Detecting problematic customers

OUTLINE

flaschenpost.de

- Who Am I?
- What is flaschenpost?
- Why did we look at problematic customers?
- How can we categorize them?
- How can we identify them at all?
- How does it all work in production?
- Did it work at all?

OUTLINE

flaschenpost.de

- Who Am I?
- What is flaschenpost?
- Why did we look at problematic customers?
- How can we categorize them?
- How can we identify them at all?
- How does it all work in production?
- Did it work at all?

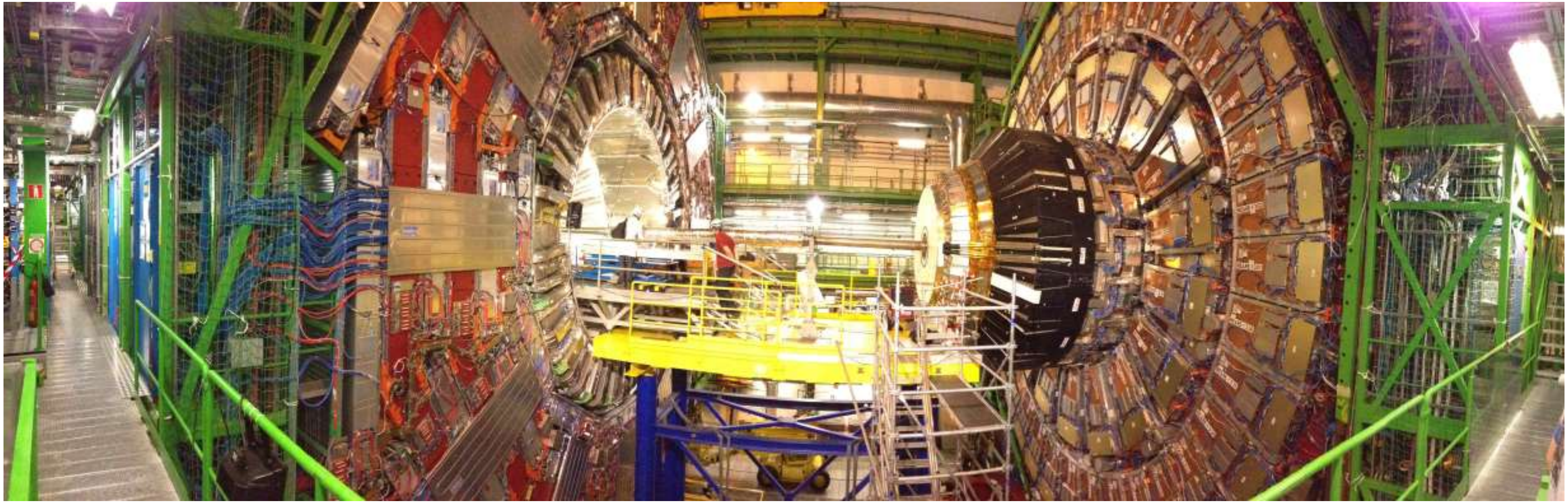
Who am I?

- Study at RWTH Aachen University
  - Physics Bachelor
  - Physics Master
  - Physics PhD
- Specialized in experimental higher energy particle physics
- Worked through all studies at the CMS experiment at CERN
- Searched for Dark Matter, Supersymmetric Particles and Quantum Black Holes
- Published my results in scientific Journals

What I did in the past?

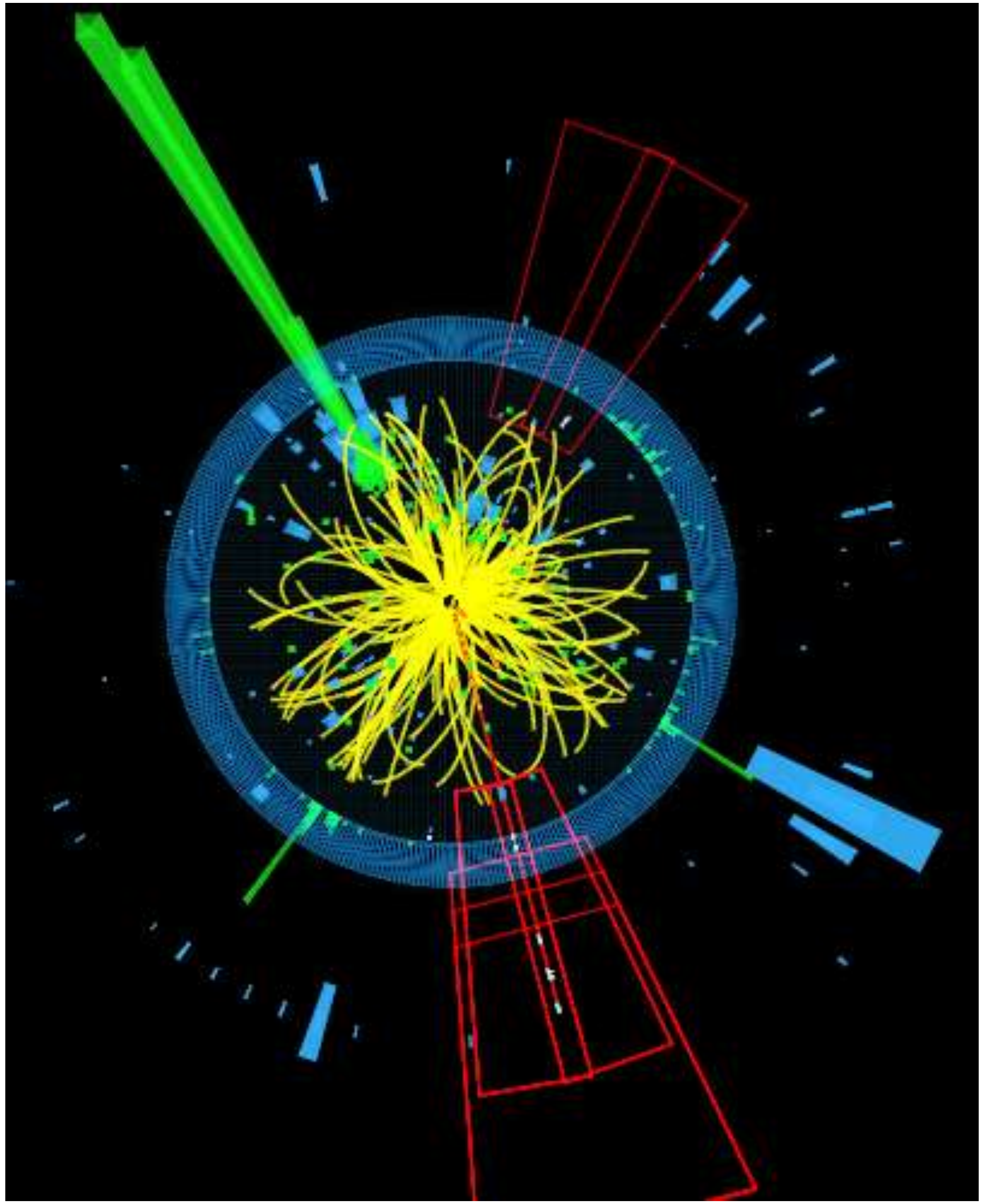
flaschenpost.de

- The biggest particle accelerator in the world collides protons
  - At a centre of mass energy of 14 TeV
  - At a rate of 40 million collisions per second
  - At ~50 interactions per beam crossing
- From this 2 billion collisions per second we have to select the interesting ones
- Interesting collisions (e.g. black holes) might occur a few time per year



What I did in the past?

- You have to analyse the data:
  - Filter all problems
  - Correct all known effects
  - Select only the interesting events
  - Do this on a grand scale (a few PB of data)
  - Compare everything to simulation
- You might find events like these:
  - Could be a black hole
  - Could be a background process (e.g. WW)
- Apply some statistics to decide if you have found something
  - In this case: Everything is compatible with the background
  - Reference: **JHEP 1804 (2018) 073**



My next steps?

- So the main task as an experimental high energy physicist is analysing data
  - Other tasks are supporting lectures and students
  - Giving conference presentations
  - Some time of the year be an on-call expert for some part of the detector at CERN
- After finishing my studies it was obvious to start working as a data scientist
- That’s how I got to flaschenpost SE

HEAD OF DATA ANALYTICS AT FLAPO

- Founding member of the data analytics team at flaschenpost
- Team lead regarding all data science topics
  - Forecasting
  - Simulation
  - Automatization
  - Analysis
- Since beginning of the year Head of the Data Analytics Department
  - Responsible for the data science team
  - Now also for the Business Intelligence Team
    - Reporting infrastructure
    - All reporting needs
  - Team size of 9 people

OUTLINE

- Who Am I?
- What is flaschenpost?
- Why did we look at problematic customers?
- How can we categorize them?
- How can we identify them at all?
- How does it all work in production?
- Did it work at all?

Business Model - Overview

flaschenpost.de

Customer

Process

CLASSIC PURCHASE PROCESS

carrying empty, dirty bottles

stressful deposit return & shopping

carrying full, heavy bottles

KG

flaschenpost business model

convenient online order

no delivery fee

within 120 minutes

& order reception & deposit return

Direct bulk purchases

Streamlined & scaled logistics

Efficient routing algorithm

Direct customer contact

vertical integration

Our customers love us

vertical integration leads to profitability

STATIONARY BEVERAGE DEALER

Indirect sourcing

Wholesale distribution

Expensive retail space

Inefficient individual trips

shop-order

new website will be available in responsive design

app is coming soon

flaschenpost.de

desktop | tablet | mobile | iOS | android

homepage | app

shop-order

highly efficient picking process

top-sellers: pick-by-light
are picked more than 50% faster using a pick-by-light system

replenishment of top-sellers is provided by a palet-robot

slower moving products
are picked using an in-house developed software system

routing algorithm

trip packing

delivery to the customer

top-seller picking system

digital order picking system

shop-order

picking process

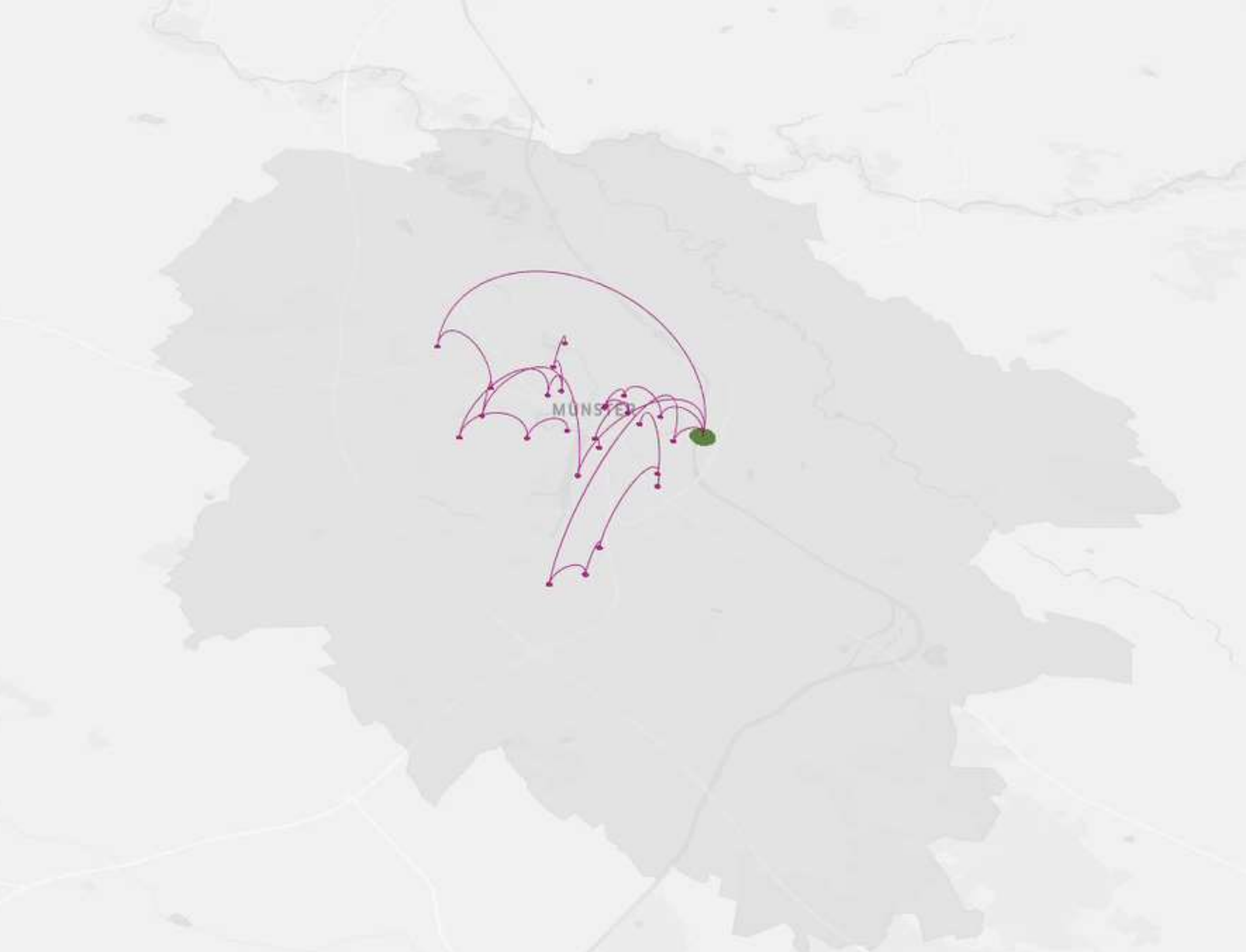
routing algorithm

we have developed our own proprietary routing algorithm

best-in-class leading performance in solving complex decisions about the composition of orders per trip

trip packing

delivery to the customer



routing algorithm

shop-order

picking process

routing algorithm

trip packing

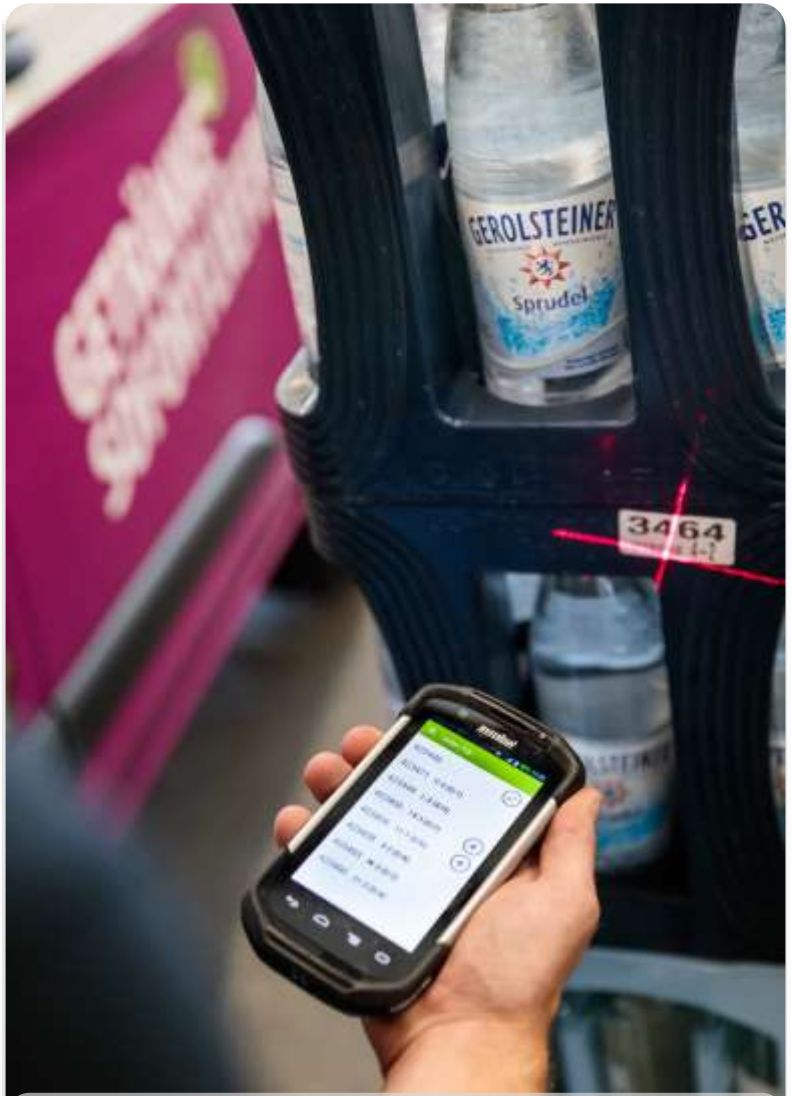
every single crate is clearly assigned

all drivers are equipped with a handheld

our software covers all activities during the delivery process, such as scanning of crates and navigation

this allows us to have highly efficient and fail-safe processes

delivery to the customer



the handheld

fail-safe processes

shop-order

picking process

routing algorithm

trip packing

delivery to the customer

the driver hands over the ordered beverages

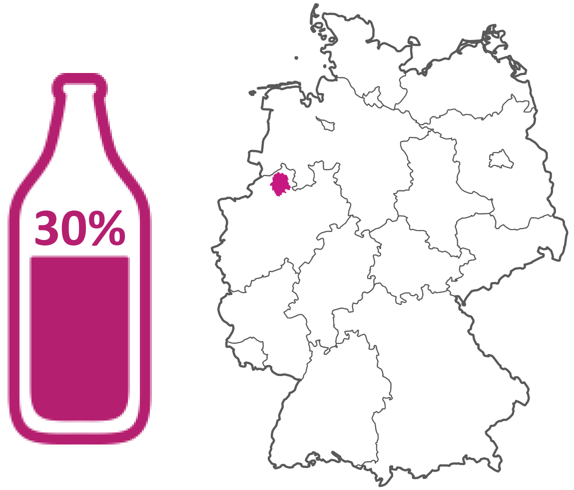
payment and deposit return are managed by using the handhelds

delivery & deposit return

electronic billing

OUTLOOK – AND THIS IS JUST THE BEGINNING

Market share Münster



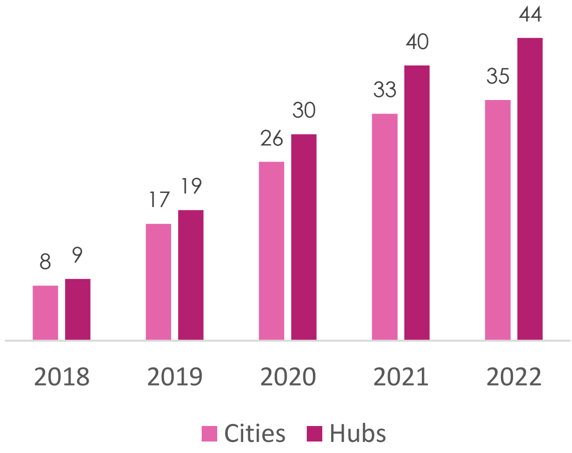
In Münster we have reached a market share of > 30% within just two years.

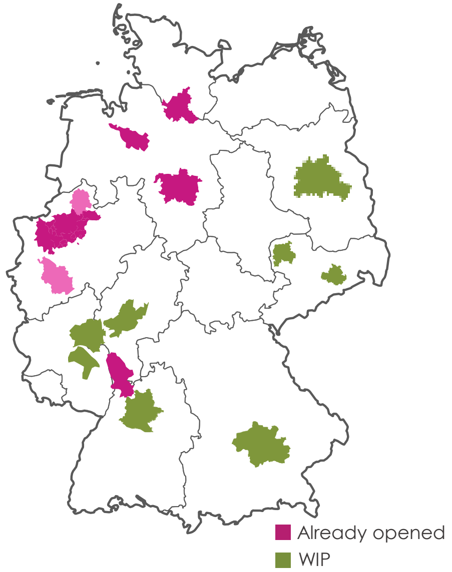
This corresponds to a

12 mn EUR

run rate in Münster

Expansion with one new hub per month





Currently we are opening one new hub per month.

If we reach Münster-level market shares across Germany, this opens an opportunity of at least

13 bn EUR

market share in Germany

…to all of Europe

We are largely unrivalled across Europe with our concept.

An expansion therefore offers the chance of a

75 bn EUR

market share in Europe

- Who Am I?
- What is flaschenpost?
- Why did we look at problematic customers?
- How can we categorize them?
- How can we identify them at all?
- How does it all work in production?
- Did it work at all?

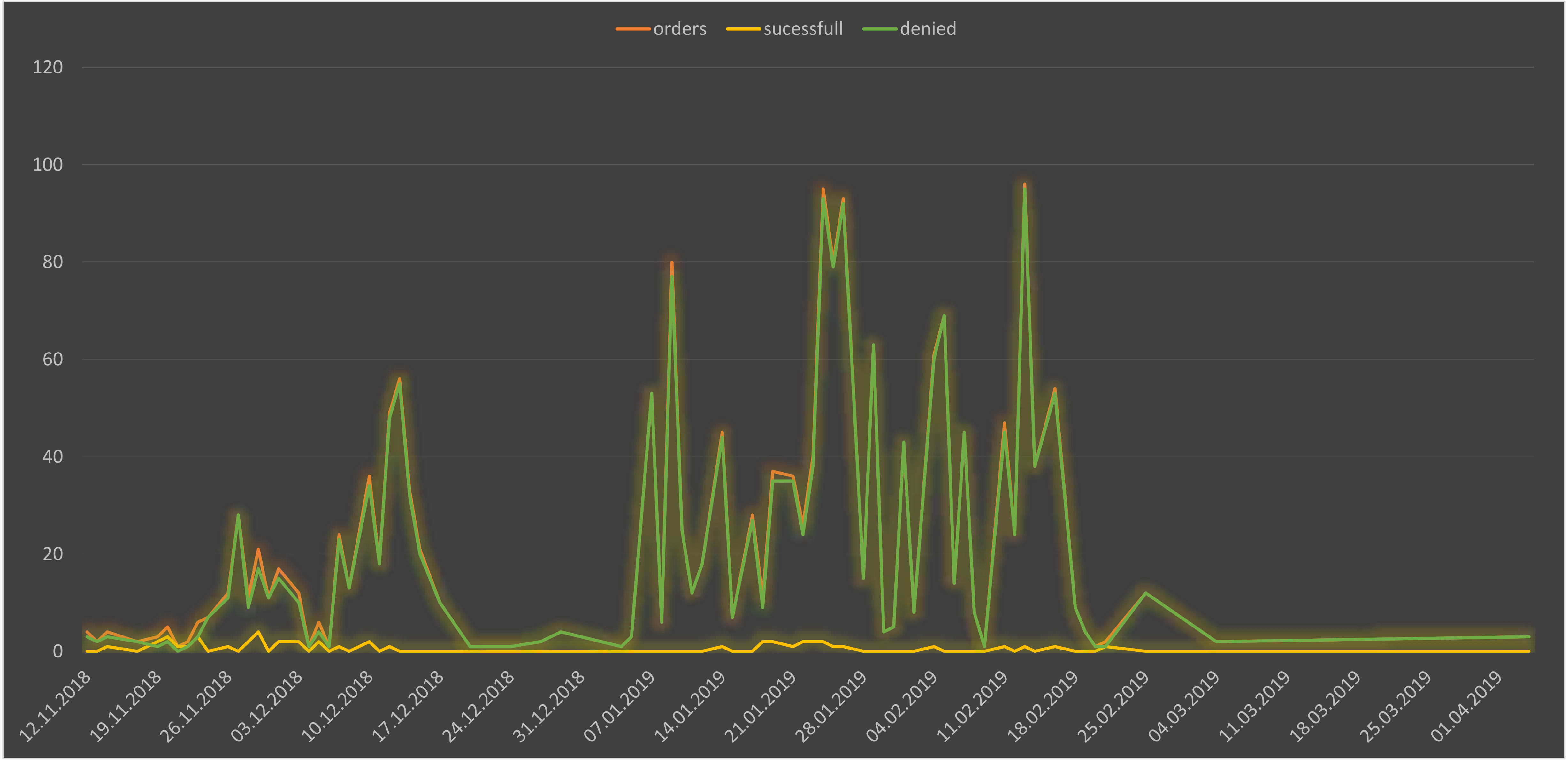
How it started? (1)

flaschenpost.de

- Middle of December 2018:
  - Some strange orders occur (e.g. Acceptance of delivery on the parking lot)
  - On a given day they come from the same IP
- Filter IPs which give a lot of orders
  - Whitelist static IPs from known customers
  - This allows to veto most of the problematic orders
- The orders come with very different features
  - Changed Names
  - Often different addresses
  - Different shopping carts
- Most orders can be vetoed by this manual process
- The need for an automated system is clear

How it started? (2)

flaschenpost.de



How it continued?

- The number of dunning procedures doubles suddenly
  - In these orders the shopping cart stands out
  - Many orders just contain cola for more then 100€
  - Just orders of about 6 crates cola

- Just checking manually orders for this behaviour significantly reduces the number of dunning procedures
- But this is another indication for an automated filtering process
- These among other examples lead to the development of an automated process

To what it actually ammounts

- From 1000 orders:
  - 23 are somewhat problematic for us
  - 16 because the delivery is not successful
  - 6 because we get not paid

- This means payment default in the amount of >10k€ per Month

- To identify these problems as early as possible is therefore crucial
  - The customer service team can make a human decision
  - But which customers should be checked?
  - We need an automated first level decision

- Who Am I?
- What is flaschenpost?
- Why did we look at problematic customers?
- How can we categorize them?
- How can we identify them at all?
- How does it all work in production?
- Did it work at all?

FRAUD CLASSIFICATIONS

flaschenpost.de

- Different versions of (problematic) customers:

1. Good customer
2. Cancelled orders
3. Payment only after the first payment reminder
4. Delivery not accepted
5. No payment even after payment reminder
6. Customer not encountered

- Classify 1. as good customers
- Classify everything else as problematic customers

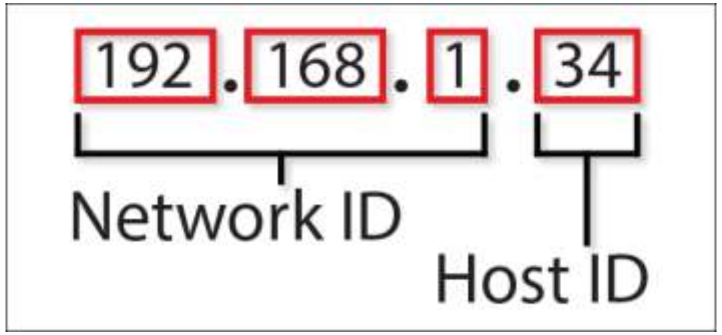
OUTLINE

flaschenpost.de

- Who Am I?
- What is flaschenpost?
- Why did we look at problematic customers?
- How can we categorize them?
- How can we identify them at all?
- How does it all work in production?
- Did it work at all?

Look at different features

- IP



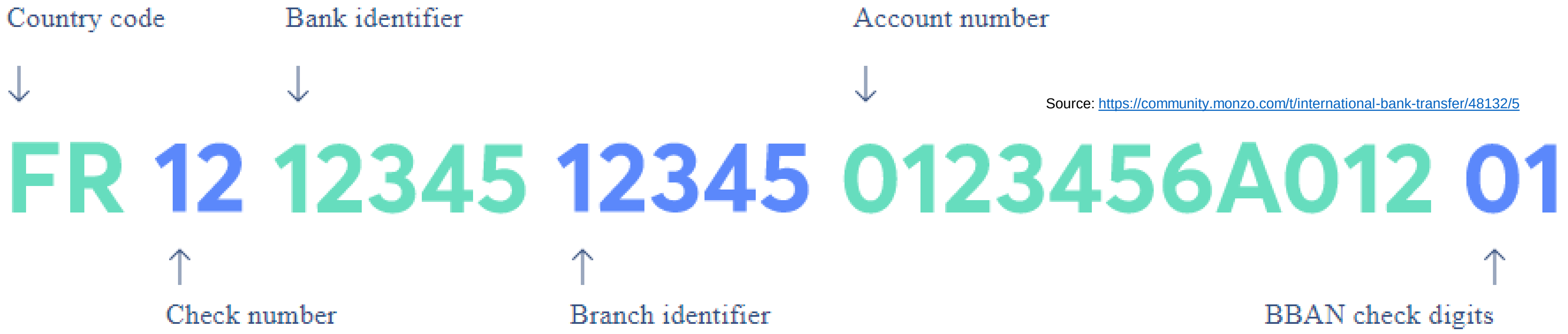
Source: http://www.itplatz.net/wie-funktionieren-ip-adressen/

- Allows to look at the ISP of the customer
- Allows to observe strange order patterns (see first story)
- Allows to identify big business customers
- But can only give a vague indication about the customer

LOOK AT DIFFERENT FEATURES (IBAN)

flaschenpost.de

- IBAN (International bank account number)



Source: https://community.monzo.com/t/international-bank-transfer/48132/5

- First: check for validity (check numbers)
- But also valid IBANs can be problematic
  - It could be public IBANs which do not belong to the customers
  - The check numbers could be okay, but the account doesn't exist

LOOK AT DIFFERENT FEATURES

- Obviously there is not just one (few) features(s) which allows to classify customers
- Therefore we use as many features as possible
- Here are some examples out of the used feature set:
  1. Webshop behaviour: e.g. how fast do the customer adds items to the his shopping cart?
  2. Time and day: At which weekday and at which time is the order placed
  3. Shopping cart: What is the content of his shopping cart, e.g. only water or only alcohol?

Source: https://www.dlf.org.uk/factsheets/safety

QUALITY CRITERION

flaschenpost.de

- Example:
  - 920k Orders

- allocation rate

		True	False	Numbers	rate
Actually good customers 900k:	good	99%	1%	9k	36%
Actually bad customers 20k:	bad	20%	80%	16k	64%

Decision by customer service

True: Customers recognized as good

False: Customers recognized as problematic

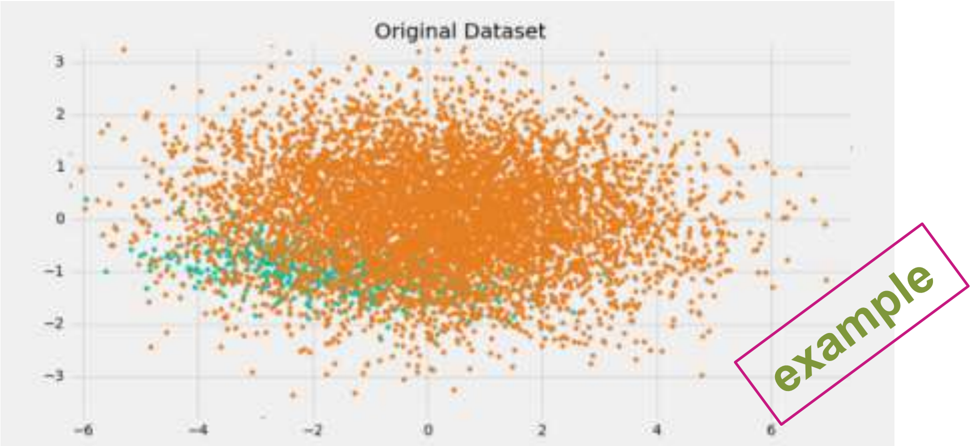
Numbers: Absolute numbers that would arrive at customer service

rate: Percentage of good and bad customers in manual inspection

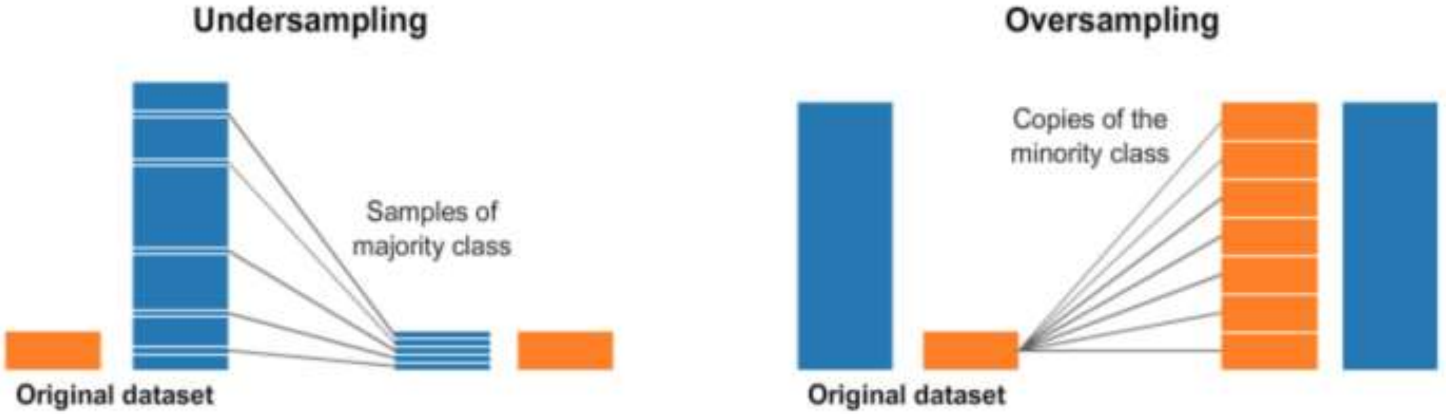
- The numbers in must be as large as possible
- The numbers in must be as small as possible

IMBALANCED DATASETS

- As discussed previously only ~1% of orders are problematic
- Therefore the dataset is highly imbalanced
- To get a valid training result, different balancing approaches are used
  - Undersampling: Using the bigger class only partially for training (Reduces statistics)
  - Oversampling: Copying the smaller class for training (can enhance artefacts of the data)
  - Creating artificial data: Needs clever algorithms to create new data with the same features as the original data
- For now we used undersampling as the easiest method with the least problems



Source: https://medium.com/james-blogs/handling-imbalanced-data-in-classification-problems-7de598c1059f

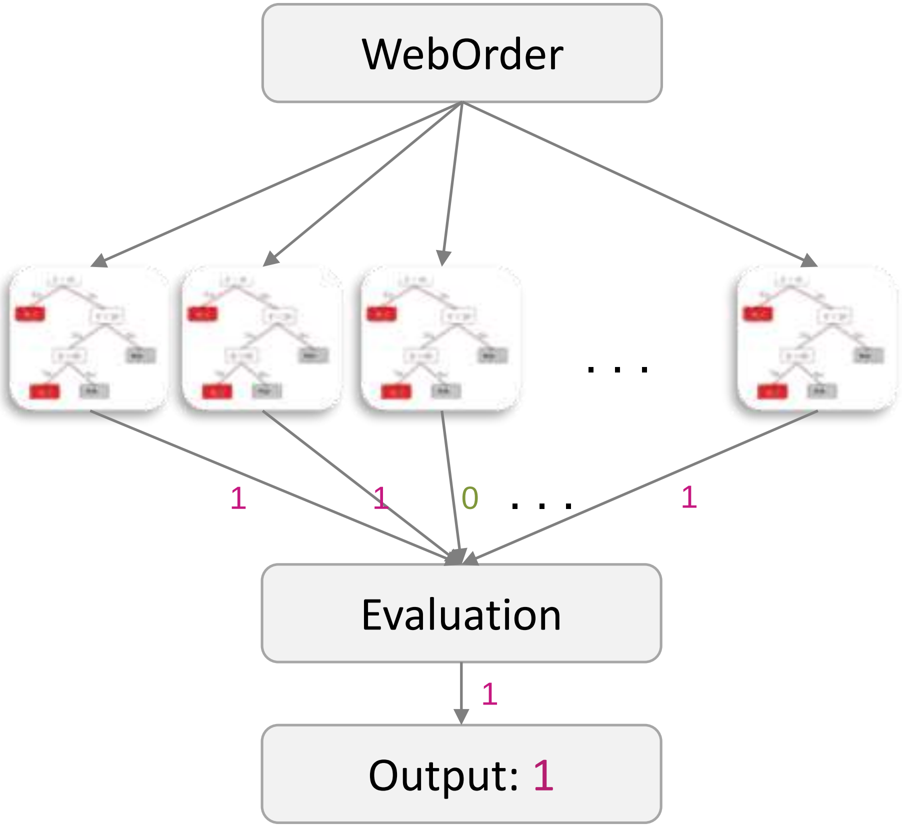


Source: https://towardsdatascience.com/having-an-imbalanced-dataset-here-is-how-you-can-solve-it-1640568947eb

EXAMPLE MODEL: RANDOM FOREST

Characteristics

- Many random decision trees make decisions according to the majority principle



Results

- Training on 14000 balanced customers:
  - 7673 Non-Fraud / 6327 Fraud
- Gives non optimal results

	True	False	Numbers	rate
good	78%	22%	215	95.5%
bad	57%	43%	10	4.5%

Of 1000 orders

RESULT COMPARISON

flaschenpost.de

Random Forest

	True	False	Numbers	rate
good	78%	22%	215	95.5%
bad	57%	43%	10	4.5%

AdaBoosting

	True	False	Numbers	rate
good	80%	20%	195	95%
bad	53%	47%	11	5%

Gradient Boosting

	True	False	Numbers	rate
good	84%	16%	156	94%
bad	57%	43%	10	6%

Best-Choice Modelle

Extra Tree

	True	False	Numbers	rate
good	78%	22%	215	95.5%
bad	56%	44%	10	4.5%

Of 1000 orders

FRAUD CLASSIFICATIONS (OLD)

- Different versions of (problematic) customers:

1. Good customer
2. Cancelled orders
3. Payment only after the first payment reminder
4. Delivery not accepted
5. No payment even after payment reminder
6. Customer not encountered

- Classify 1. as good customers
- Classify everything else as problematic customers

Fraud classifications (New)

- Different versions of (problematic) customers:

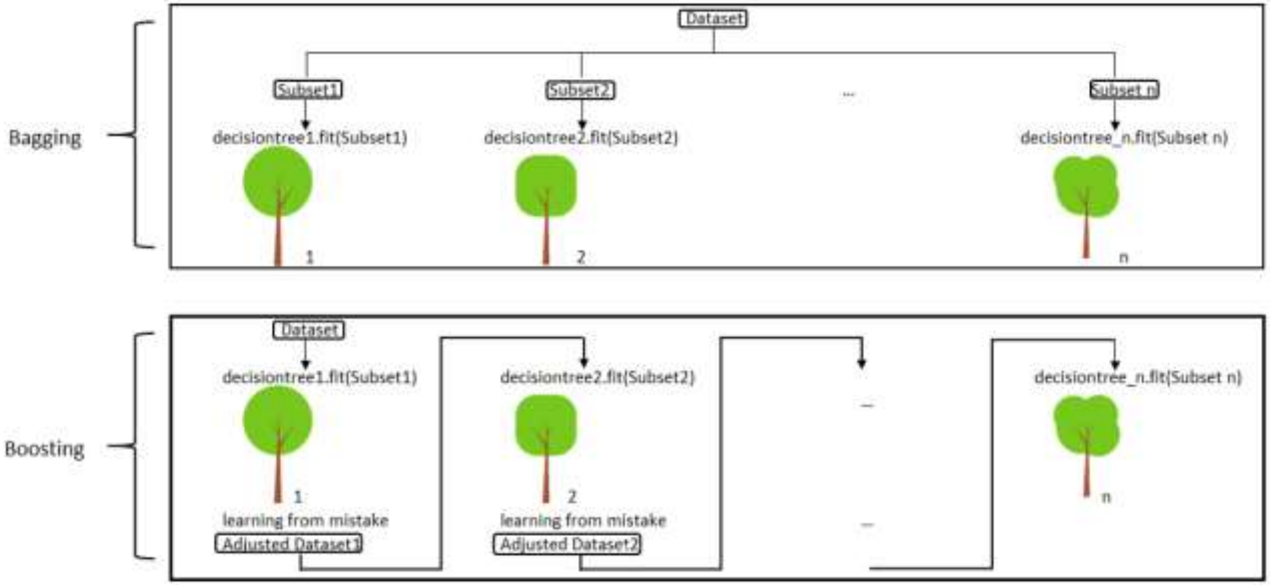
1. Good customer
2. Cancelled orders
3. Payment only after the first payment reminder
4. Delivery not accepted
5. No payment even after payment reminder
6. Customer not encountered

- Classify 1., 2., 3., 4. and 6. as good customers
- Classify only 5. as problematic customers

Modell: Gradient Boosting

Characteristics

- Combined different decision trees to get a better decision
- Either use bagging (e.g. Random Forrest)
- Or use boosting (e.g. Adaptive Boosting)



- Best results are obtained with the Gradient Boosting algorithm

Results

- Training on 14000 balanced customers:
  - 7673 Non-Fraud / 6327 Fraud
- Comparison of results with previous classification
- Final results:

	True	False	Numbers	rate
good	97%	3%	29	63%
bad	28%	72%	17	37%

- Compare to Training on 14000 balanced

	True	False	Numbers	rate
good	84%	16%	156	94%
bad	57%	43%	10	6%

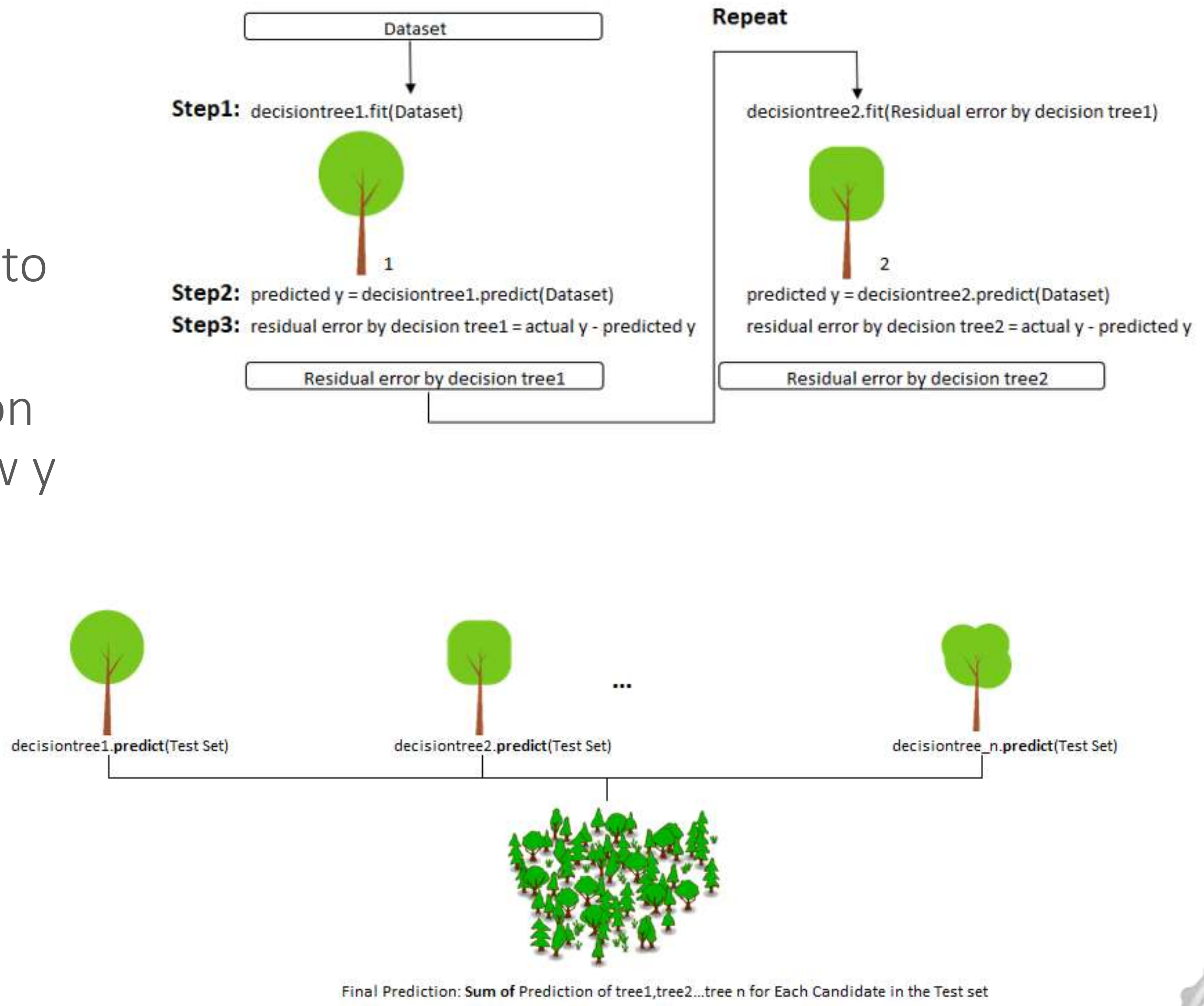
Of 1000 orders

Source: https://towardsdatascience.com/basic-ensemble-learning-random-forest-adaboost-gradient-boosting-step-by-step-explained-95d49d1e2725

Modell: Gradient Boosting

- How does the gradient boosting work:

1. **Train** a decision tree
2. **Apply** the decision tree just trained to predict
3. **Calculate** the residual of this decision tree, Save residual errors as the new y
4. **Repeat** Step 1 (until the number of trees we set to train is reached)
5. **Make the final prediction as sum of predictions**



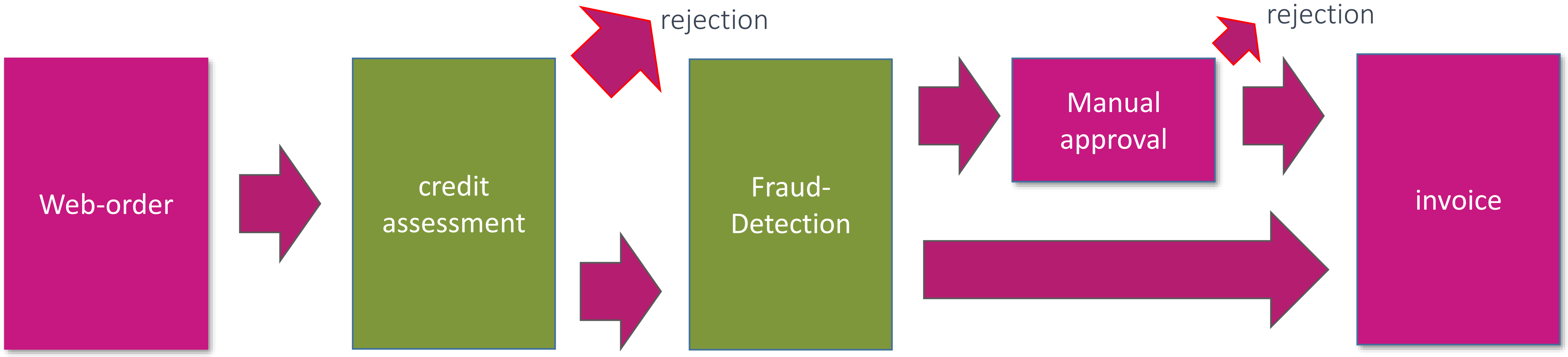
Source: https://towardsdatascience.com/basic-ensemble-learning-random-forest-adaboost-gradient-boosting-step-by-step-explained-95d49d1e2725

OUTLINE

flaschenpost.de

- Who Am I?
- What is flaschenpost?
- Why did we look at problematic customers?
- How can we categorize them?
- How can we identify them at all?
- How does it all work in production?
- Did it work at all?

System concept



- Credit assessment:
  - Determination of the default probability for already known customers
  - Based on past behaviour
- Fraud-Detection:
  - Determination of the default probability for unknown customers
  - Based on his behavior

Run enviroment

flaschenpost.de

Quick version

- Define and train the model
- Compile everything into a .exe file
- Provide the file to the web shop process
  - Let it call by the web shop
  - Use the return value as evaluation
- Very inconvenient regarding updates
- Limited performance

Long term solution

- Use a dedicated machine
- Connect to the shop system via API
- Allows for easy maintenance
  - Definition of complete environment
  - Easy update possibilities
- Better scalabilities
- Bigger requirements regarding availability

Source: https://www.giga.de/extra/api/

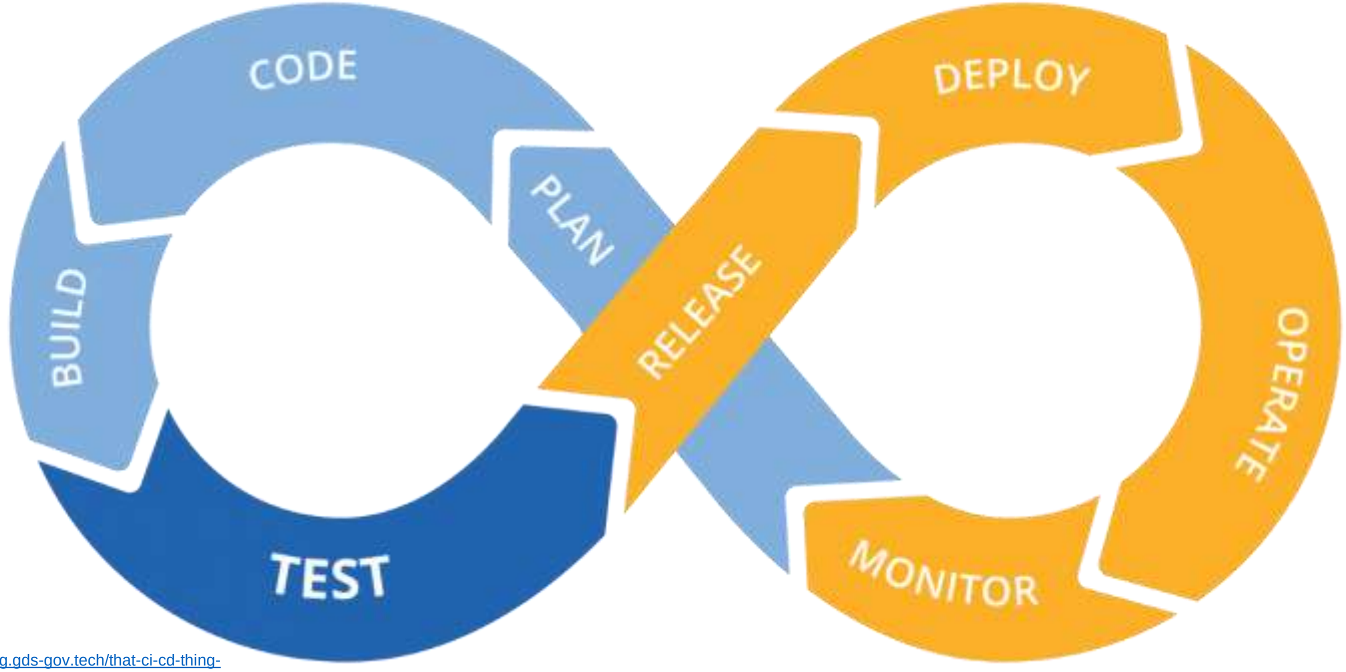
OUTLINE

flaschenpost.de

- Who Am I?
- What is flaschenpost?
- Why did we look at problematic customers?
- How can we categorize them?
- How can we identify them at all?
- How does it all work in production?
- Did it work at all?

LIVE RUN

- Sub steps are already productive
  - credit assessment
  - Filtering of various special cases
- Test system runs with the current webshop
  - In the preliminary variant per exe
  - Final evaluation only possible after receipt of all payment reminders
- Preliminary results show good agreement with test results
- Still many tests and improvements left



Source: https://blog.gds-gov.tech/that-ci-cd-thing-principles-implementation-tools-aa8e77f9a350

CONCLUSION/NEXT STEPS?

- We got a nice system working to identify problematic customers
- But there are many more improvements possible:
  - Improve the models itself (parameter tuning)
  - Try out some more models
  - Try other methods to handle the imbalanced dataset (e.g. artificial data)
  - Update the run environment
- A lot of data science left

Data science

Source: https://www.american.edu/uploads/standard/original/spa-datascience-news.jpg

flaschenpost.de

Getränke für Alle. Sofort.

Dr. Sören Erdweg

Head of data analytics, flaschenpost SE

soeren.erdweg@flaschenpost.de