flaschenpost

Precision in Prediction: Smarter Delivery Estimates

STADS Datathon 2025 – Team 7

Our Team

Simon
Schumacher

Emil
Schallwig

Paul
König

Elise
Wolf

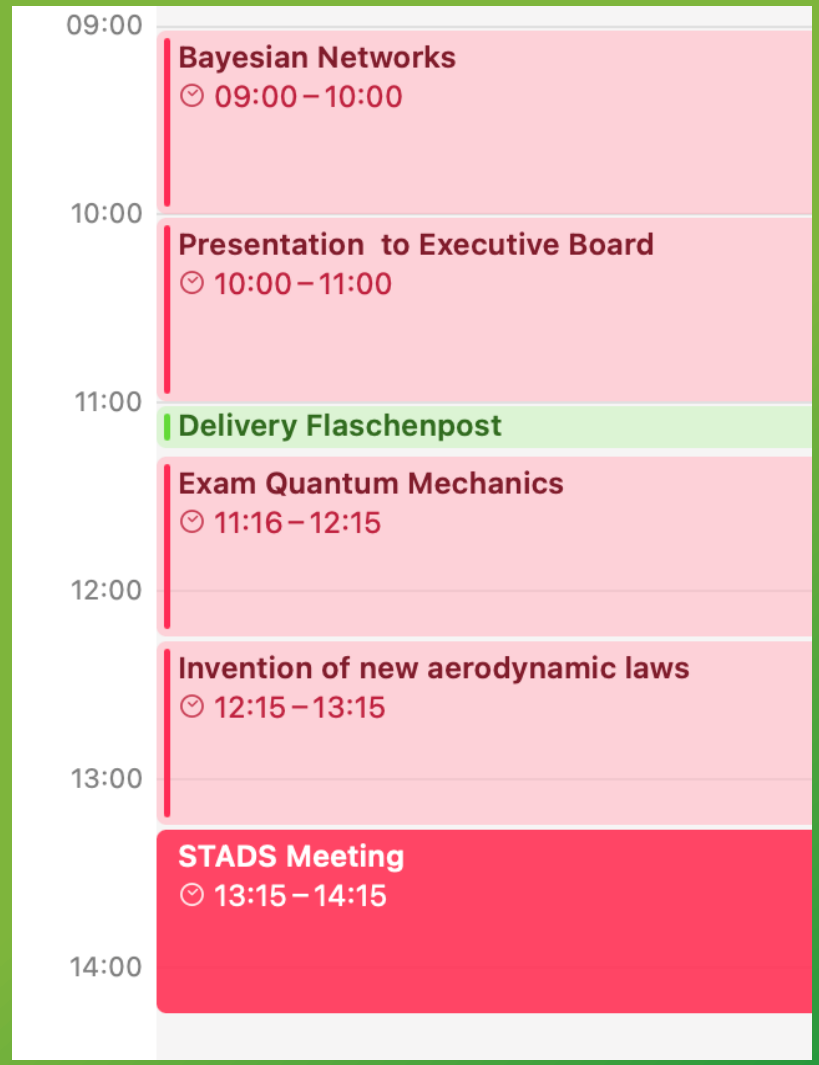
Making sure you get your groceries on time!

Agenda

- Challenge
- Approach
- Our Solution

EXPECTATION

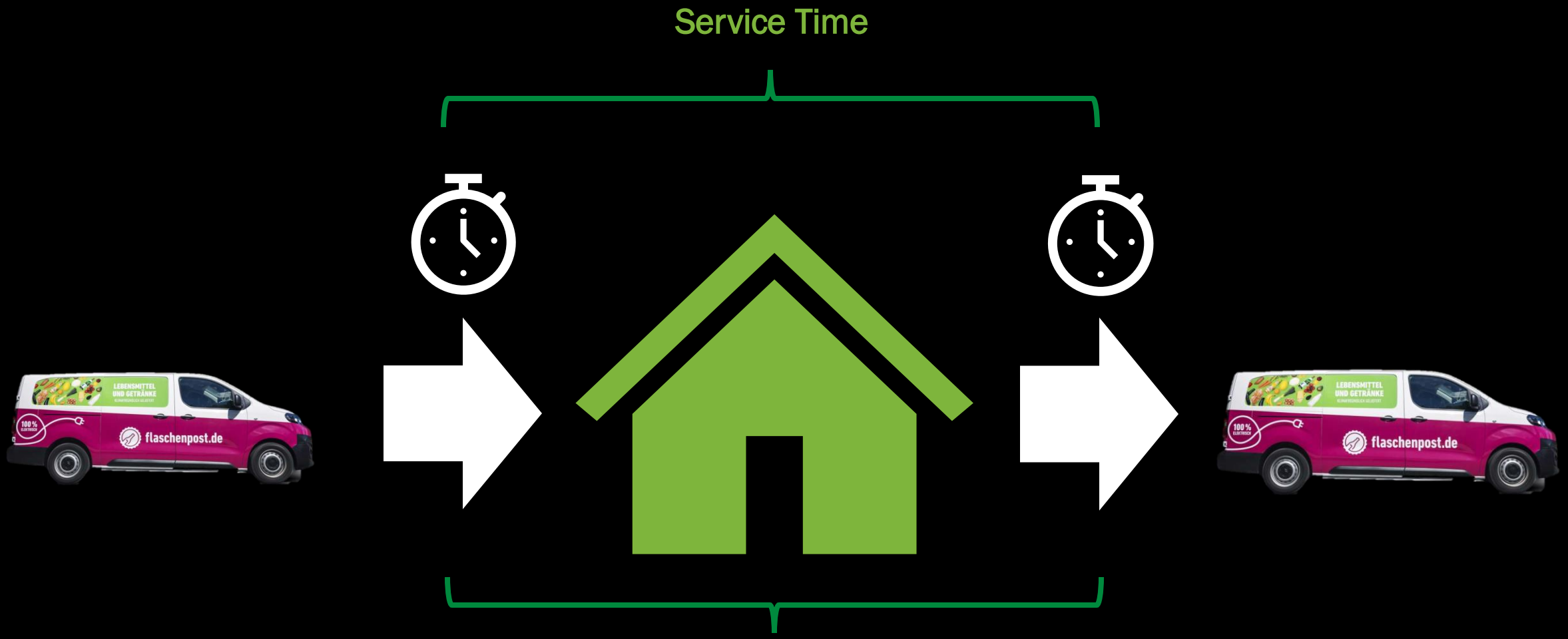
Flaschenpost's Business Model



A typical day in our lives...

WHY

Motivation



Too low → Customer Dissatisfaction

Too high → Loss of Efficiency

Data on hand

- Number of **stories**
- Existence of an **elevator**
- **Business** order
- **Weight** of the items

- **Arrival time** of the driver
- Flaschenpost **Warehouse**
- **Driver** ID
- **Customer** ID

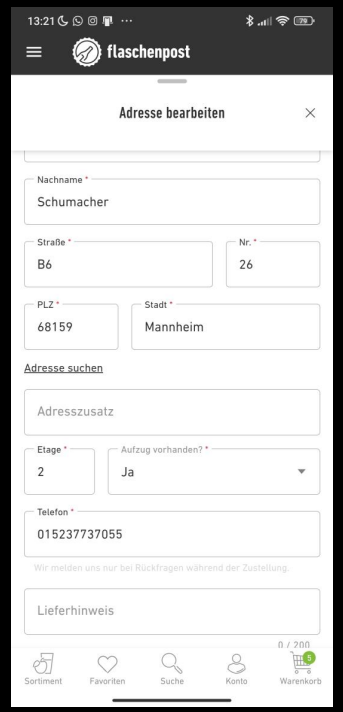
Agenda

Challenge

Approach

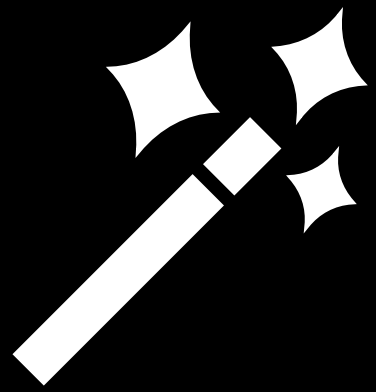
Our Solution

Feature Engineering



Preprocess available features...

... aggregate information ...



... and extract new features.

Enhancing data through Business Understanding

Experience of the drivers

Number of crates and boxes to be delivered

Worker's growing fatigue

Recurring customers

Removing outliers

Enhancing data through Business Understanding

Experience of the drivers

Number of crates and boxes to be delivered

Worker's growing fatigue

Recurring customers

Removing outliers

Enhancing data through Business Understanding

Experience of the drivers

Number of crates and boxes to be delivered

Worker's growing fatigue

Recurring customers

Removing outliers

Enhancing data through Business Understanding

Experience of the drivers

Number of crates and boxes to be delivered

Worker's growing fatigue

Recurring customers

Removing outliers

Enhancing data through Business Understanding

Experience of the drivers

Number of crates and boxes to be delivered

Worker's growing fatigue

Recurring customers

Removing outliers

Challenge

Approach

Our Solution

Our Models

We experimented with a diverse set of models…

Shared Datasets

Linear Regression
Random Forest
XGBoost
BART
LinexXGB
LightGBM
Deep Gaussian Process
HierarchicalBayes
Neural Network
Bayesian Ridge Regression

MAE
MSE
R2
Confidence Interval

Streamlined preprocessing…

… applied to different models …

… and shared evaluation.

Our Results

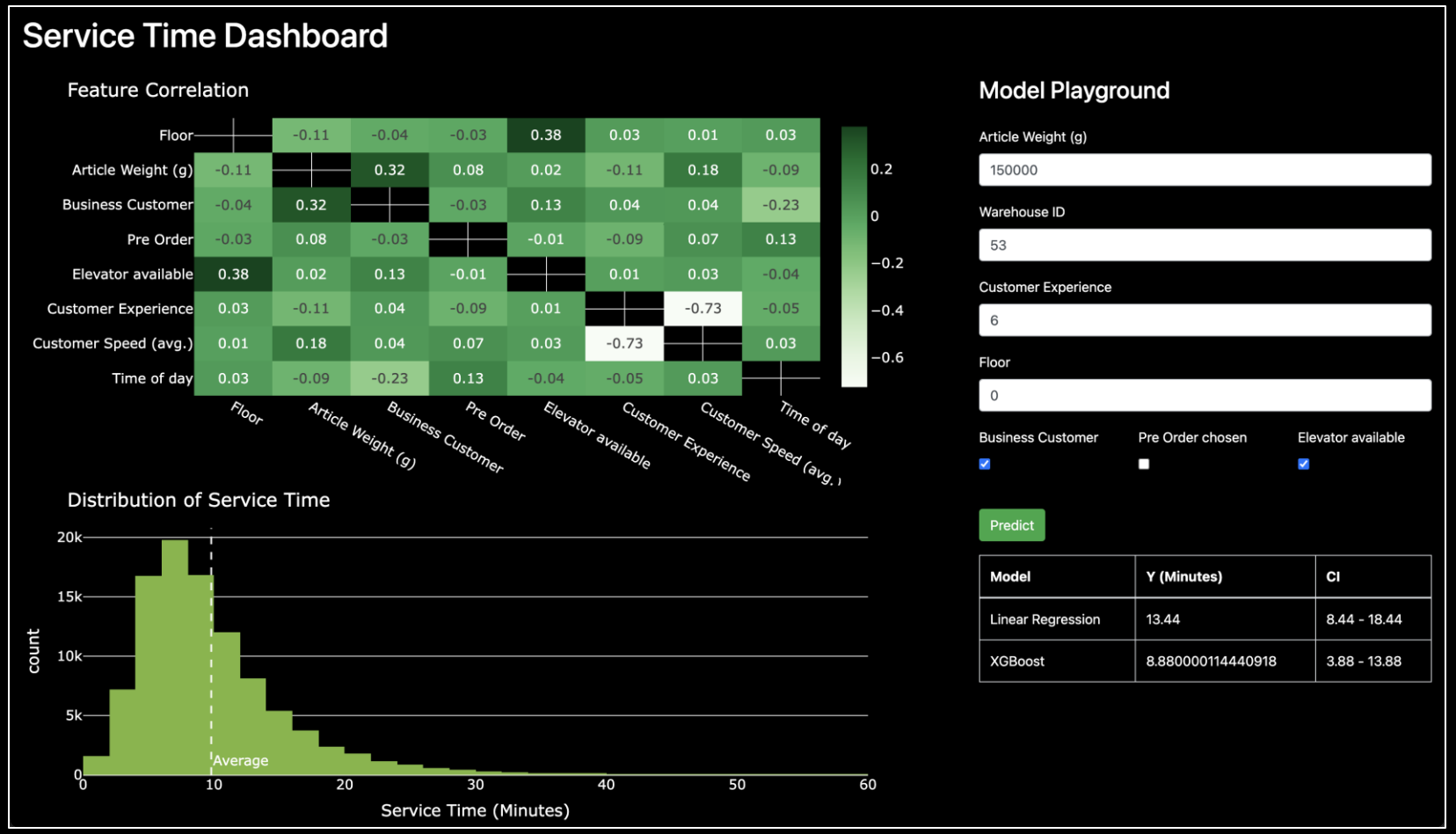
… that achieved similarly low MAE values - substantially outperforming our baseline predictions."

Model	MSE	MAE	R2	Confidence Interval
Baseline	36.965	4.330	--	(-9.801; 9.801)
XGBoost	**11.299**	2.316	**0.769**	**(-9.759, 9.787)**
Neural Network	26.658	3.5201	0.279	(-9.469, 9.493)
Bayesian Ridge Regression	26.932	3.579	0.271	(-9.714, 9.738)
Linear Regression	26.995	3.582	0.270	(-9.718, 9.742)
LightGBM	133.025	9.801	-2.599	(9.777; 9.824)

OUR SOLUTION

Service Time Dashboard

We help the Flaschenpost experts by providing them with all information at hand…



Next Steps

New Features based on additional Data

Integrate additional or external data (e.g., weather, traffic) to derive novel features that boost model performance.

Value Analysis of Customers

Quantify customer value by analyzing historical behaviors to drive targeted retention strategies.

Anomaly Detection of Drivers

Identify unusual driver performance through advanced anomaly detection to mitigate risks.