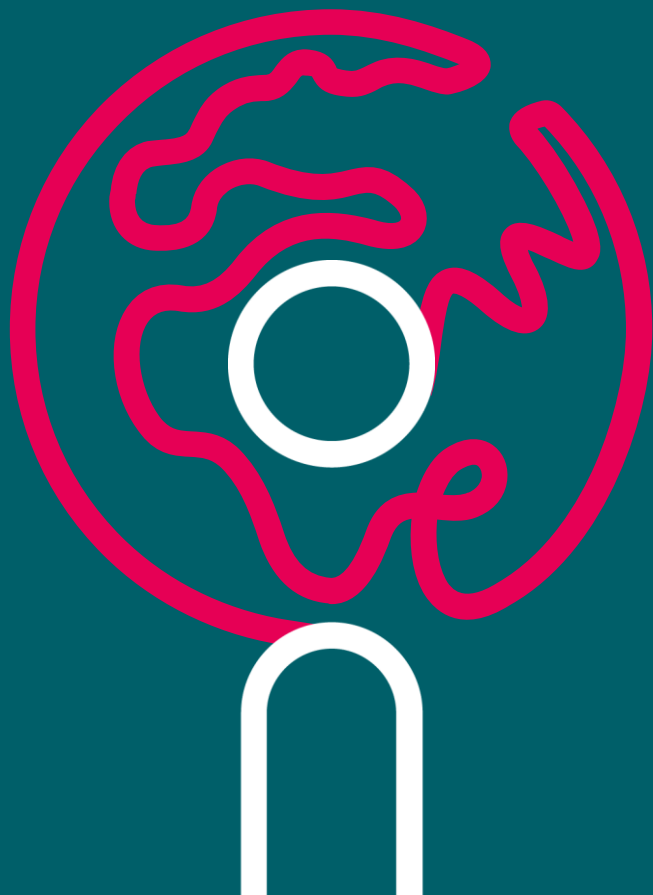




innogy

Data Science Anwendungen bei innogy – Netz & Infrastruktur (G&I) - Grid Research and Insights Platform (GRIP) -

Essener Energiegespräche

25. Oktober 2018

Philipp Clasen, Freier Mitarbeiter

1

Einleitung

Data Science bei innogy G&I

2

Data Science Terminology:

Beyond the buzzwords

3

Use Case 110 kV Transformers:

Predictive maintenance



4

Use Case e-prognosis:

Prognose von Ladesäulen

5

Use Case Stromausfall.de:

Web analytics

6

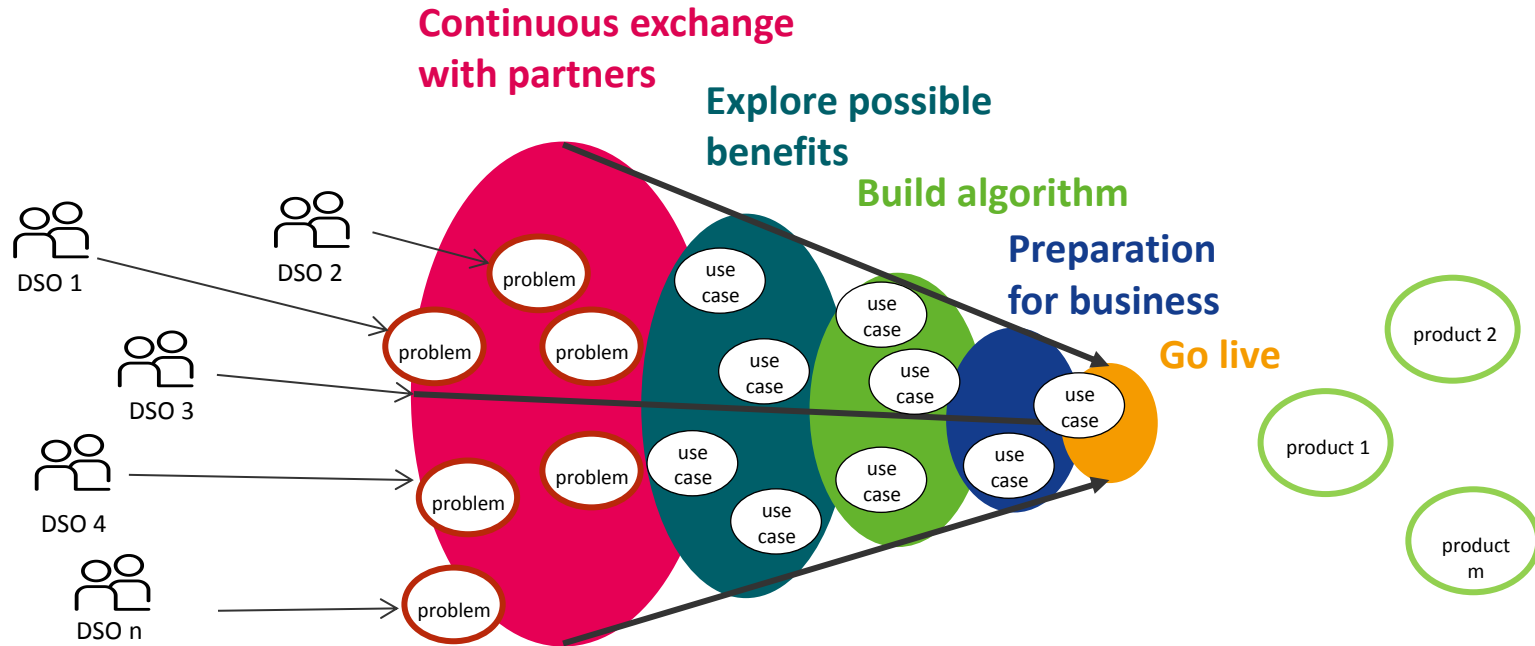
Fazit



EINLEITUNG

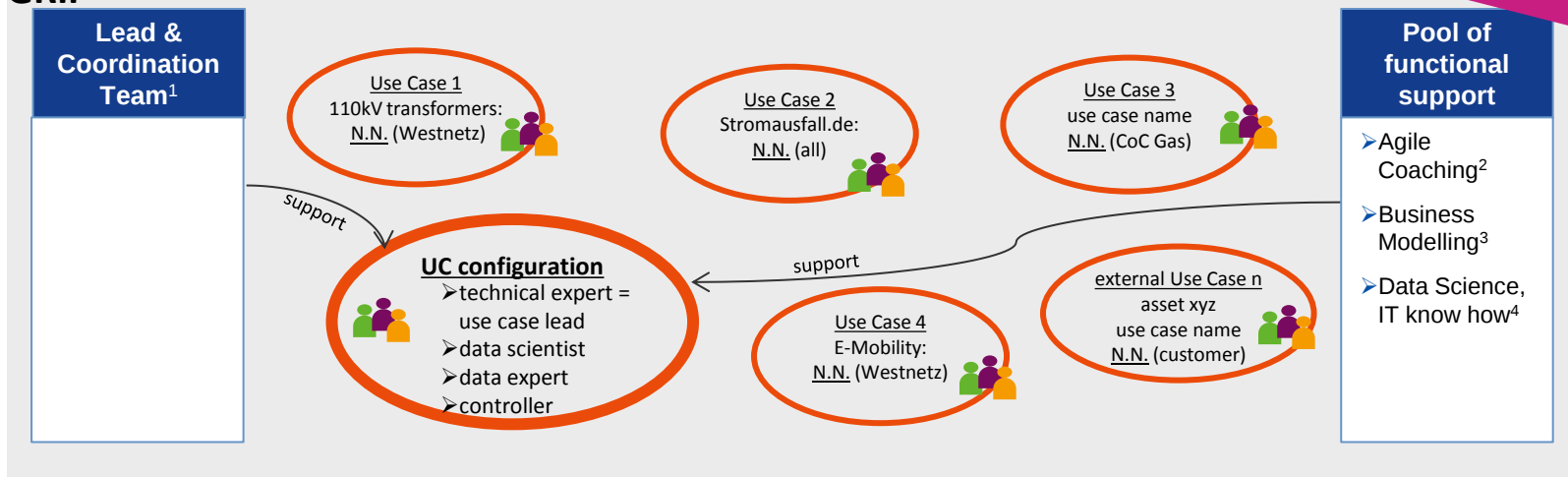
Data Science bei innogy G&I

GRIP – Grid Research and Insights Platform



Lighthouse GRIP: Overview of Project Structure

GRIP



Tasks of teams:

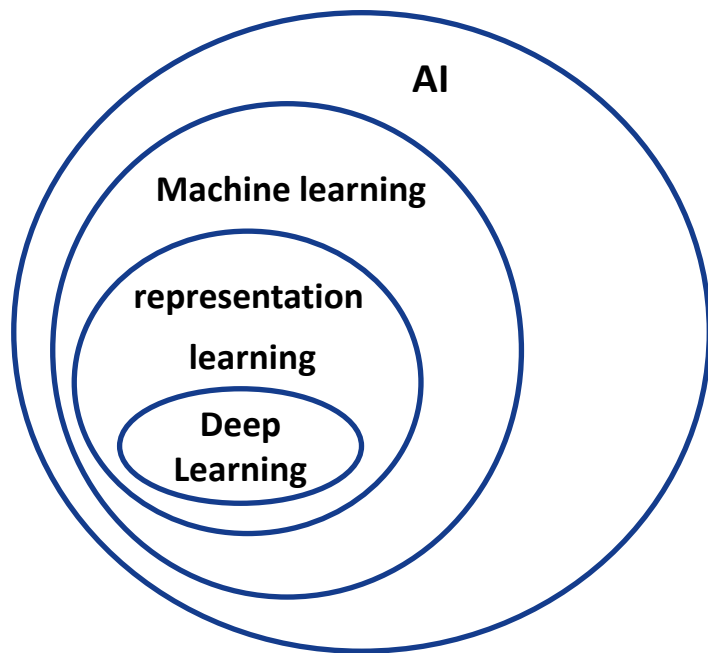
1. Organization and coordination of Use Cases, stakeholder management, budget, quality management
2. Support in agile working methods
3. Buildup of business models, strategic junction of use case value and company's strategy
4. Provision of additional know how (if necessary) like data scientists, data experts, IT know how



DATA SCIENCE TERMINOLOGY

Beyond the buzzwords

Artificial Intelligence (AI)



“The **true challenge to AI** proved to be solving the tasks that are easy for people to perform but hard for people to describe formally

—

problems that we solve intuitively, that feel automatic, like recognizing spoken words or faces in images”

Learning Tasks

Machine Learning = learn a task from experience

Task

z.B. handschriftliche Wörter aus Bildern erkennen

Performance measure

z.B. % richtig erkannter Wörter

Experience

z.B. Datenbank mit Bildern
handschriftlicher Wörter (und Beschriftung)

Kleiner Exkurs in die Botanik



Iris Versicolor



Iris Setosa



Iris Virginica

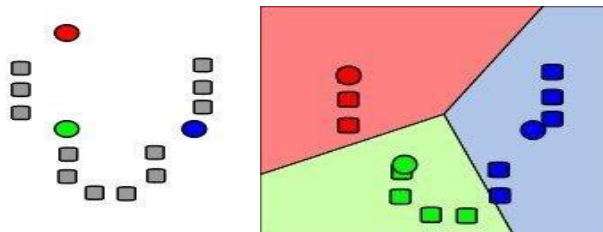
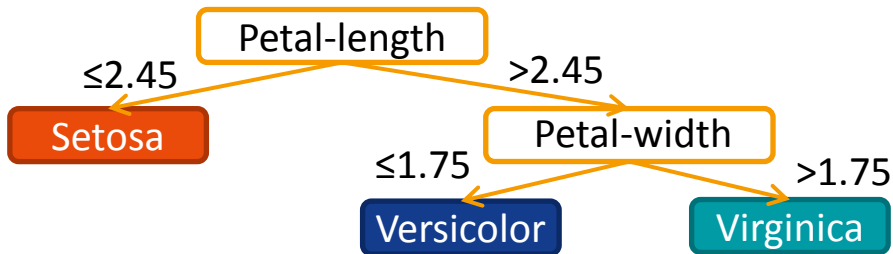
Machine Learning – A Simplified Overview

Supervised Learning:
„Learning a Function“, e.g.

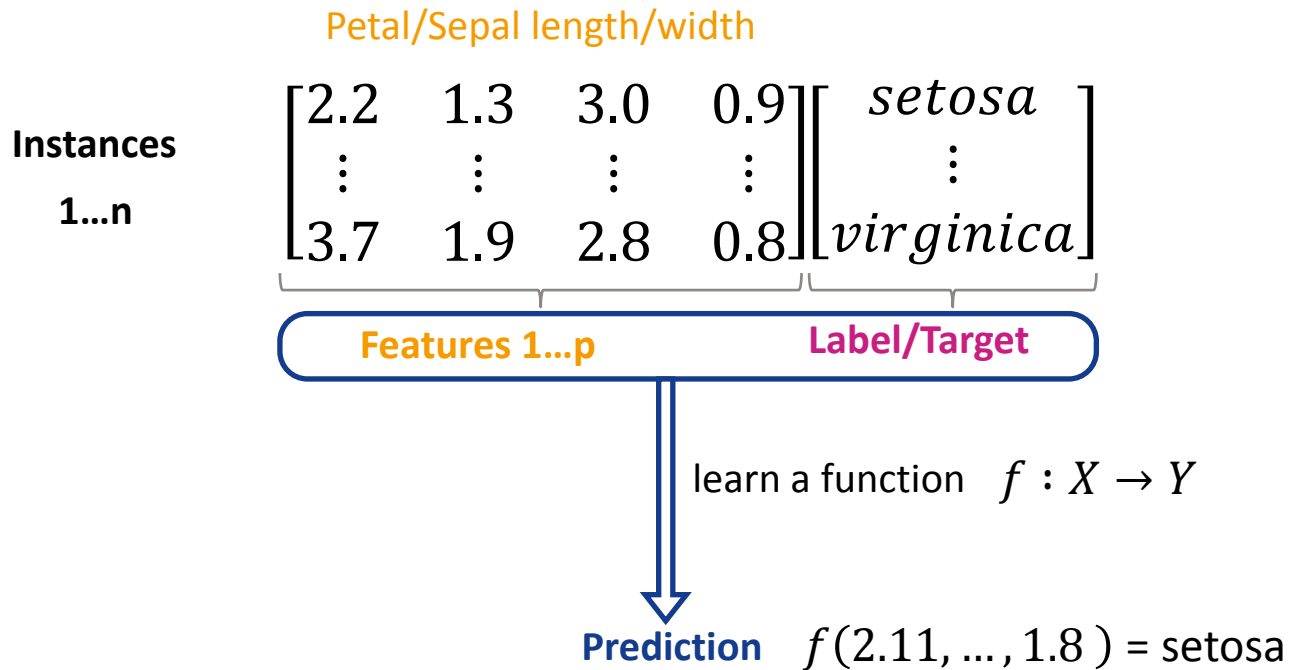
- Regression
- Classification

Unsupervised Learning:
„Finding Patterns“, e.g.

- Clustering
- Dimensionality Reduction



Supervised Learning



How to Learn a Function

Data

+

Inductive Bias (assumptions)

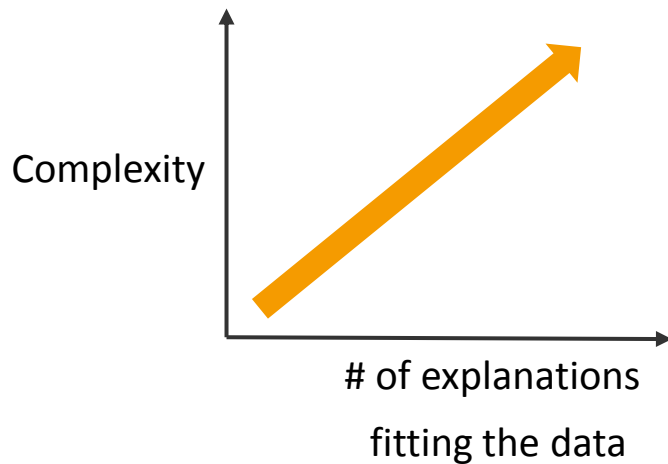
➤ Hard Bias

e.g. only consider linear functions

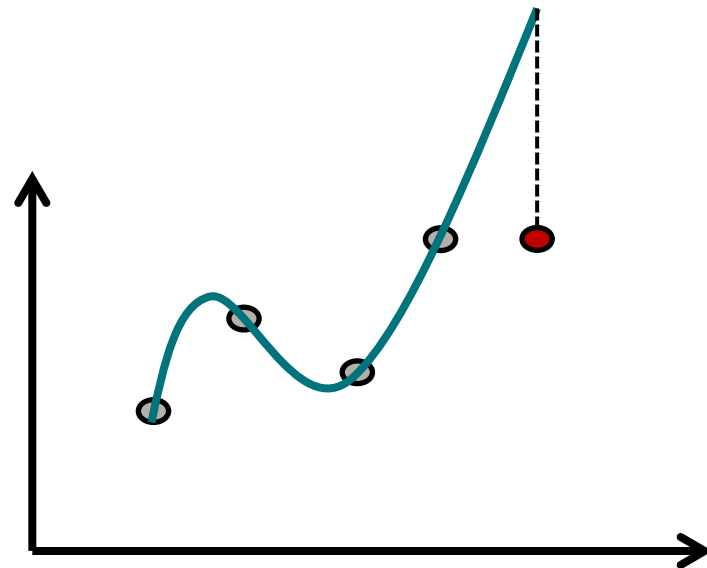
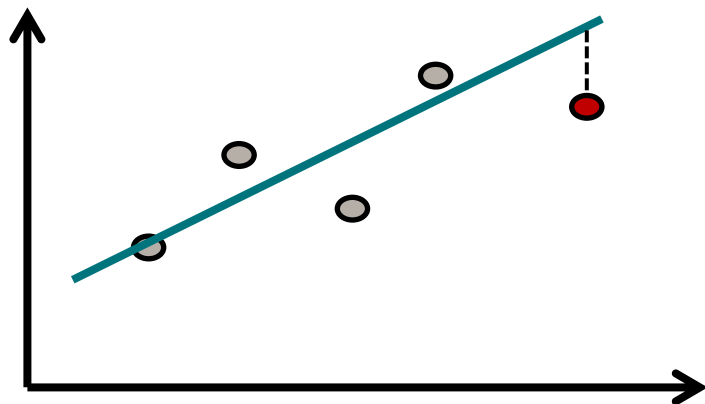
➤ Soft Bias

e.g. prefer small coefficients (i.e. less complex functions)

... the rest is just math (optimization/search)



Avoid Overfitting





USE CASE 110KV TRANSFORMERS

Predictive maintenance

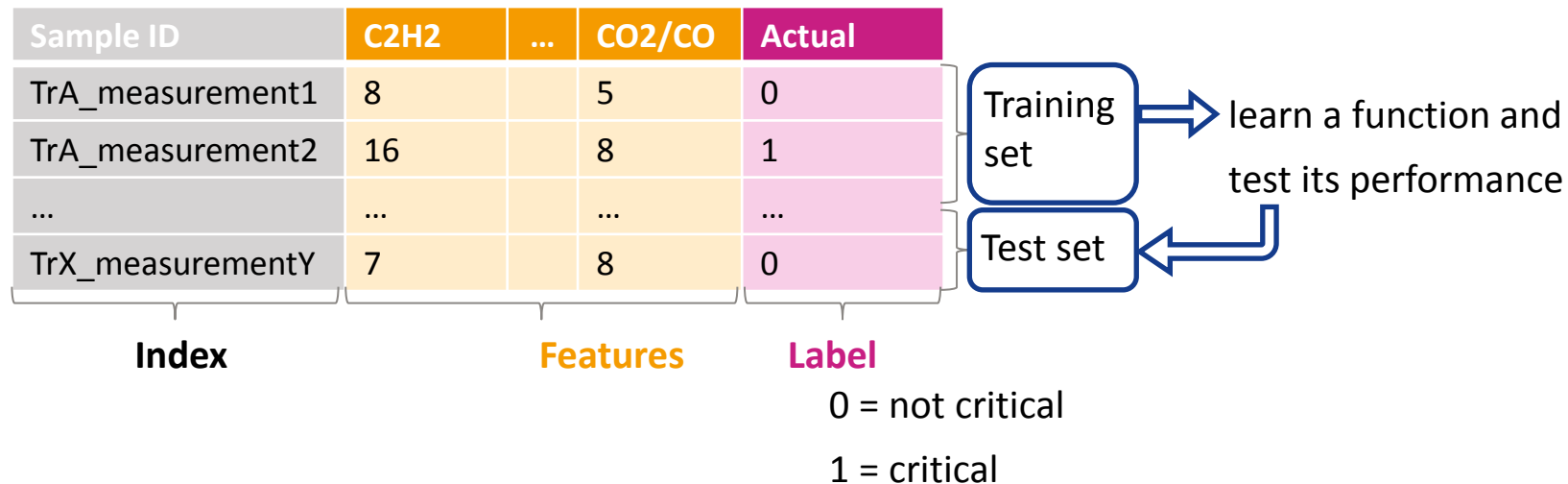
110 kV Transformatoren sind robust – doch Wartung und Austausch sind kostspielig



Vorhersage des (zukünftigen) Zustands eines Trafos

fiktive
Daten

Classification Problem



Testing the Function

Make predictions on the test set

C2H2	...	CO2/CO
10		5
15		7
...		...
7		8



$$f(x_{i,1}, \dots, x_{i,p})$$



Prediction	Actual
0.1 $\approx$ 0	0
0.8 $\approx$ 1	1
...	...
0.3 $\approx$ 0	0

test set with 1500 instances

- 75 critical
- 1425 not critical

The trained model (function) classified
99% of the instances in the test set correctly.

Is that good or bad?

Cost Sensitive Classification

Predictions for the test set

Prediction	Actual
0.1	0
0.8	1
...	...
0.3	0

Choose a threshold
e.g. 0.3

test set with 1500 instances

- 75 critical
- 1425 not critical

Confusion Matrix		Actual	
		0	1
Predict.	0	1110	11
	1	315	64



USE CASE E-PROGNOSIS

Prognose von Ladesäulen

E-Mobility-Prognose für die Netzplanung

- Sie erhalten die Anzahl an privaten, gewerblichen, halb-öffentlichen und öffentlichen Ladepunkten sowie die Haupttreiber 100 mal 100 Meter genau



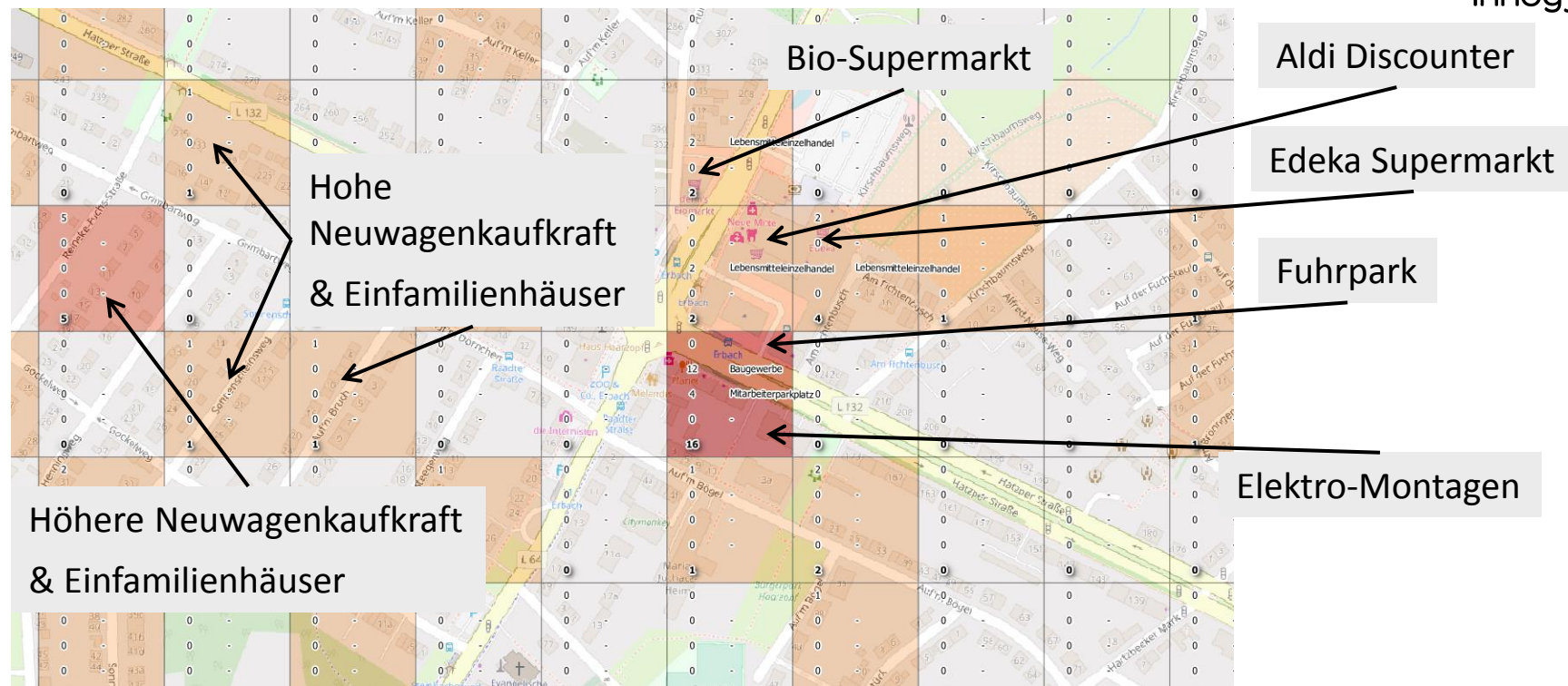
Privat:	2 Ladepunkte
Gewerbe:	2 Ladepunkte
Haupttreiber:	Verkehr und Lagerei
Halböffentlich:	4 Ladepunkte
Haupttreiber:	Parkhaus
Öffentlich:	0 Ladepunkte
Haupttreiber:	–



Beispielerggebnis e-prognosis in Essen Haarzopf



innogy

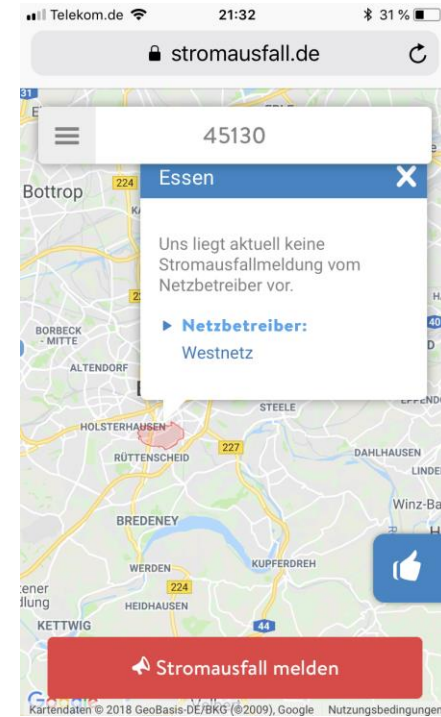
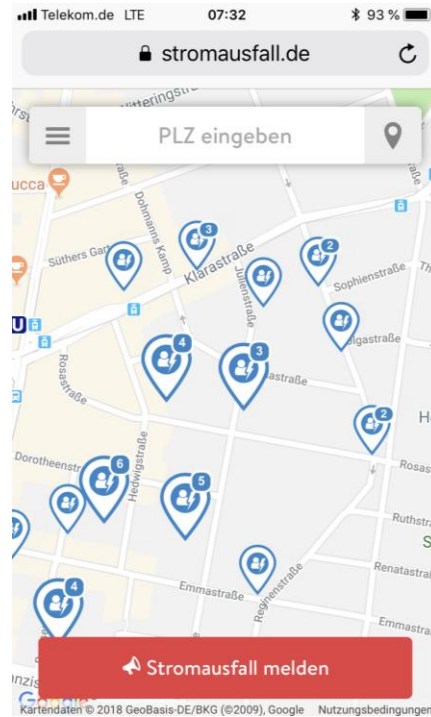
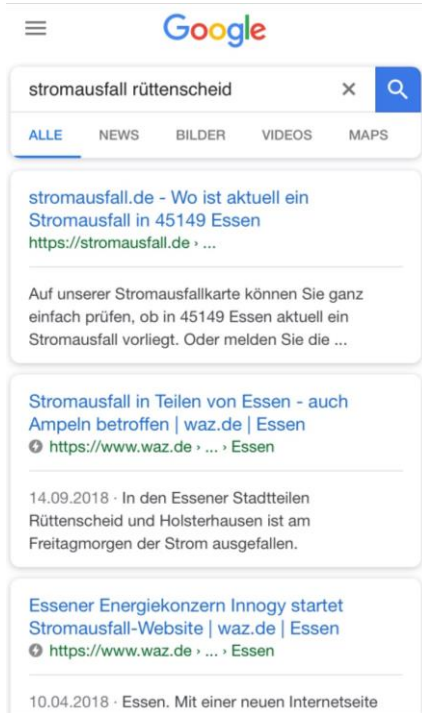




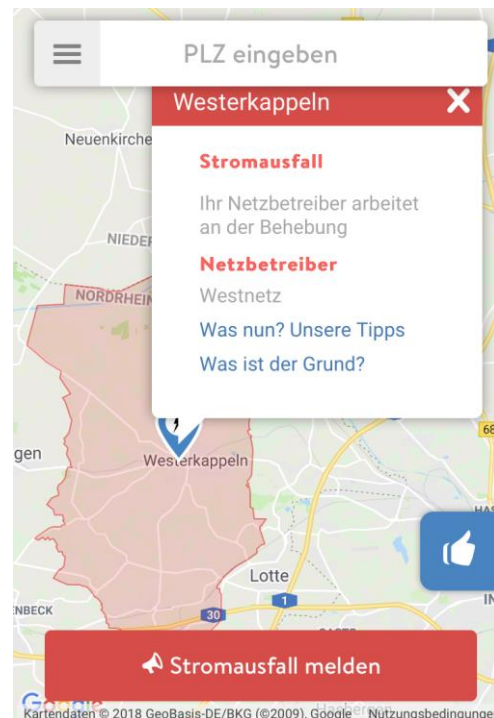
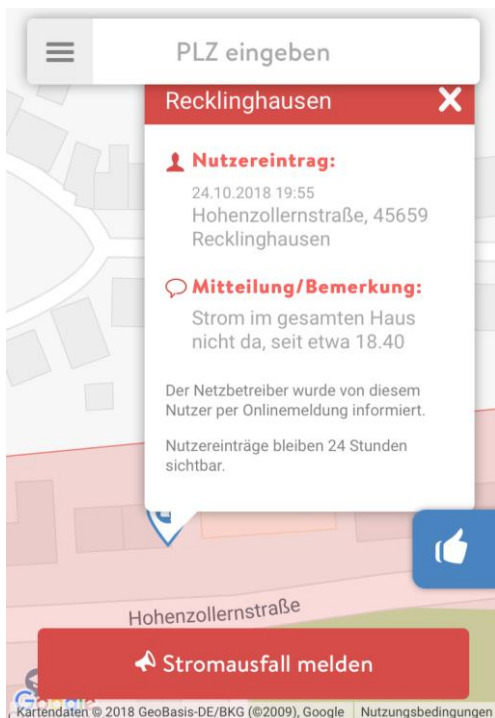
USE CASE STROMAUSFALL.DE

Web Analytics

Stromausfall in Rüttenscheid



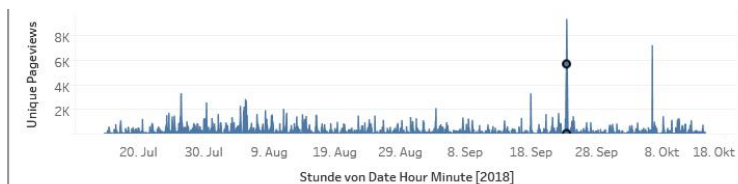
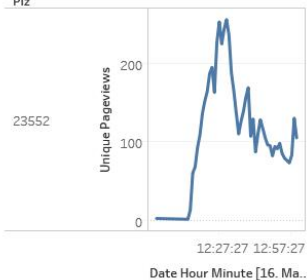
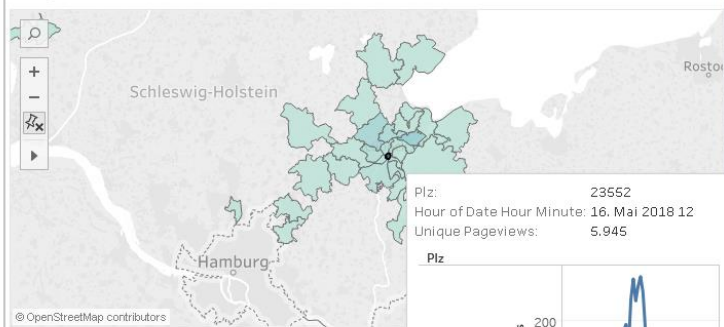
Karteneinträge auf www.stromausfall.de



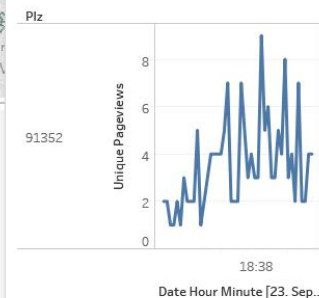
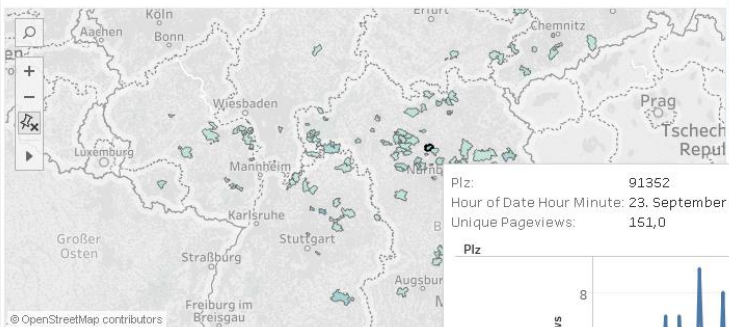
Datenvisualisierung für www.stromausfall.de



Map Overview



Map Overview





FAZIT

Fazit

Data Science ist das, was funktioniert, um einen Mehrwert aus Daten zu erzeugen

Machine Learning ist ein essentieller Teil von Data Science

Daten und ein Data Scientist alleine lösen meistens nicht das Problem



FRAGEN