Machine learned search: setting up a production pipeline

Zalando SE

Maximilian Werk

Senior Research Engineer

20-01-2020

Picture from Pexels.com

Machine learned search: setting up a production pipeline

Zalando SE

Maximilian Werk

Senior Research Engineer

20-01-2020

Programming vs. Software Engineering

Who is a Software Engineer?

Who does programming in their day-to-day work?

Who has a machine learning background?

DAMEN | HERREN | KINDER

zalando

DE | EN Anmelden Wunschzettel Warenkorb

Get the Look NEU Bekleidung Schuhe Sport Accessoires Beauty Premium Marken Sale %

boho kleid

Bekleidung
Schuhe
Sport
Accessoires
Premium
Beauty
Sale
Geschenkgutscheine

'boho kleid'

Sortieren | Größe | Marke | Farbe | Preis | Obermaterial | Muster | **Mehr Filter**

33 Artikel

Gesponsert

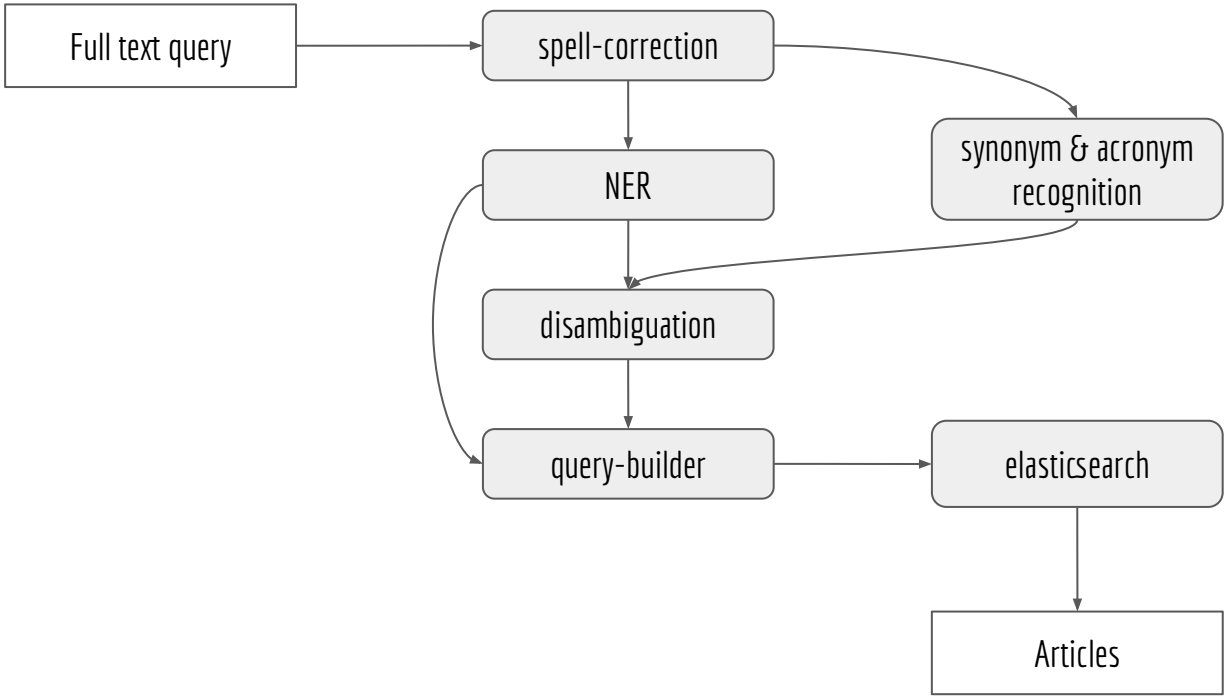
WEEKEND MaxMara 374,95 €
TORNADO - Strickkleid - anthrazit

YAS 49,99 €
Freizeitkleid - sudan brown

Nly by Nelly ~~54,95 €~~
BOHO TUNIC DRESS - Freizeitkleid... **49,49 €**

zalando

Our Information Retrieval Pipeline



zalando

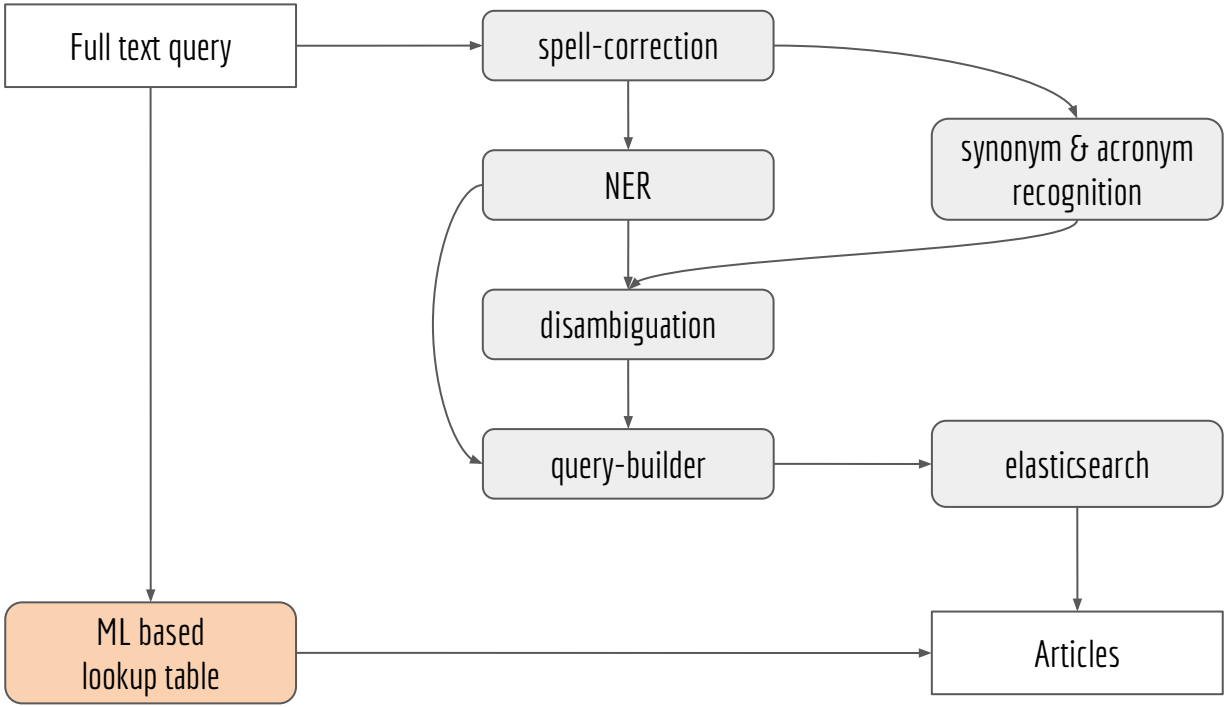
Failing classical information retrieval

“pullover patchwork”

“top figurumspielend”

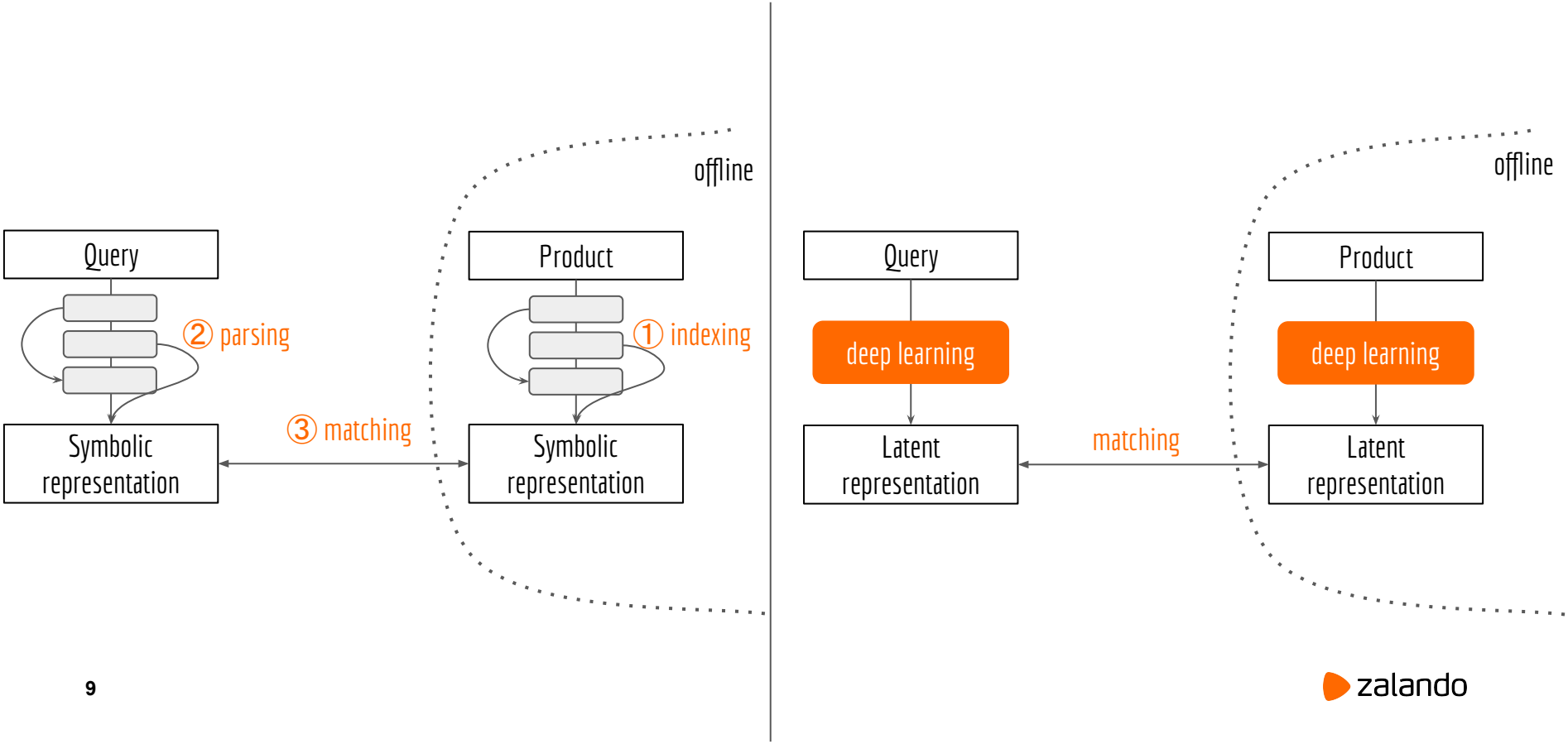
“abendkleid tattoospitze”

Adding ML based solution



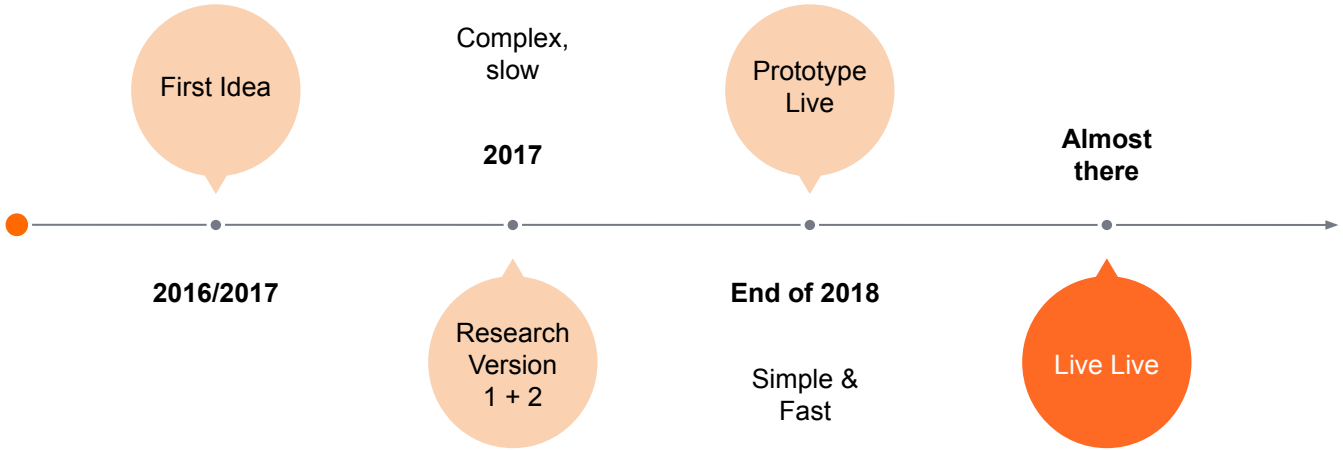
zalando

Classical system vs. end-to-end product search system



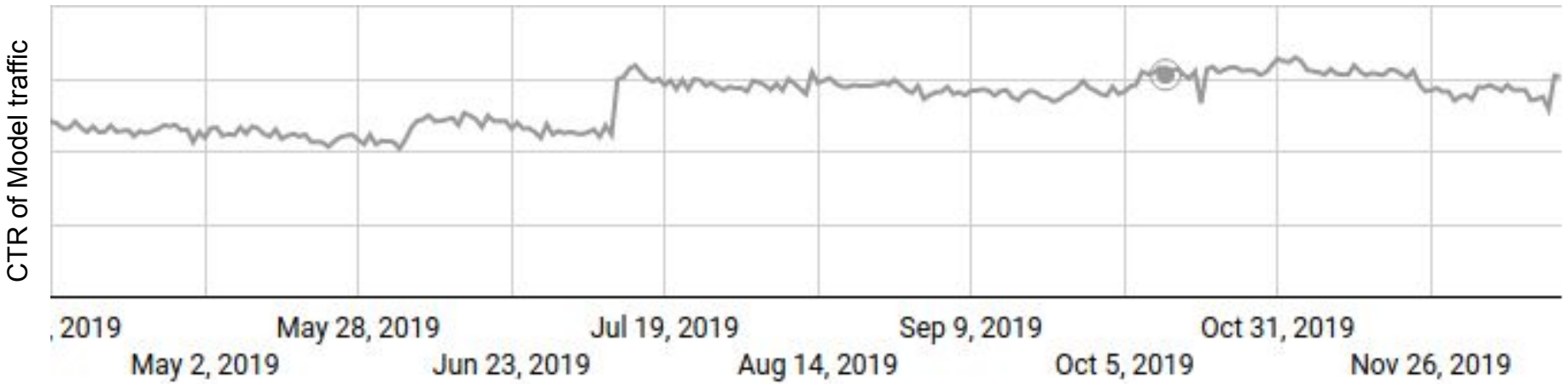
zalando

History



zalando

Model degradation



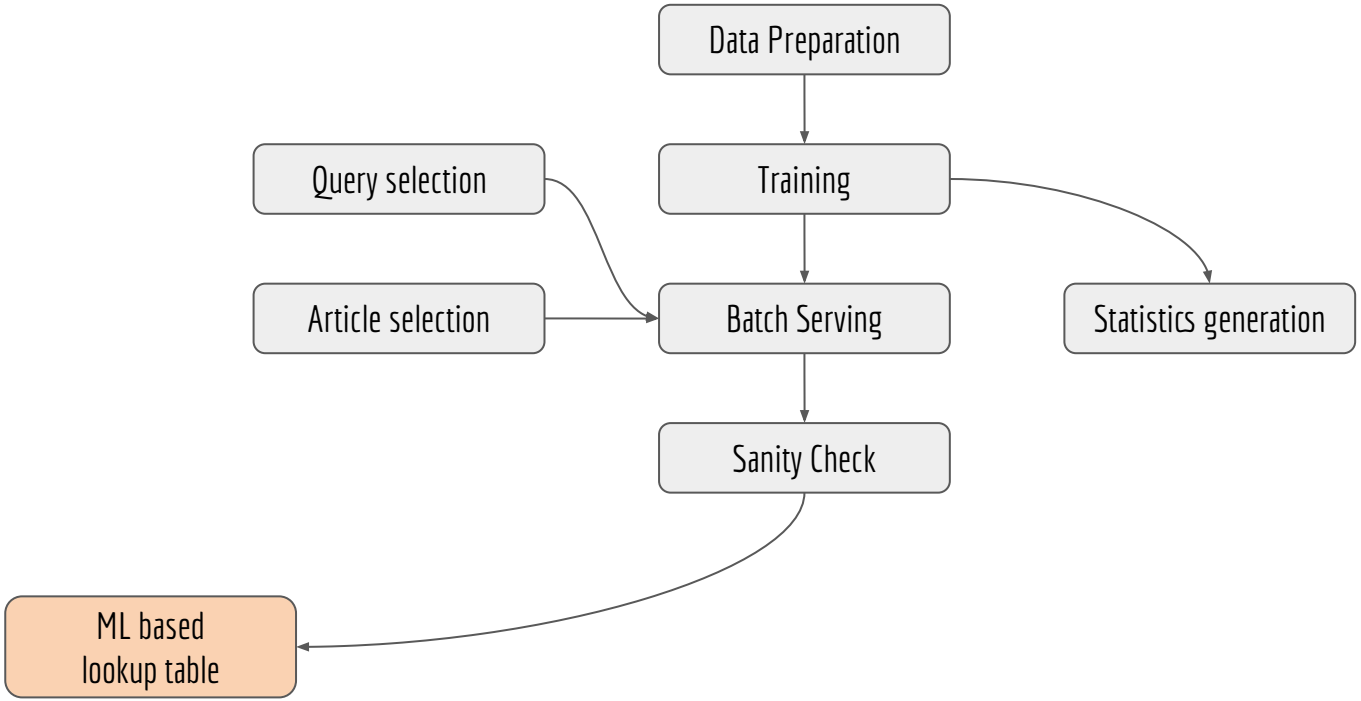
zalando

Building a pipeline

 Picture from Pexels.com

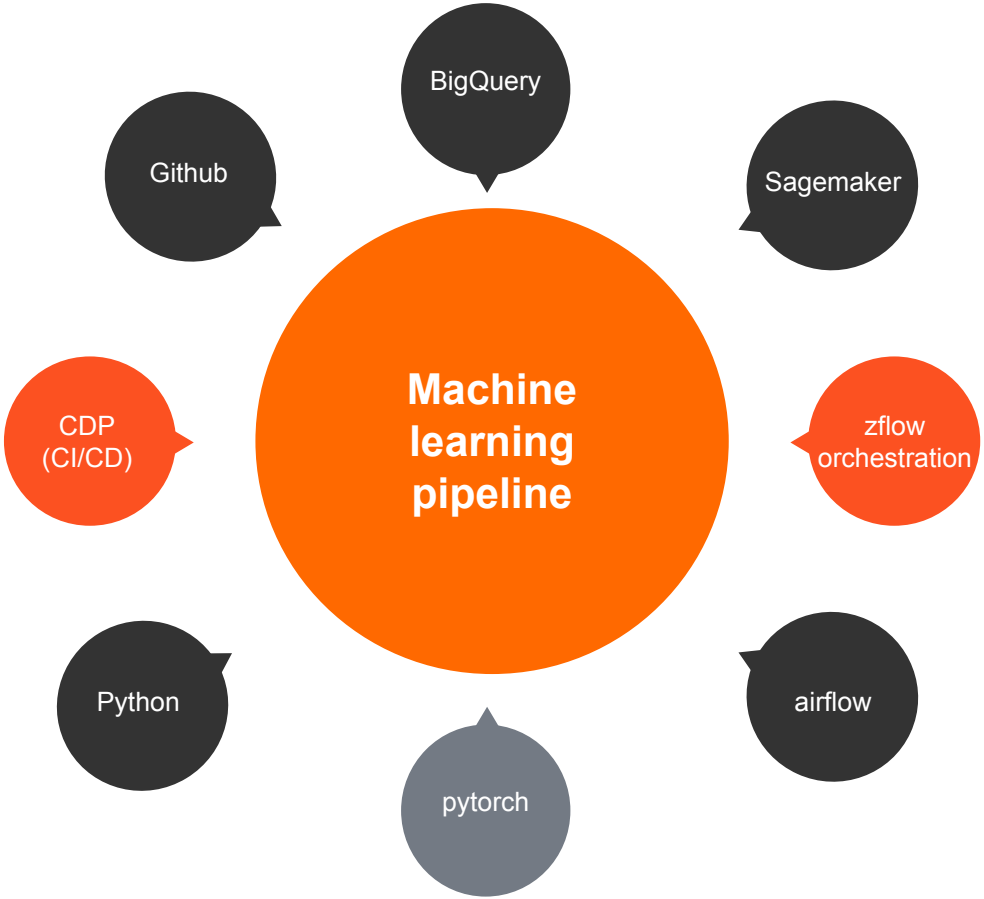
zalando

ML pipeline



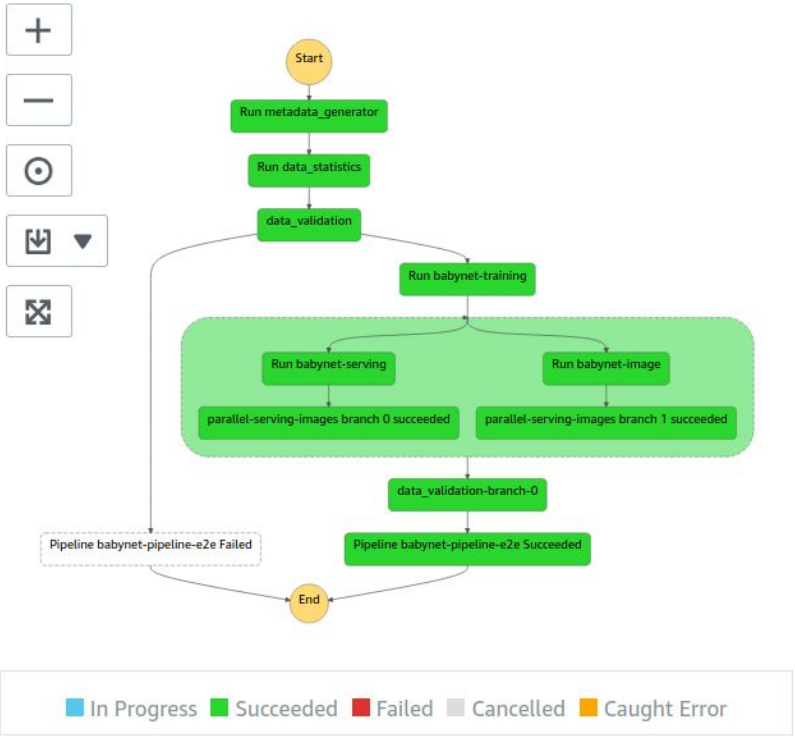
zalando

Our Technology Stack



zalando

Visual workflow



Code | Step details

Name: Run babynet-image

Type: Task

Status: Succeeded

Resource: [redacted]

▶ Input

▶ Output

▶ Exception

zalando

One-time job vs. continuous development
“Obvious decisions”

Complex model vs. simple model

Batch vs. live

Manual vs. automated

Training cadence (daily, weekly, monthly, irregularly)

One-time job vs. continuous development
“Hidden decisions”

Fast now vs. fast, maintainable, robust later

Scripting vs. Software Development

Talk, talk talk!

 Picture from Pexels.com

Testing is hard

 Picture from Pexels.com

Configuration is complex

 Picture from Pexels.com

Follow standards & best practices

 Picture from Pexels.com

Do clean code

```python
def p(i):
    s = i.split("|")
    d = s[0].lower()
    # Cleaning the price string into byte number representation and converting it to cents
    p = int(float(s[1].replace(",", ".").replace("€", "")) * 100)
    return d, p
```

BÄÄÄH

```python
from collections import namedtuple

Article = namedtuple("Article", ["description", "price"])


def parse_article(article_string):
    splitted = article_string.split("|")
    description = splitted[0].lower()

    cleaned_price_string = splitted[1].replace(",", ".").replace("€", "")
    price_in_cents = int(float(cleaned_price_string) * 100)

    return Article(description, price_in_cents)
```

Nice

About good (ML) code

Correct

Simple functions

Written for others to read

Accessible business logic

Pipeline steps are independently executable

Producing good code

1) Feature
2) Correct
3) Readable
4) Simple
5) Readable
6) Feature is still correct?
    a) No => go to 1)
    b) Yes => Happy days!

Be a scout: Leave the code cleaner than you found it.

Future work

Improve Model

Monitoring

Model influences training data

Add more use cases

Replace tradition IR search

Takeaways

Simple model does the job

Pipeline building takes a lot of time

Train code craftsmanship

zalando

Maximilian Werk

Search - Team Lens

Senior Research Engineer

maximilian@zalando.de

Twitter: @maintainable_ds

We hire:

Senior Search Engineer

Principal Research Engineer - Search

Principal Product Manager - Search

20-01-2020