zalando

ML Takes Control: Converting Business Rules into Machine Learning Features

Dr. Patrick Baier

Dipl.-math. oec. Henning Esser

Zalando Payments

Oktober 2019

OUTLINE

Who we are and what we do

Motivating Example

ML Takes Over

ML vs. Decision Rules

Experiments

zalando

WHO WE ARE
AND
WHAT WE DO

SPEAKERS

Patrick Baier

- Lead Data Scientist at Zalando (~ 5 years)
- PhD in Computer Science from Uni Stuttgart

Henning Esser

- Data Scientist at Zalando (~ 7 years)
- Diploma in Management Mathematics from TU Kaiserslautern

WHAT WE DO

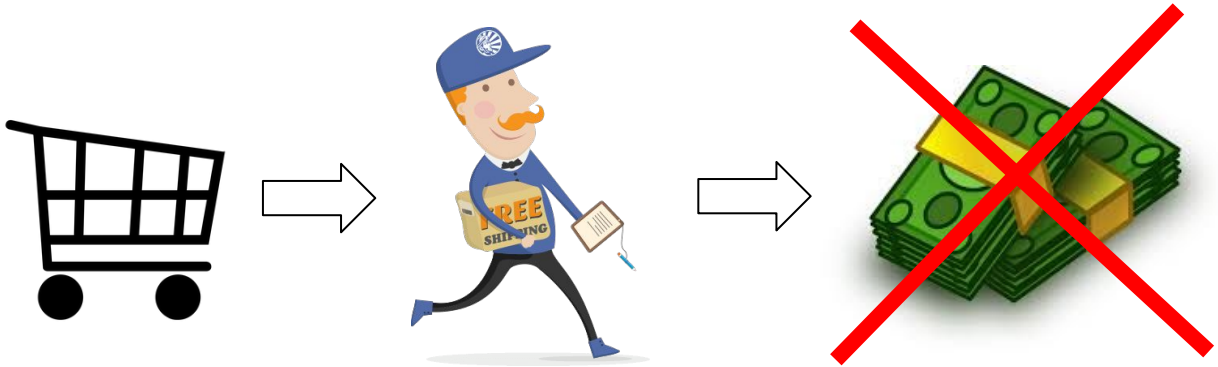
Detect and prevent payment fraud

Shopping Shipping Payment

zalando

WHAT WE DO

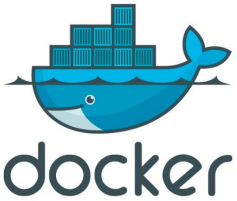
Detect and prevent payment fraud



Shopping Shipping Payment

zalando

OUR TECH STACK



Motivating Example

Once upon a time in the valley …

CAST

Founder and CEO

Head of Banking

Data Scientist

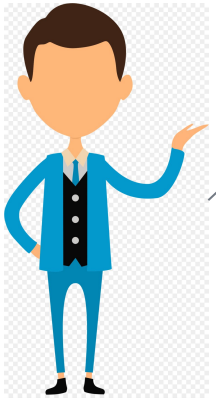
SETTING

We are at a fintech-startup:
An only online bank that wants to disrupt the banking market.

We need more customers
to attract more VC.

Let’s introduce the following feature:

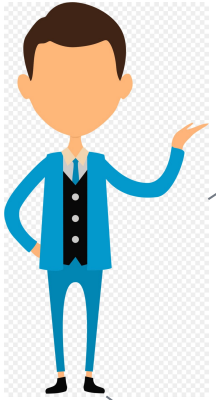
“A customer can get a free loan up to 10k immediately!”



Okay, but we need to decide which customers we offer this loan. We should only give it to those customers that will pay us the money back.

Makes sense...
Let’s use AI for this!
I will hire one of those Data Scientist!

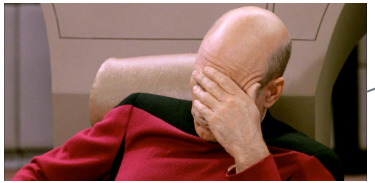
SCENARIO



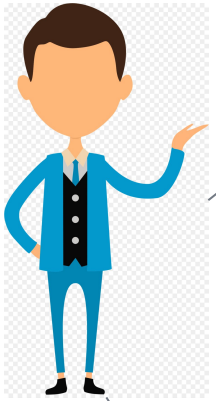
Welcome!
We need a machine learning model that predicts if a customer pays us back its loan.

Cool!
Where is the training data?

Hmm…we do not have data right now.



SCENARIO



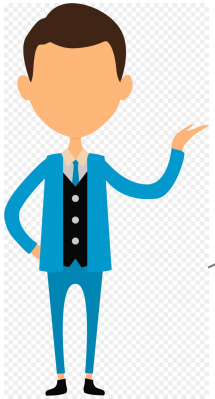
We cannot use AI, we do not have training data.

What? I thought AI can solve anything…
What can we do instead?

Let’s build some old school business rules with the help of domain experts.

...okay, but let’s pretend in front of the VCs that we do AI!

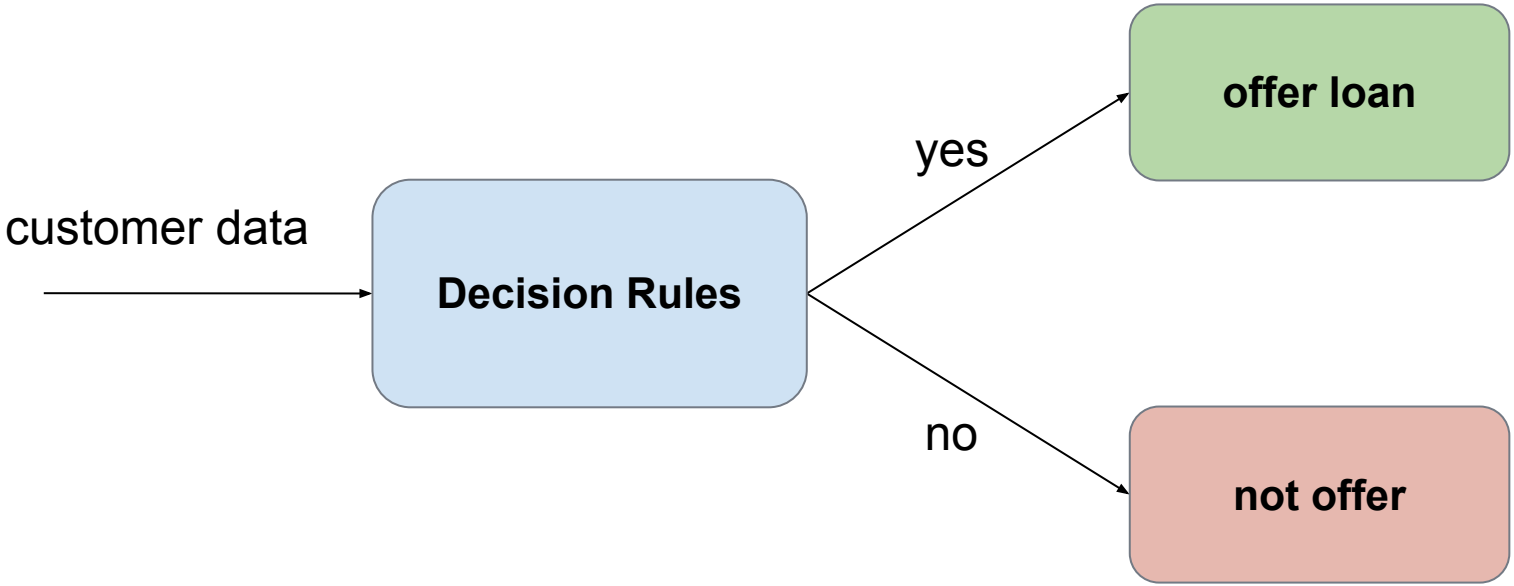
*After some back and forth, they finally came up with the following **decision rules**:*



We will offer the loan if a customer has:

- monthly income > 3k AND
- current balance > 0 AND
- external solvency score > 8

DECISION ENGINE



zalando

One year later …

SCENARIO

The loan product is great: It brought us 10M new customers and we raised 1B of funding!

Yeah...but our default rate is 5%. We are actually burning a lot of money.

Uhhh...what can we do?

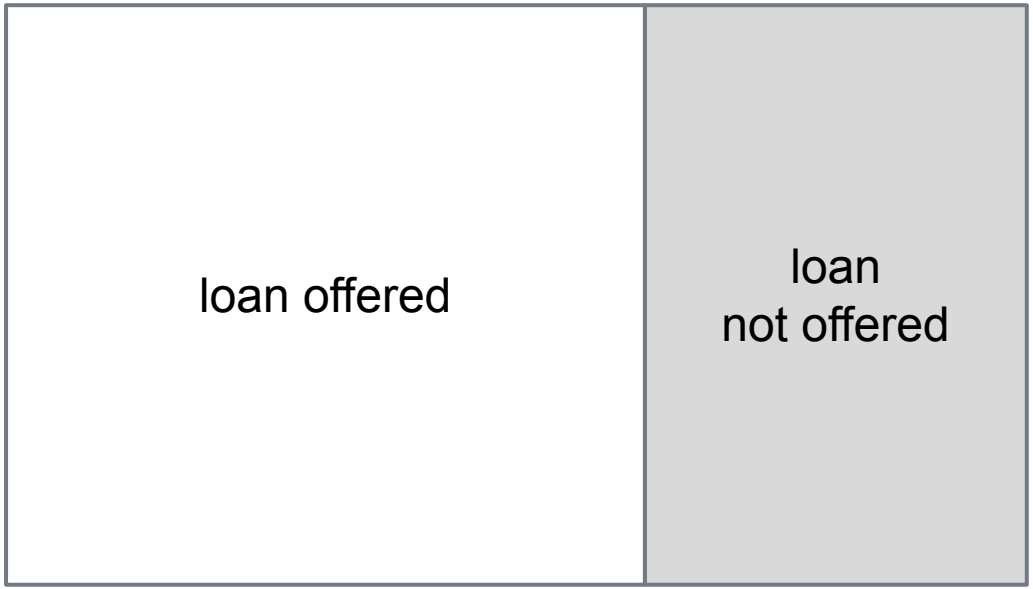
We have one year of data, let's do real AI!

And this is where the trouble starts …

Machine Learning take over!
Why it is hard to get rid off decision rules

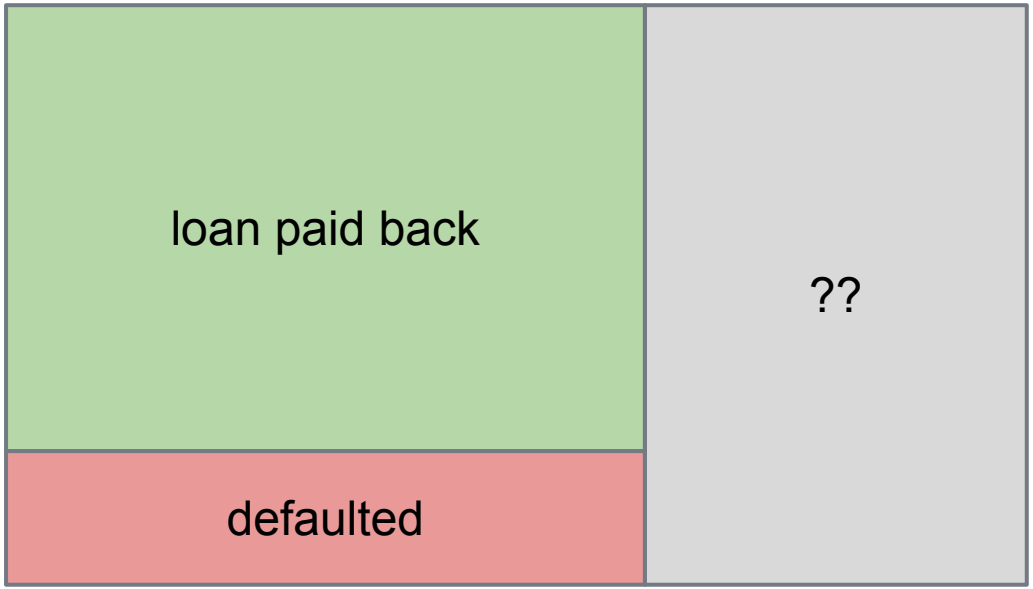

all customers

Was the loan offered?

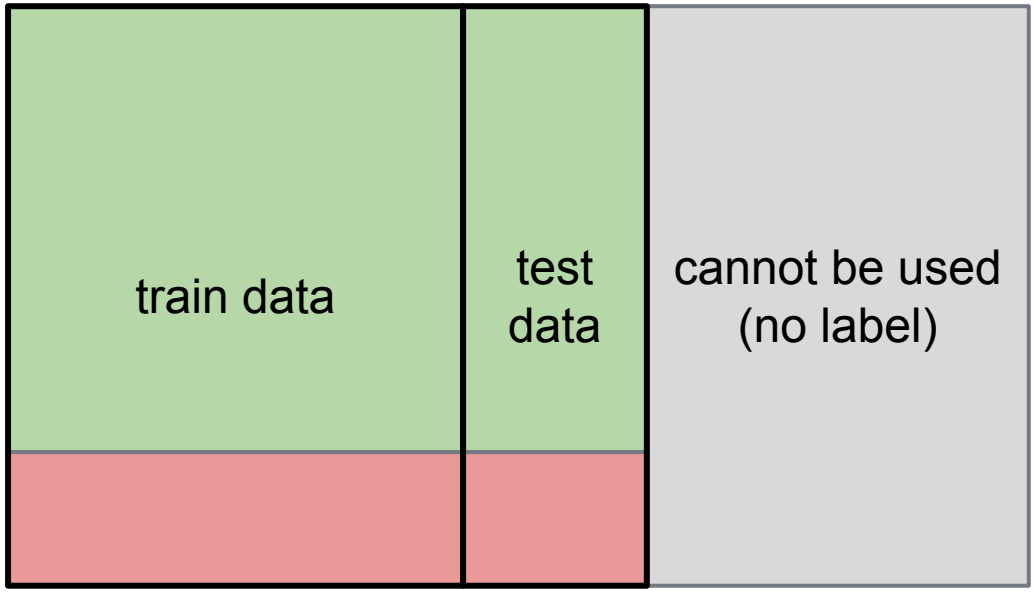


zalando

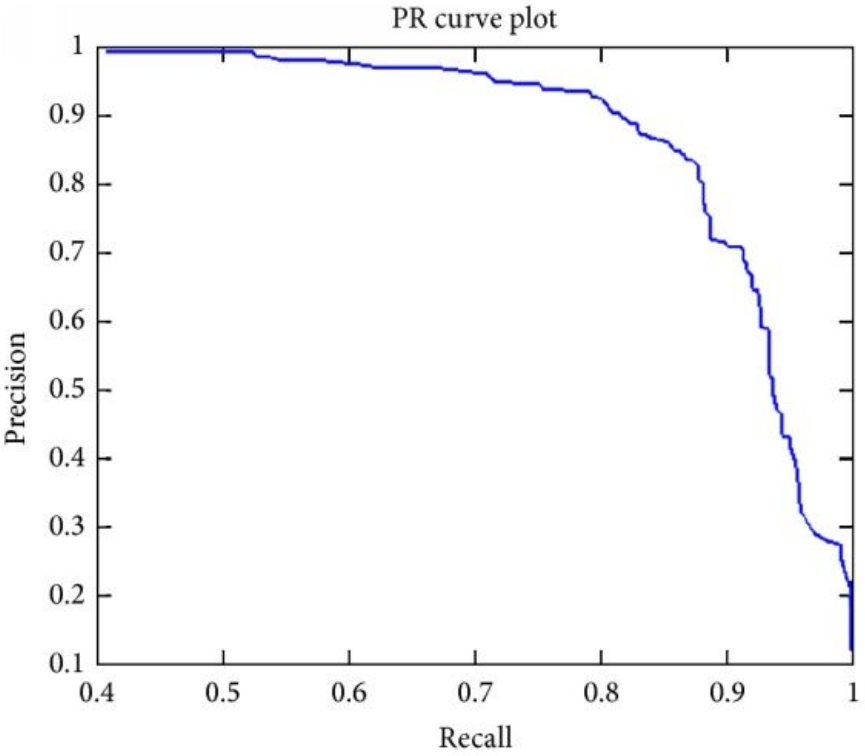
Was the loan paid back?



Training a machine learning model



PERFORMANCE



PR curve looks great!

Question:

- How can we deploy this model?

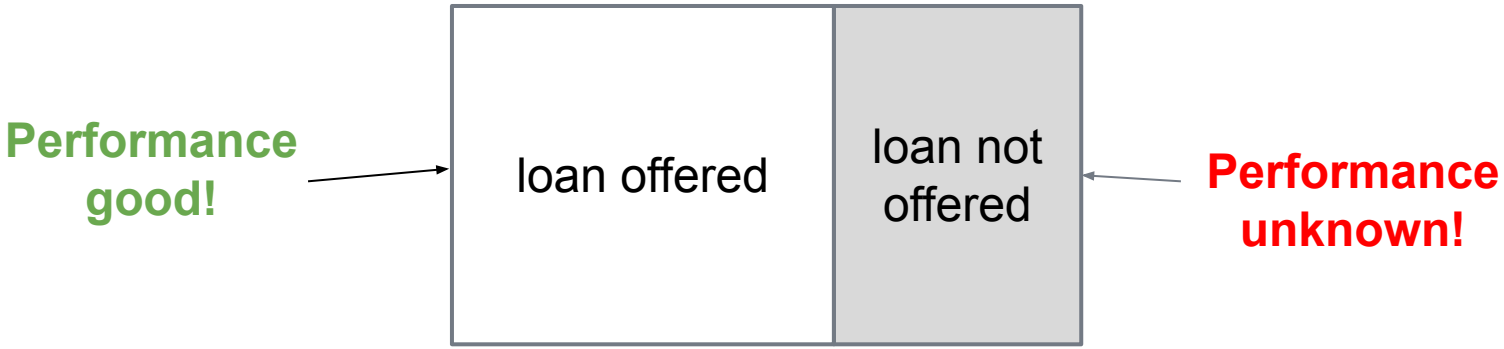
Problem:

- We only know our performance on the observed customers.

What are the options?

DEPLOYMENT

Option A: Be brave and use the model to predict on all customers.

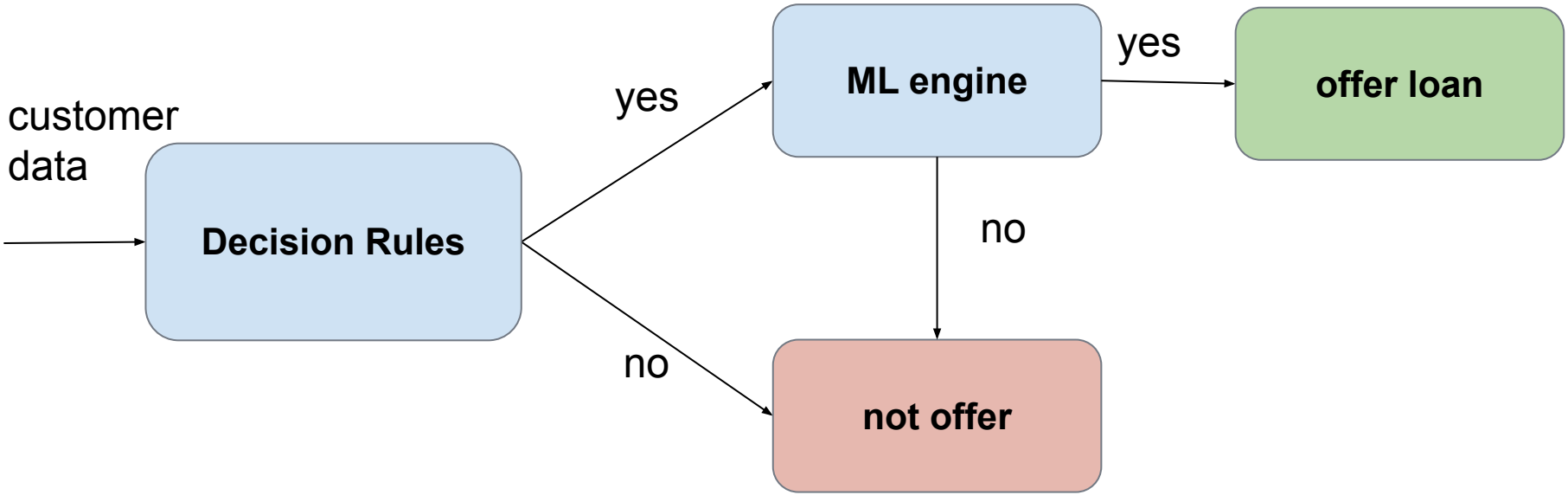


Danger zone:

- We do not know our performance on this data set.
- It could be superbad (a lot of data was not seen in training)!

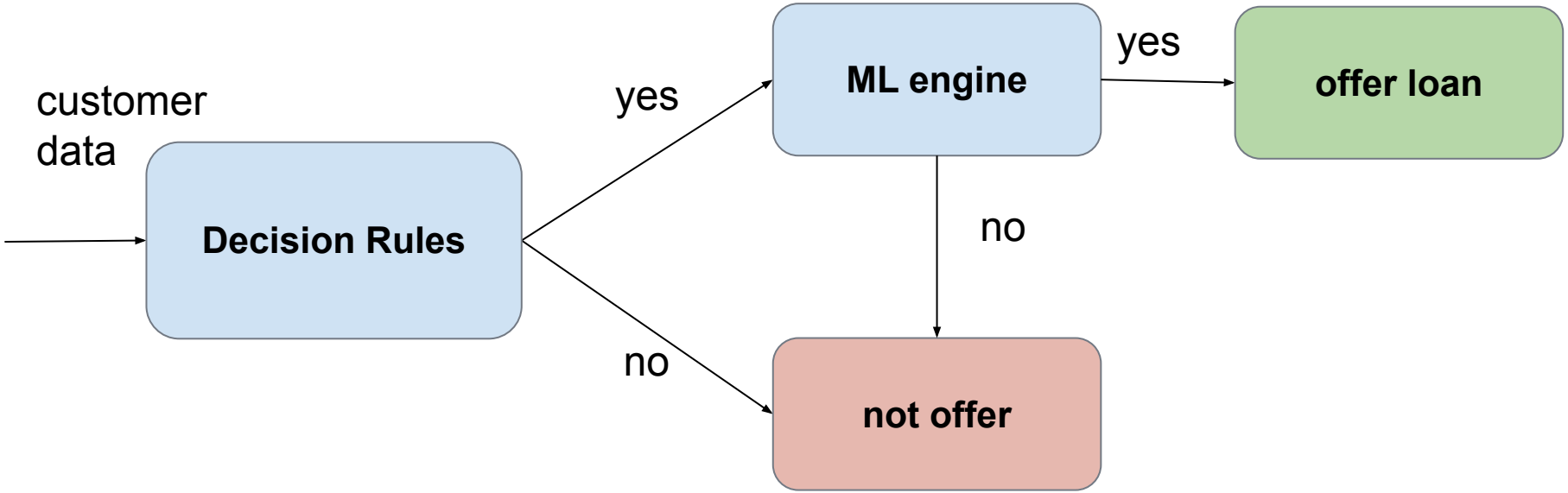
DEPLOYMENT

Option B: Use the model only for orders behind the business rule.



zalando

DEPLOYMENT

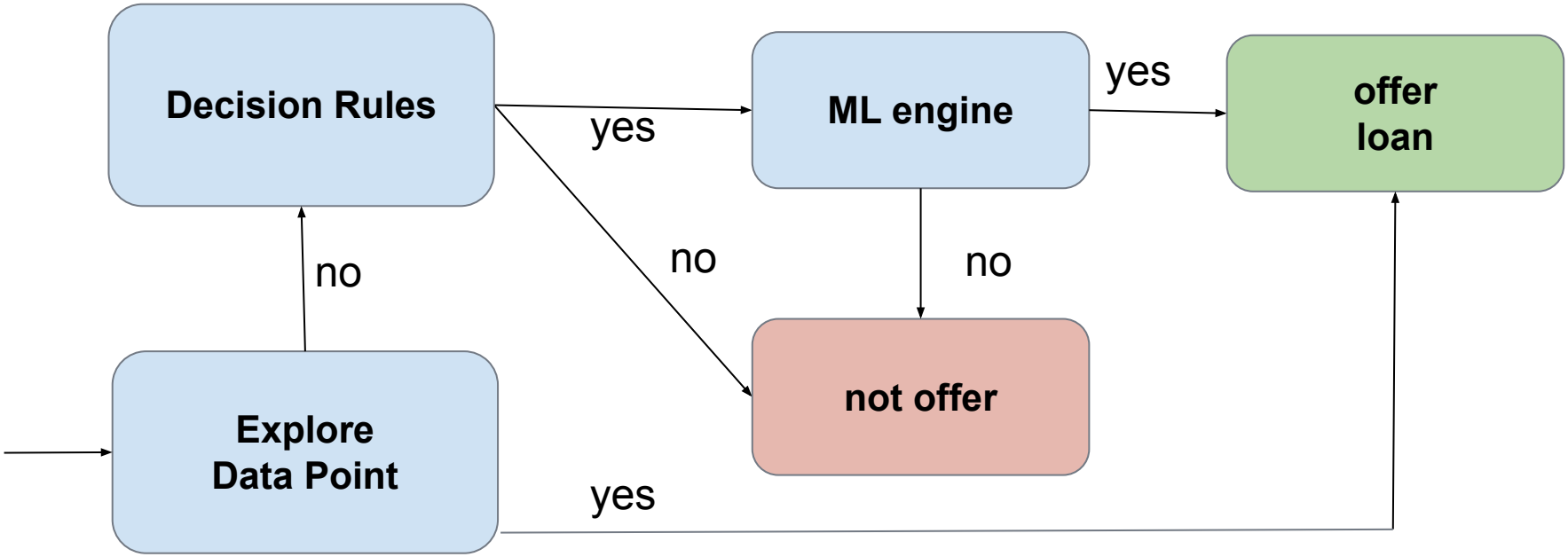


But we want to get rid of the decision rule!

zalando

DEPLOYMENT

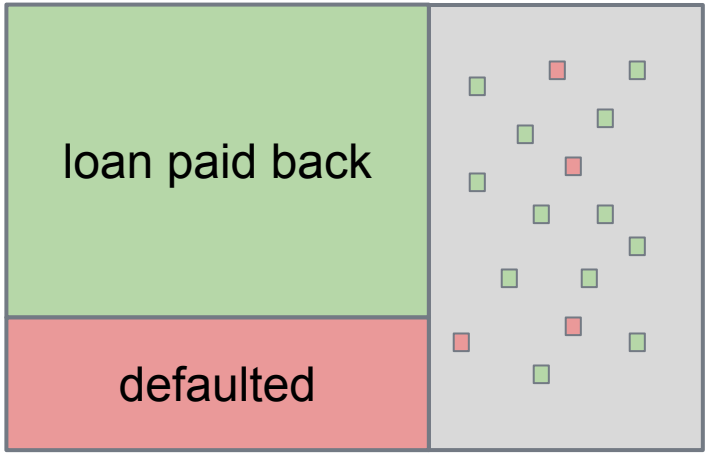
Option C: Use option B but start **exploring** the grey area.



zalando

DEPLOYMENT

Option C: Use option B but start **exploring** the grey area.



Exploration:

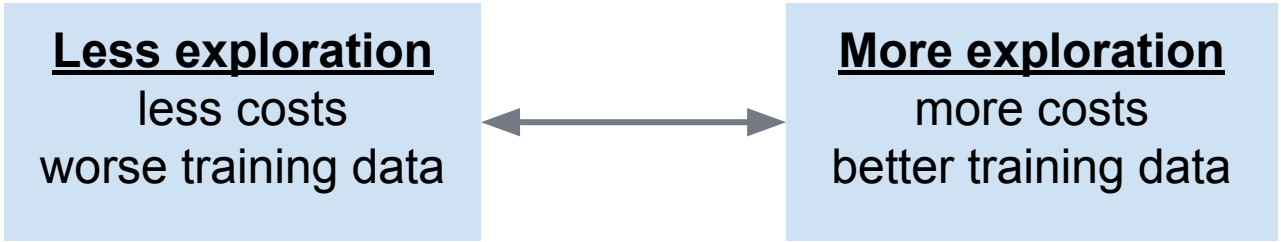
- Costs money!
- Only way to collect labels behind the decision rules

Questions:

1. How much exploration is needed?
2. When can we get rid of the rules?

EXPLORATION

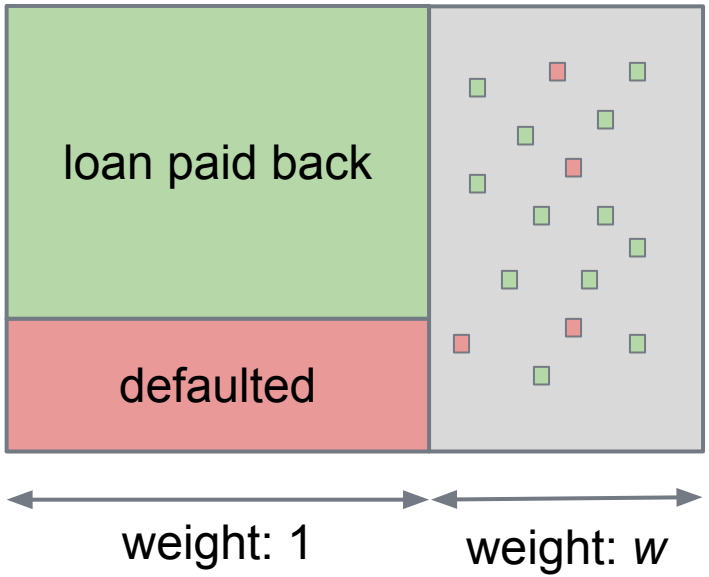
How much exploration is needed?



- Exploration vs. Exploitation tradeoff (study RL literature)
- As a start, we follow this suggestion [1] and explore 5% of our data

EXPLORATION

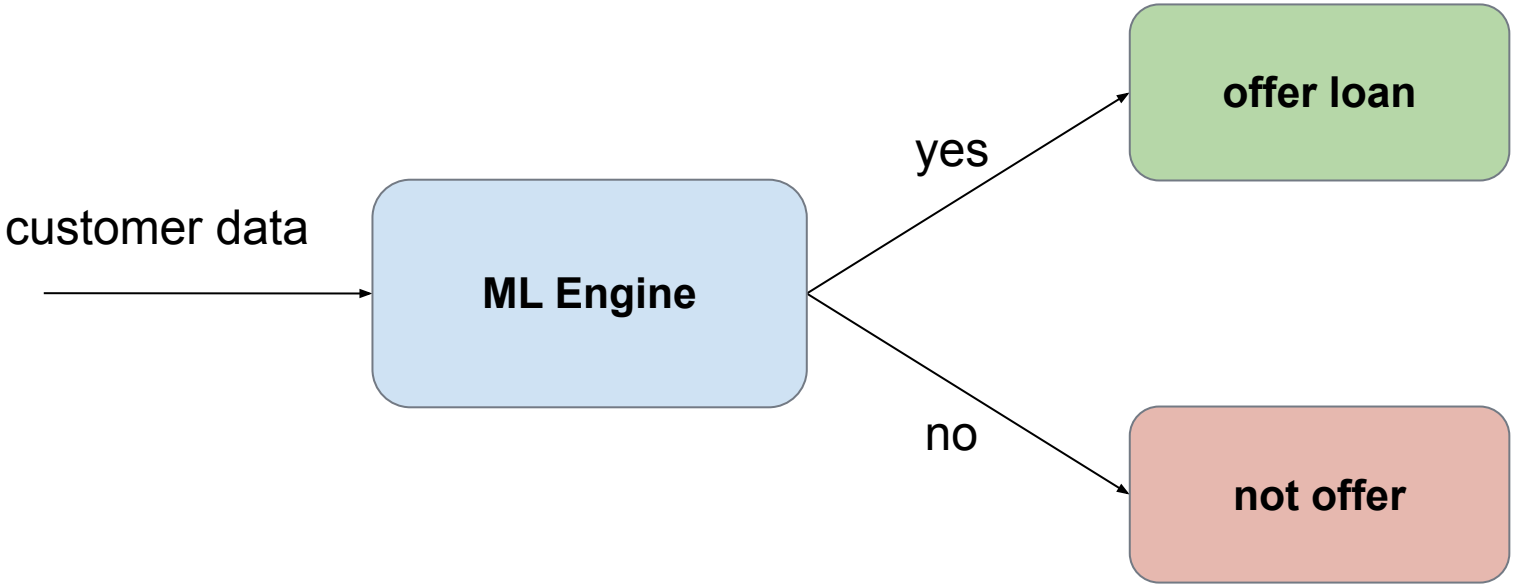
How to train a ML model?



For training, we need to up-sample the explored data points during learning accordingly:

$$w = \frac{\#\text{rejected} + \#\text{explored}}{\#\text{explored}}$$

DECISION ENGINE II



Hint: Exploration is still a good idea, if you want to use your outputs as future training data.

CELEBRATION TIME

They finally have a plan to get rid of the decision rule and to have a system that is running on pure AI …

CELEBRATION TIME

The new model is trained and deployed.

When should I switch off the business rule?

Great!
Switch it off now!

They are about to switch off the rule,

as suddenly a flashlight appears …

CELEBRATION TIME

Not so fast, young fintech pioneers!

A lot of money you might lose!

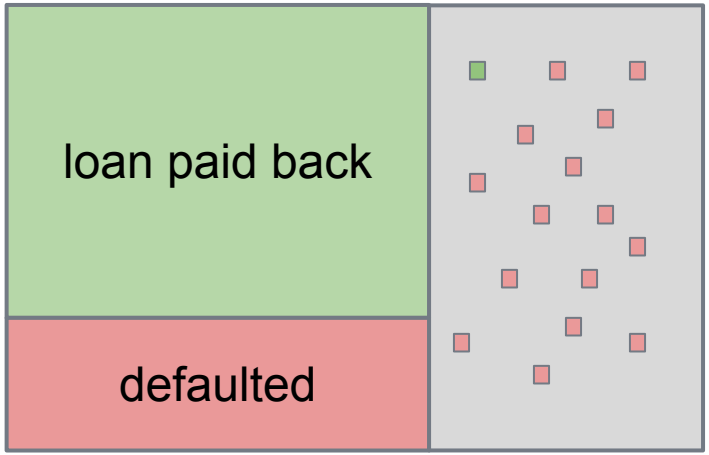
The decision rule, excellent it can be!

Switch it off now!

What does the Jedi mean?

REVISE DECISION RULE

It can be that the decision rule already does a fantastic job:



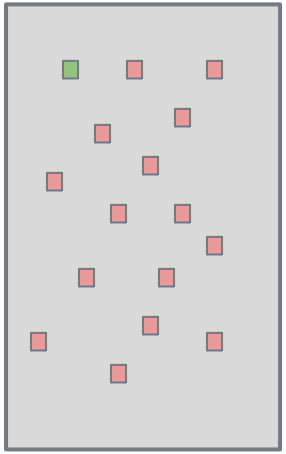
→ Almost all blocked loans are actually defaults

Is it better than the ML model?

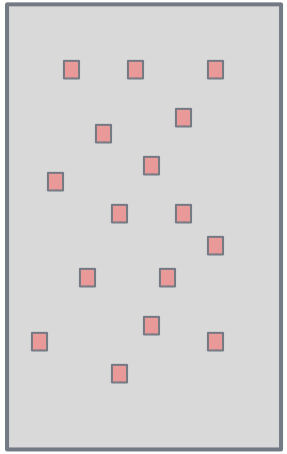
Compare the model and the rule on the explored data.

REVISE DECISION RULE

True Label

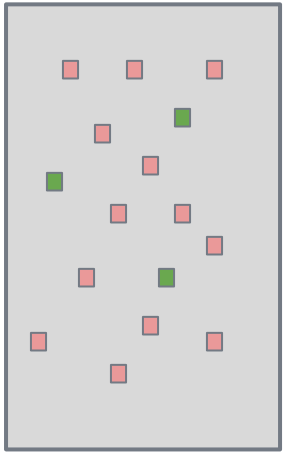


Explored Data classified by decision rule



1 false positive
0 false negatives

Explored Data classified by ML model

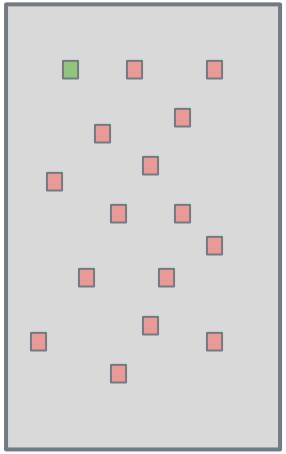


1 false positive
3 false negatives

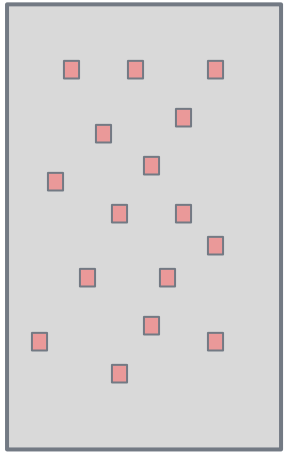
Switching rule off, money you will lose!

REVISE DECISION RULE

True Label

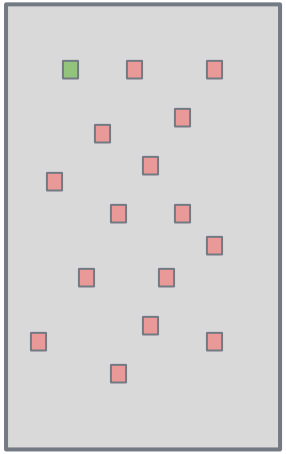


Explored Data classified by decision rule



1 false positive
0 false negatives

Explored Data classified by ML model



0 false positive
0 false negatives

Let's finally switch off the rule!

ML vs. Decision Rules

Comparison

Why can Decision Rules be good?

- ML model is usually modeling more than one pattern, focus on all patterns is limited by model complexity (which is limited by available data history)
- They are usually quicker to implement and to put in production.
- They do not require a rich data history, if designed by a domain expert (warning: not taking the data driven approach).

Comparison

Why can Decision Rules be bad?

- They can be biased and not performing when designed by a domain expert (instead of taking a data driven approach).
- They add complexity to the decision logic.
- They are hard to get rid off (as we have seen).

CONCLUSION

- Decision Rules are widely-used in practice
- They are usually used for starting off (when no data available)
- They can have a justification beyond that
- ML will usually do a better job (once you have enough data).
- Hence, ML should be the prefered way for decision making …
- … but only once you can be sure that it is doing better decisions!

THANK YOU!

Patrick Baier & Henning Esser

https://tech.zalando.com/blog/scalable-fraud-detection-fashion-platform