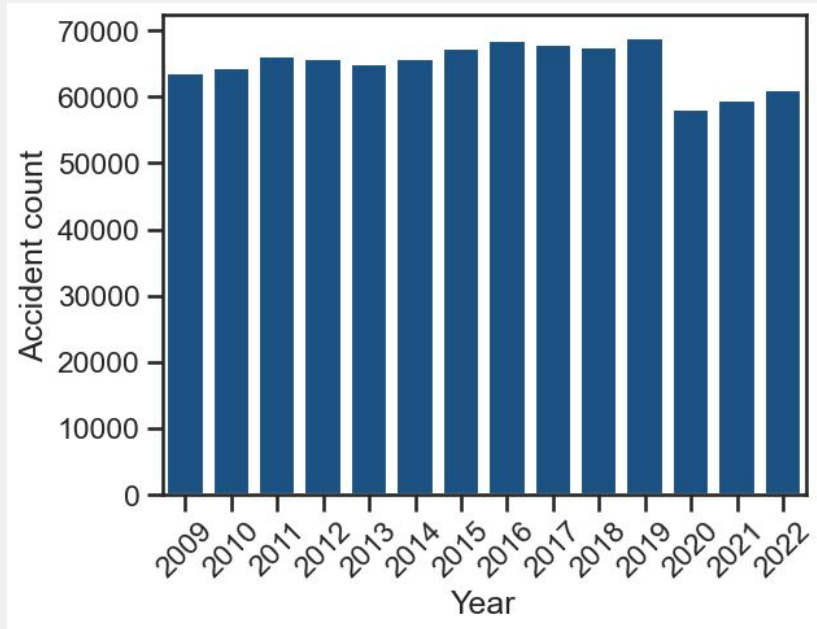
# TRAFFIC INJURIES **PER YEAR**

**+8500** light
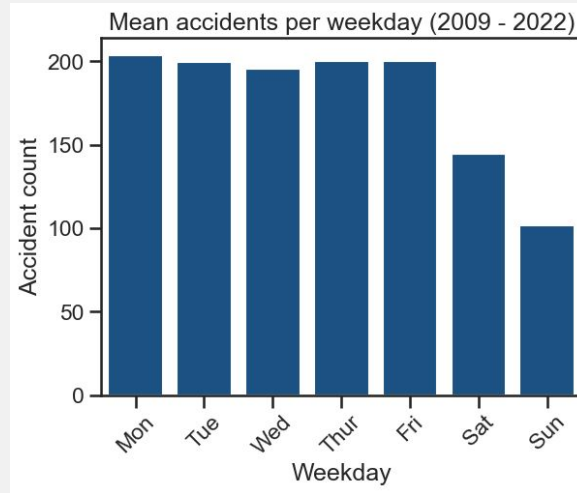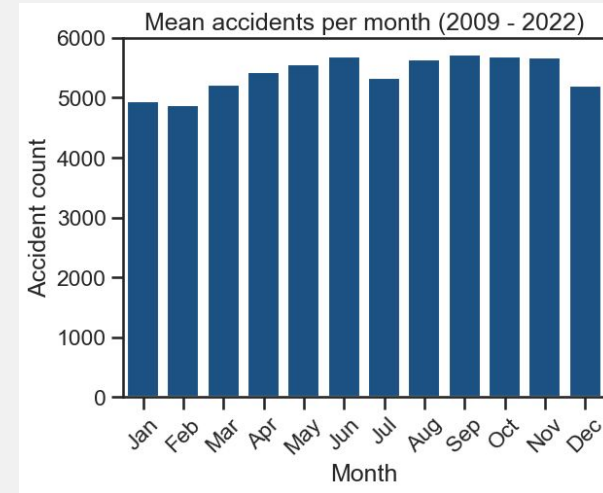**+800** severe
INJURIES

in Hamburg
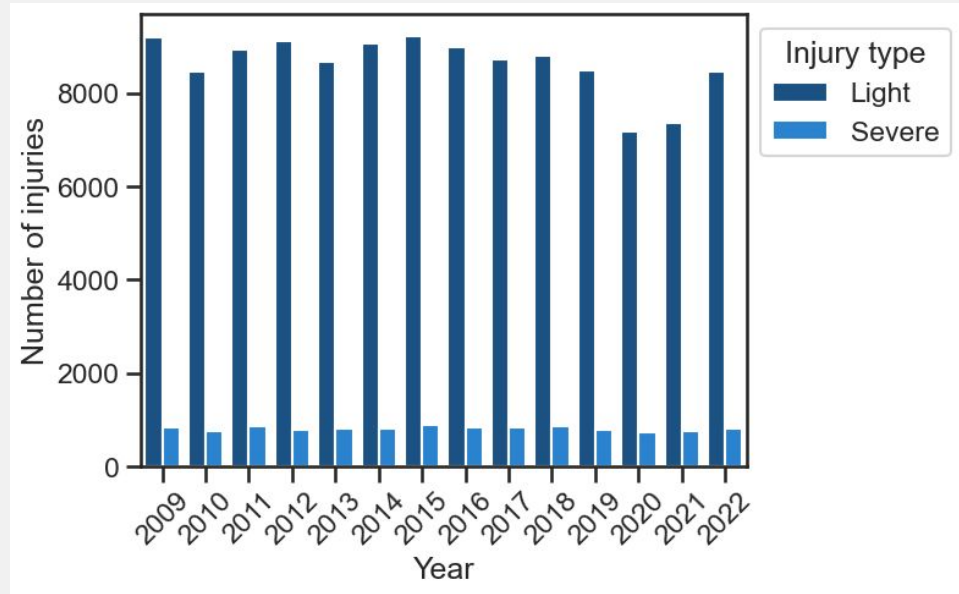
# TIME **PATTERNS**

Hour



Weekday



Month

# COVID-19 **IMPACT**



Accidents per month
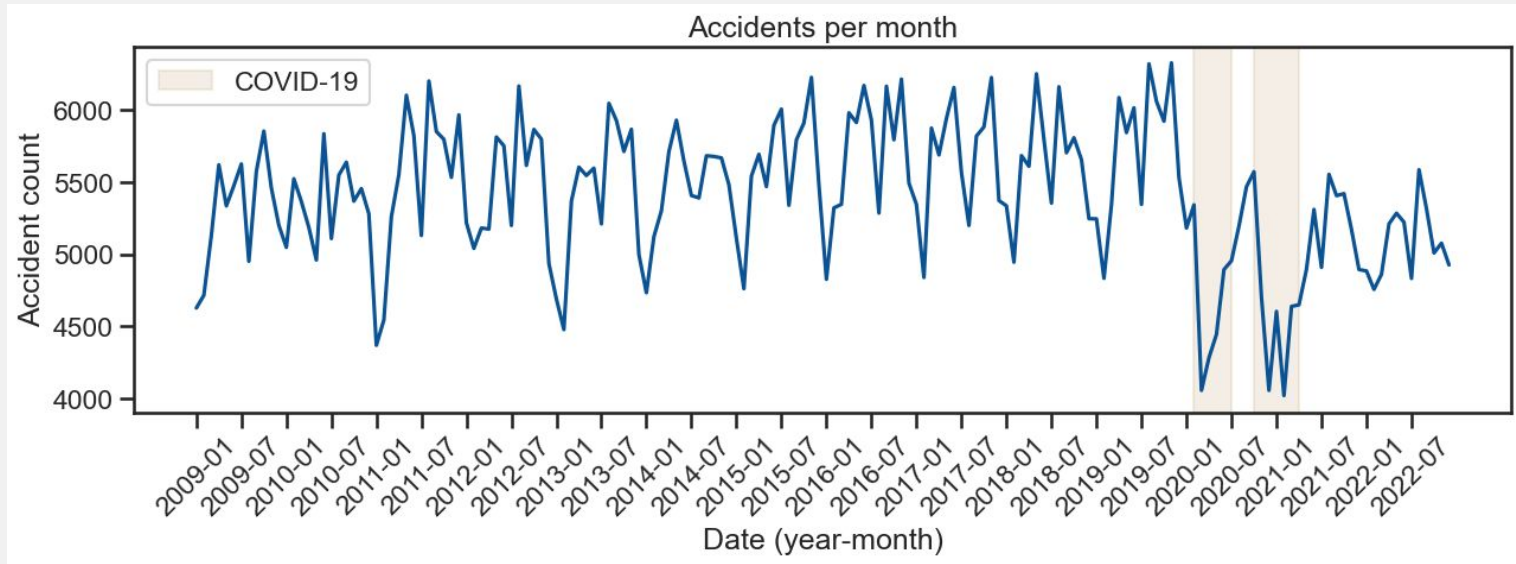
# PREDICTION **MODELS**

## SARIMAX

- Classic time-series model
- Autoregressive component (AR)
- Moving average (MA)
- Seasonality
- Additional features

## Prophet

- Time-series model
- Additive regression model
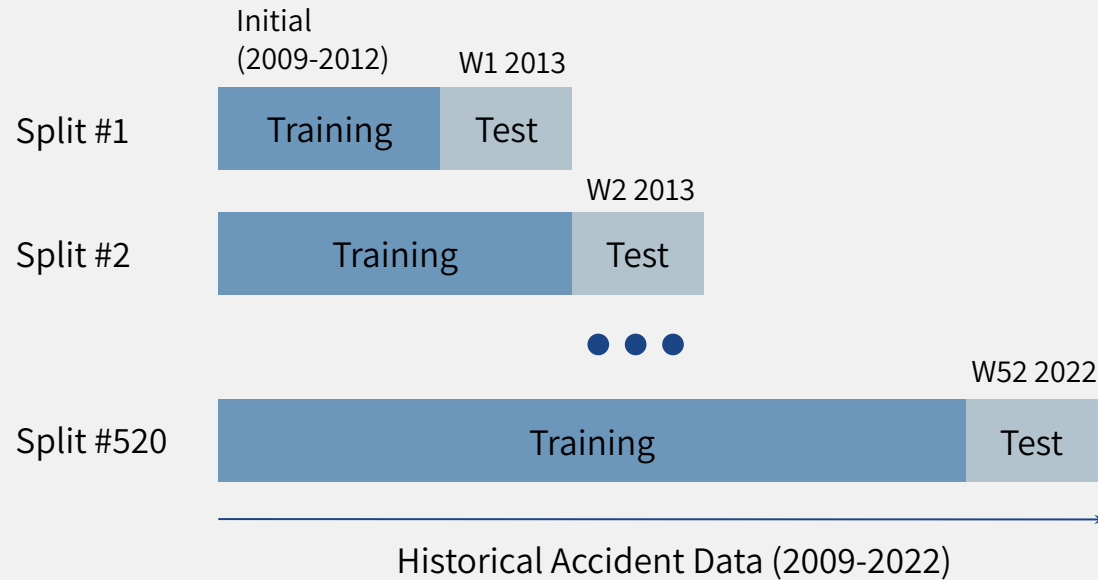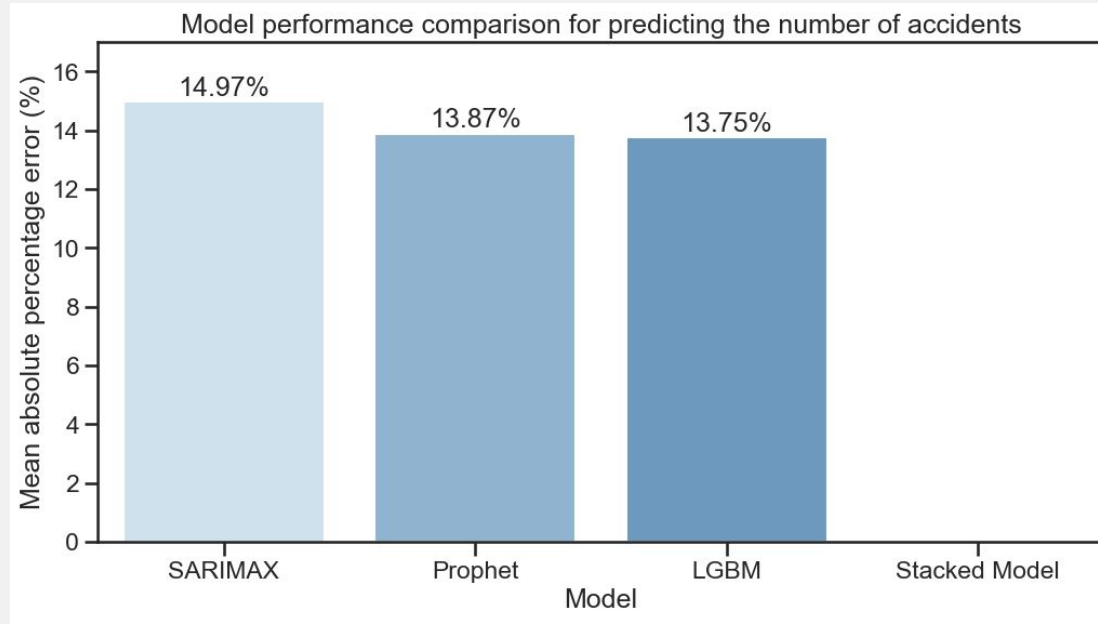- Trend, Seasonality, Holidays
- Additional features

## LightGBM

- Gradient boosting model
- Tree based learning

# EVALUATION **METHOD**

- Cross validation with weekly splits

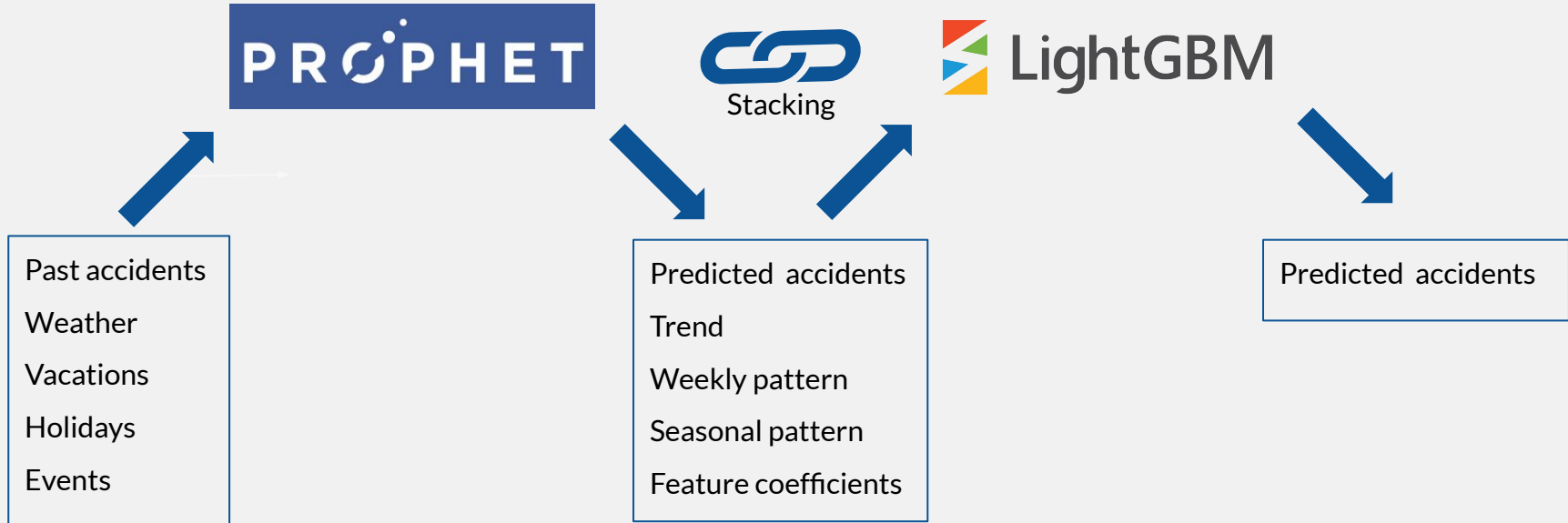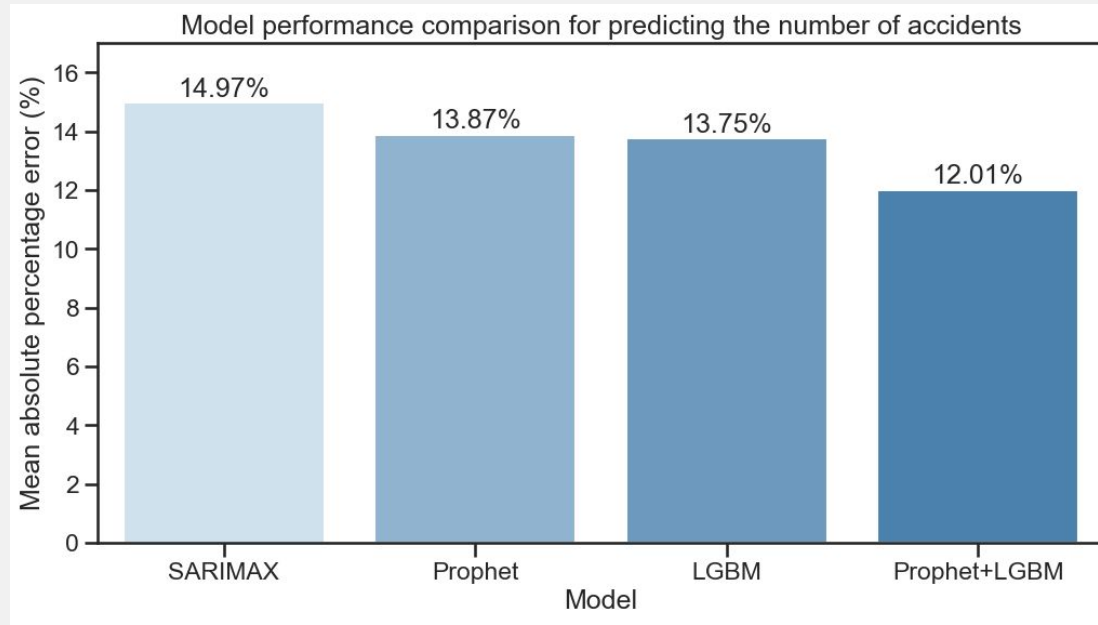- Mean absolute percentage error (MAPE) as metric



Historical Accident Data (2009-2022)

# MODEL **EVALUATION**



Model performance comparison for predicting the number of accidents

# FORECAST **MODEL**

# FORECAST **MODEL**



Past accidents
Weather
Vacations
Holidays
Events

Stacking

Predicted accidents
Trend
Weekly pattern
Seasonal pattern
Feature coefficients

LightGBM

Predicted accidents

# MODEL **EVALUATION**



Model performance comparison for predicting the number of accidents

# Tuning and **EVALUATION**

**Tuning**

- Best regressor combinations chosen by cross validation
- Hyperparameter tuning

**Evaluation**

- Only 12% MAPE for predicting the number of daily accidents
- Significant drop of prediction quality during Covid-pandemic
- 29.6% MAPE for predicting the number of daily injured people

Mean absolute percentage error by year for predicting the number of accidents
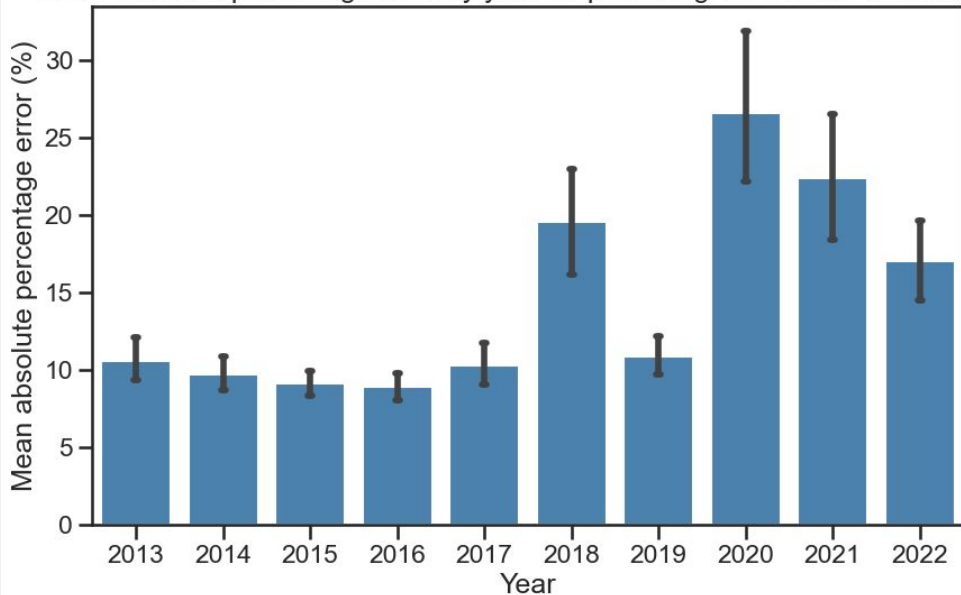
# DASHBOARD **LIVE DEMO**
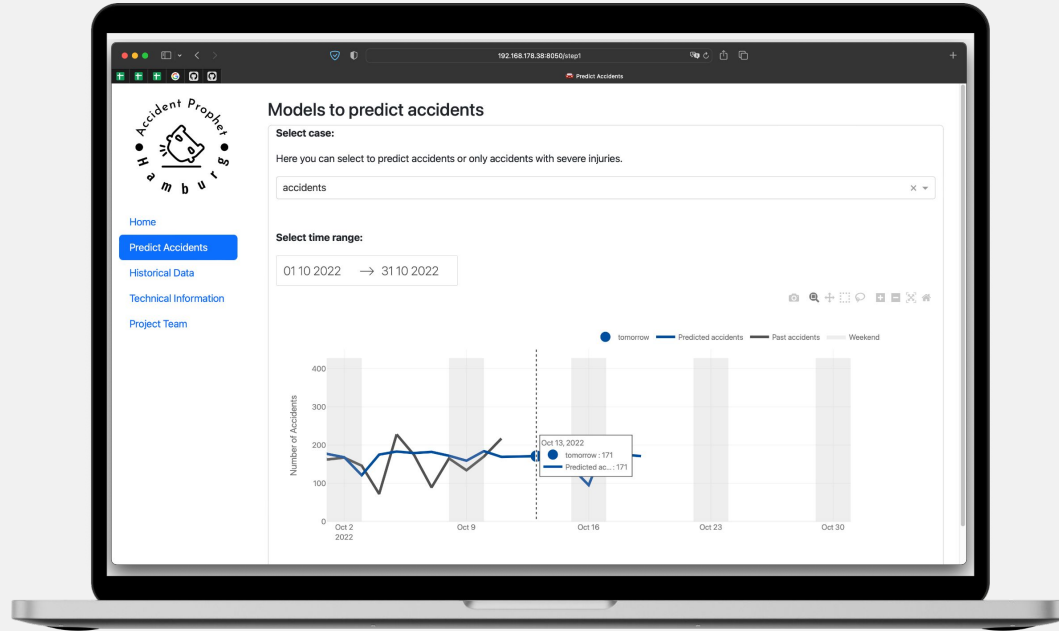
# FUTURE **WORK**

Implement real-time data
(weather, traffic etc.)

Improve model,
try more sophisticated models

User feedback

Monitoring of dashboard effect

# THANKS!

Do you have any questions?



**Felix Geyer**

Mathematician

https://www.linkedin.com/in/felix-geyer-ds/

**Andreas Kunzendorf**

Biochemist

www.linkedin.com/in/andreaskunzendorf/

**Carolina Körber**

Chemical Engineer

www.linkedin.com/in/carolina-mergulhao-körber

**Till Meineke**

Biochemist & Data Analyst

https://github.com/TillMeineke