Problem Statement



Data Acquisition

Team



Evaluation



Business Value



Congested roads and slow-moving traffic are problems faced by commuters in Darmstadt. The existing challenges necessitate innovative solutions to optimize traffic flow, reduce environmental impacts, and enhance overall transportation efficiency in the city.

Through our application, historical data can be used

to predict future traffic flow. As such, this

application can provide insights for traffic

department staff and help them improve urban

For instance, critical junctions can be identified

early on before congestion occurs, traffic can be

analyzed over long-term for urban planning and to

improve traffic concepts, redirections for

construction sites or events can be planned, and traffic light signals can be balanced which leads to

transportation in various ways.

"Datenplattform Darmstadt" provides traffic data in CSV format for one month, an overview map of all intersections, a detailed map for each intersection showing the location of sensors, and a JSON file containing sensor configuration data.

An extended dataset of two years was made available after a mail request.

Input: Traffic sensor data from "Datenplattform

Output: Number of cars per hour (past / present /

Methodology: Machine learning models: Prophet,

Regression for time-series data was implemented.

and predictive performance.

Success Criteria

Evaluation was carried out using a combination of

A regulated flow of traffic increases the safety and well-being of all road users and residents. error metrics: MAE, MSE, RMSE, R2, SMAPE, and additionally, a confusion matrix (for the congestion Unlike in other services, the end user is a control categories). This approach allows for a nuanced evaluation of the model's accuracy, error magnitude.

center agent from the road authority of Darmstadt. Monitoring and prediction are key in planning the future traffic flow. Agents can determine other ways to organize traffic and difficult traffic conditions beforehand. This can be applied to road planning and instantly adjusting traffic lights.

Utilizing the city sensor data allows precise recording of the traffic situation which results in more accurate future predictions.

Solution



Analytics Formulation

Random Forest, LightGBM, CatBoost.



Objective: the model performs well with respect to the evaluation metrics. It identifies congestion in historical data and provides reliable predictions.

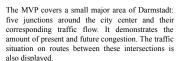
Subjective: users can easily recognize critical moments in the current and future traffic situation. Furthermore, there is a user interface representation that shows the traffic in different colors, based on the amount of congestion, and provides an accurate traffic prediction. The predictions can keep up with and outrun predictions from other tools such as







Google Maps and Waze.



Modeling



number of cars per hour. For this purpose, we evaluated the following models: Prophet, LightGBM, CatBoost, and Random Forest. Based on the metrics. Random Forest was deemed as best to use for our application.

After prediction, data was binned to different congestion categories (low / high / medium).

Constraints We performed time-series data prediction for the



Kev Actors



Updating the model with real time data is a constraint. Unpredictability exists due to regional factors like strikes. Also, there are variations in the sensors. Sensor hardware failure leads to imputation of data which leads to falsification of the prediction. In addition, there is ambiguity in the information about the dataset, sensors, and their positions.

Internal: traffic sensor data

Customer stakeholders: traffic department of Darmstadt (and other cities), vehicular commuters. general public

External: employees from the mobility office or road planning department of Darmstadt.

Users & Use

less slow-moving traffic.



Our application is aimed at control center agents monitoring and controlling traffic in Darmstadt. Moreover, it offers support in recognizing and planning critical situations. This relieves and supports them in their work.

Indirectly, road users and residents of Darmstadt will benefit the most, as they will be able to move around in the city more quickly and safely and plan routes reliably.

Data Preparation



First, we had to gain deep understanding of the dataset (rudimentary / missing explanation), then data cleansing was performed e.g. intersections that are out of service due to roadworks, merging data from various intersections, connecting which sensors are installed where and at which intersection, and interpolation for missing values.

Technology stack



Backend: Python and Python libraries (including numpy, pandas, sklearn, prophet, catboost, lightgbm, matplotlib)

Frontend: Figma, Anvil, Python

Project Management: Trello, Google Meet, Telegram

Development: Github, Jupyter Notebook, Google Colab, Visual Studio Code