
Use case definition

Sprint 1 - Week 1

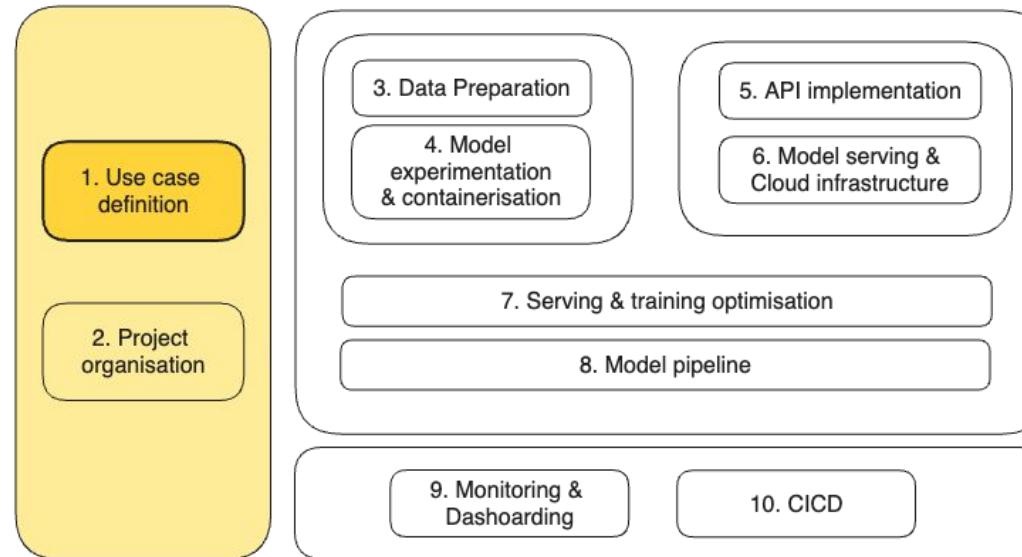
INFO 9023 - Machine Learning Systems Design

2024 H1

Thomas Vrancken (t.vrancken@uliege.be)
Matthias Pirlet (matthias.pirlet@uliege.be)

Status on our overall course roadmap

Sprint 1:
Project organisation



Agenda

What will we talk about today

Lecture (2 hours)

1. Use case deep dives
 2. Project phases & challenges
 3. Project definition framework
-
- Bonus: Explainable AI (XAI)

Use case deep dives

Real-estate valuation
assistant

Context & Problem Statement

...heard of Fednot?



Fednot

- = Koninklijke **Federatie van het Belgisch Notariaat**
- = **Fédération Royale du Notariat belge**
- = Royal **Federation of the Belgian Notaryship**

Fednot supports the notary studies with juridical advice, office management, IT solutions, trainings, and information for the general public.

Valuation assistant.

FEDNOT | Waarderingsassistent immo

Thomas UYTTELENHOVE
TEST ETUDE 12

TERUG | RESULTAAT

Dataset van het pand

Adres
10 Sportstraat, 9000 - Gent
Perceelnummer
4480910810/00F006

Waardering

Resultaat van de waardering

Indicatieve prijs
€ 397 000

Indicatieve prijs en distributie van de geïndexeerde verkoopprijs van 319 huizen binnen een straal van 1 km.

HOE HEEFT HET MODEL DEZE INDICATIEVE PRIJS BEREIKT?

Services 1.8.0 UI 8.9.0

How the ML model conceptually works

Known values that the model will use as input to make predictions						What the model needs to predict
Feature variables						Target variables
ID	Size (sqft)	Bedrooms	Bathrooms	Distance to City Center	Garage	House Price (k\$)
1	2200	3	2	5	Yes	300
2	1800	4	2	3	No	275
3	1400	2	1	10	Yes	200
...
80 000	3000	5	4	4	Yes	400
80 001	1600	3	2	12	No	? (Test)
...
100 000	2100	4	2	6	Yes	? (Test)

Single house

Train set

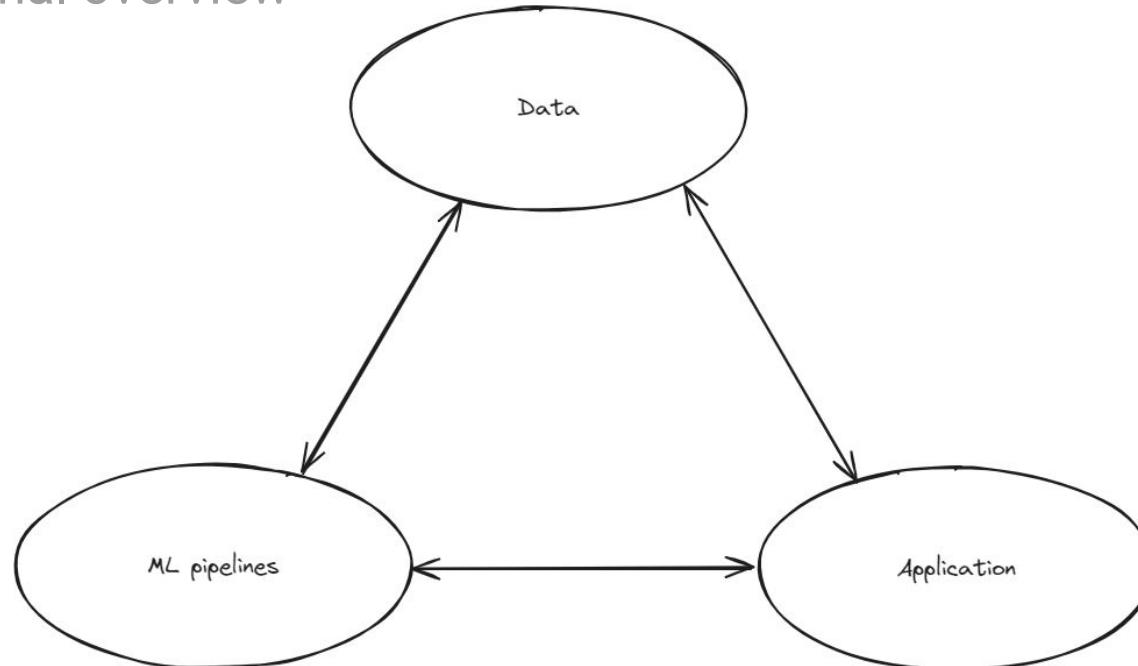
Test set

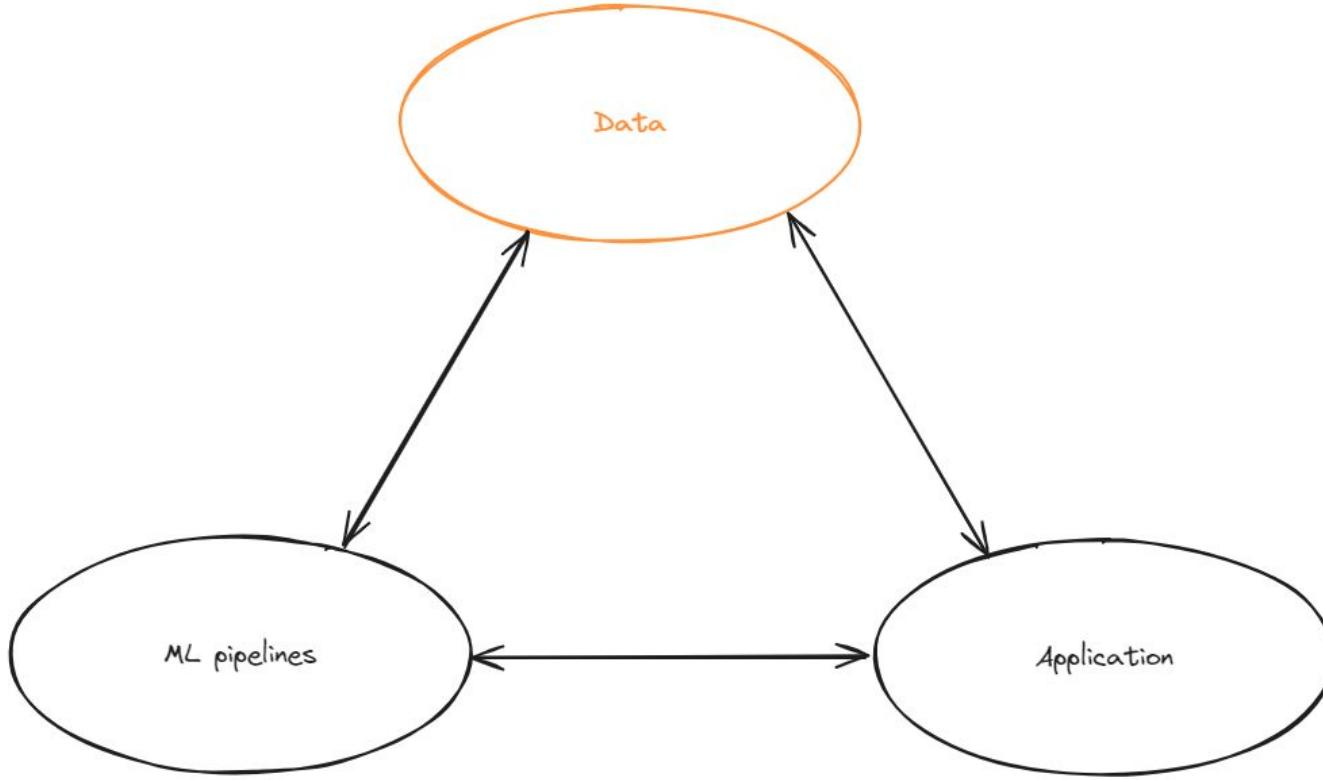
The ML model will see many **observations** (houses) defined as a set of features (information, variables). From it the model will learn patterns and what impacts **target variable** (house prices).

If given new observations, the model can **predict** the target variable based on the input features.

Solution architecture.

A functional overview





Valuation features.

Starting from unstructured PDF documents

Buimtelijke Ordening:
De koper verzaakt uitdrukkelijk aan de mogelijkheid om de nietigheid van onderhavige verkoop in te roepen bij gebrek aan informatie.

9) Risicozone overstroming:
Ingevolge oproeping gedaan op **25 januari 2018** verklaart ondergetekende Notaris in navolging van artikel 129 §4 van de wet betreffende de verzekeringen van 4 april 2014, dat het hierboven vermelde goed **niet** gelegen is in een risicozone voor overstromingen.

Ingevolge zelfde oproeking, verklaart ondergetekende Notaris in navolging van artikel 17bis van het Decreet van 18 juli 2003, gewijzigd in 2013 betreffende het integraal waterbeleid, dat het hierboven vermeld goed:

- niet gelegen is in een mogelijk overstromingsgevoelig gebied;
- dient beschouwd te worden als gebieden die uitsluitend bij heel extreme weersomstandigheden of bij een defect aan de waterkering overstromen;
- niet gelegen is in een effectief overstromingsgevoelig gebied;
- dient beschouwd te worden als gebieden waar recent nog een overstromingsincident of gebieden waarvan modellen aangegeven dat er in de honderd jaar (of frequenter) een overstroming plaatsvindt;
- niet gelegen is in een afgebakend overstromingsgebied;
- niet gelegen is in een afgebakend overzomen.

10) Postinterventiedossier (Koninklijk Besluit van vijf en twintig januari tweeduizend en één):
De verkopers verklaarden dat er aan het verkochte goed sinds één maal tweeduizend en één werken werden uitgevoerd. De verkopers verbinden er zich toe dit dossier aan de kopers te overhandigen uiterlijk binnen de 6 maanden te rekenen vanaf heden.

11) Stoekolietank
De verkopers verklaarden dat er in voorschreven goed geen stoekolietank aanwezig is.

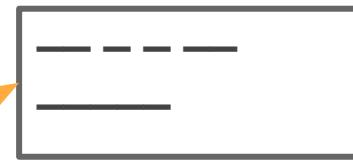
12) Elektrische installatie. De verkopers verklaarden dat het onroerend goed, voorwerp van huidige verkoop, een woonhuis is in de zin van artikel 276 bis van het Algemeen Reglement op de Elektrische Installaties van 10 maart 1980.

De verkopers overhandigen bij deze aan de koper, heropen deze erkent, het proces-verbaal van controleonderzoek opgemaakt door de Vereniging Zonder Winstoogmerk OCB op 3

Clauses



Entities



Features



Valuation features.

Processing the texts from deeds

We want to extract **named entities!**

- Persons
- Dates
- Addresses
- Locations
- ...

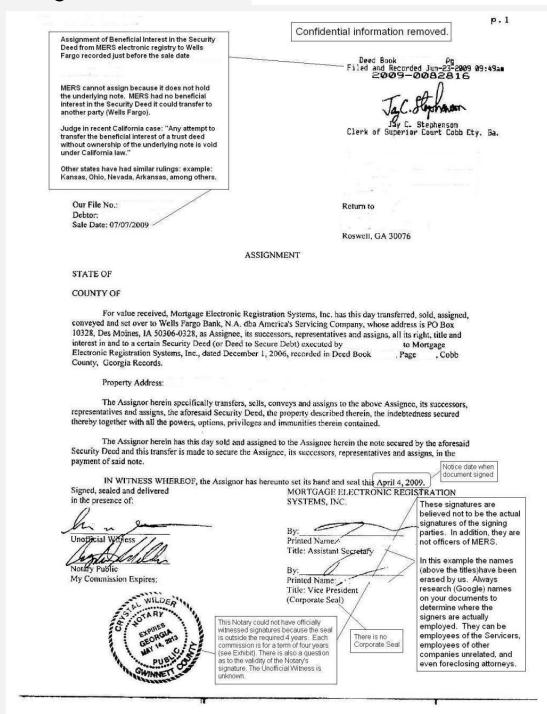
Erfdienstbaarheden

Het goed wordt verkocht met al zijn gekende en ongekende, zichtbare en onzichtbare, voortdurende en niet-voortdurende erfdiestbaarheden en zakelijke rechten en verplichtingen. De verkoper verklaart geen weet te hebben dat het goed is bezwaard met onzichtbare erfdiestbaarheden behoudens de erfdiestbaarheid vermeld in de eerder aangehaalde eigendomsakte verleden voor notaris **Hans Deprez** op **14 oktober 2012**, waarin letterlijk het volgende wordt vermeld:
“Ten titel van inlichting en zonder de bedoeling exhaustief te zijn, worden de volgende erfdiestbaarheden aangehaald, zoals deze vermeld staan in de akte verleden door notaris **Hendrickx te Antwerpen** op **7 april 2014** overgeschreven inhoudende verkoop door de vennootschap aan de heer **Peeters Tom**.
“Over het verkochte goed te **Stationsstraat 148** wordt onvergeld en eeuwigdurend een erfdiestbaarheid van doorgang behouden in het voordeel van het erf van de bewoner.
Deze doorgang zal mogen gebruikt worden zowel te voet als met gemotoriseerde voertuigen, om de terreinen achter de bestaande gebouwen van de verkoper en deze gebouwen zelf langs hun achterzijde te kunnen bereiken of verlaten, doch enkel in nood gevallen (bijvoorbeeld door de brandweer of andere hulpdiensten).
Om in voorkomend geval te allen tijde en ten gerieve van de aanpalende erven van verkoper de vrije doorvaart te verzekeren, zal vijf meter achter de bestaande gebouwen van verkoper, nooit enige constructie, aanplanting of hindernis mogen opgesteld worden, inclus enige obstructie ten belope van de bedrijfsuitvoering van **Dynamo BVBA** die hieronder omschreven staat als zijnde de primaire en op het moment van schrijven enige...”

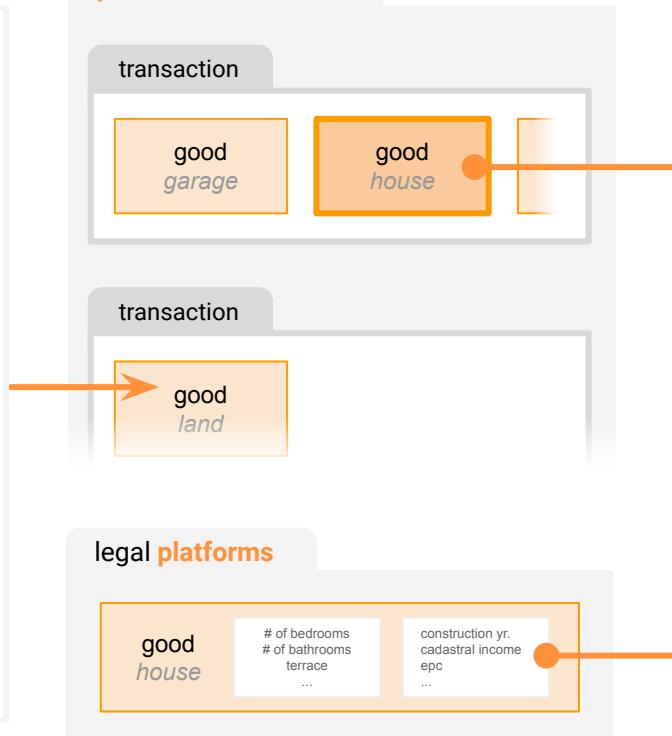
Valuation features.

Legal real-estate data

legal document



processed document



Legal ('AI') features

- Transaction date
- Good (sub)type
- Price(s)
- Energy Performance Certificate (EPC)
- Newly built/existing
- Cadastral income
- ...

Valuation features.

Legal real-estate data

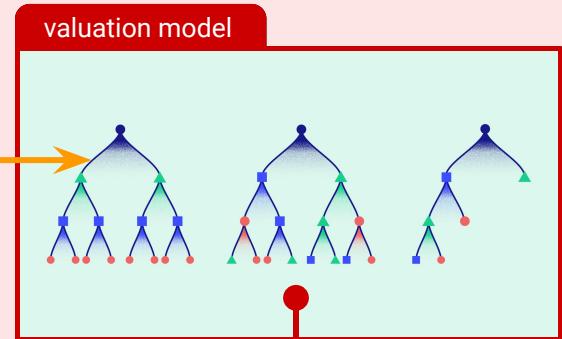
processed document

legal platforms

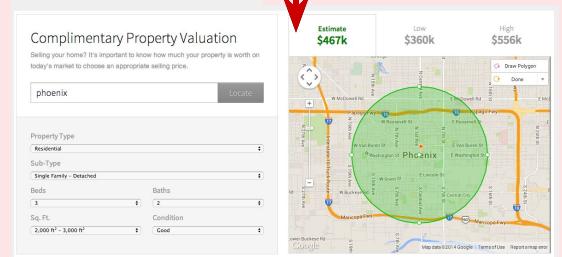


Legal ('AI') features

Legal features alone do not
capture sufficient
information to accurately
predict real estate prices...



valuation tool

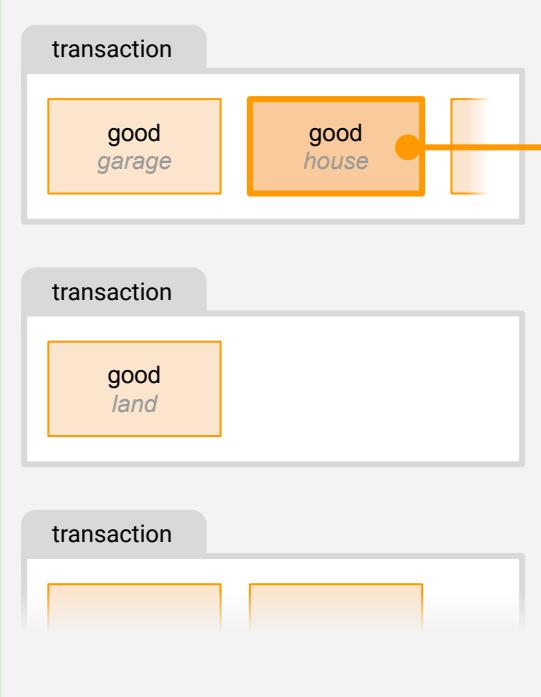


Valuation features.

Open real-estate data

processed document

legal platforms

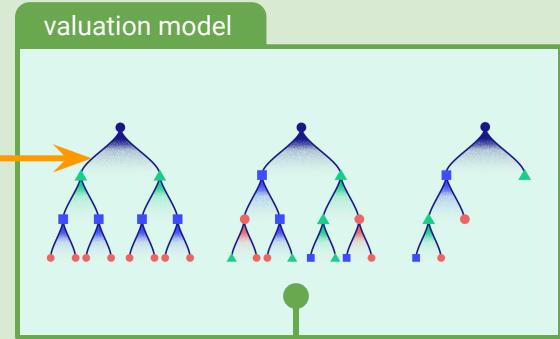


Legal ('AI') features

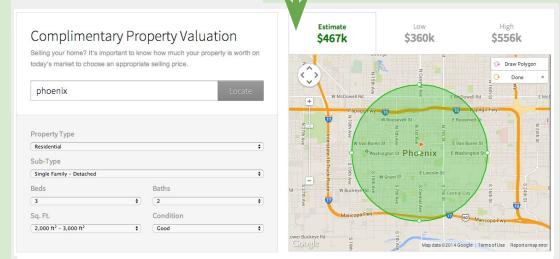


Leverage open data to expand limited feature set

- **Various types** of data sources provided by the government or community
- (Mostly) **freely accessible**
- Opportunities for **complex, more informative features!**



valuation tool



Valuation features.

Open real-estate data

■ Cadastral information

- Parcel area, street width, ratio, and orientation;
- Building area, type (“open”, “half-open”, “closed”), facade width, and orientation



Valuation features.

Open real-estate data

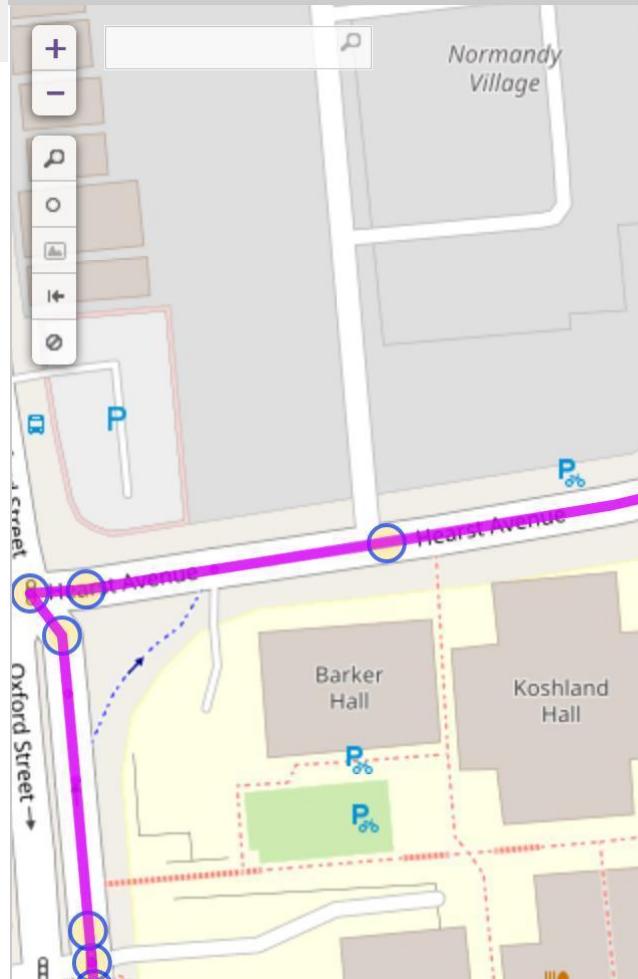
- Cadastral information
- Height information
 - Building height, volume, number of stories.



Valuation features.

Open real-estate data

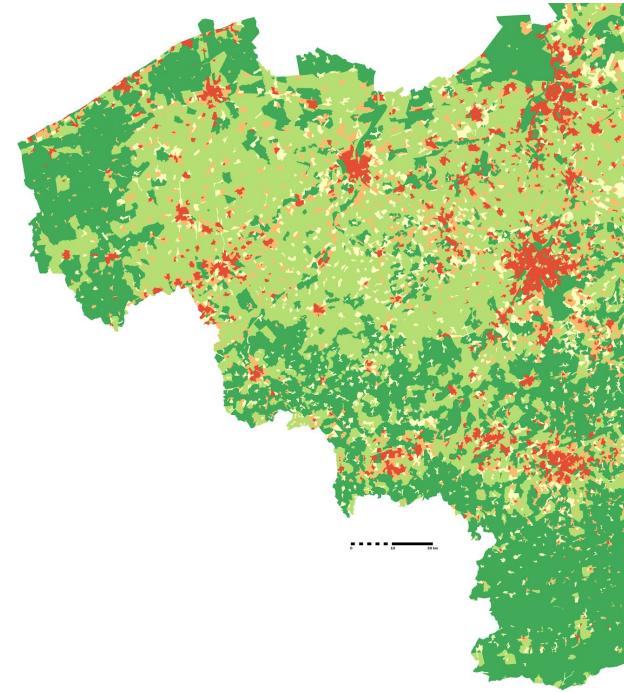
- Cadastral information
- Height information
- Location information
 - Distance to major cities and to nearest city center, highway (entry), primary road, railway, station, bus stop, etc.



Valuation features.

Open real-estate data

- Cadastral information
- Height information
- Location information
- **Local socio-economic and demographic statistics**
 - Municipality population size, tax percentage, prosperity index, avg. income;
 - Statistical sector cadastral income percentiles



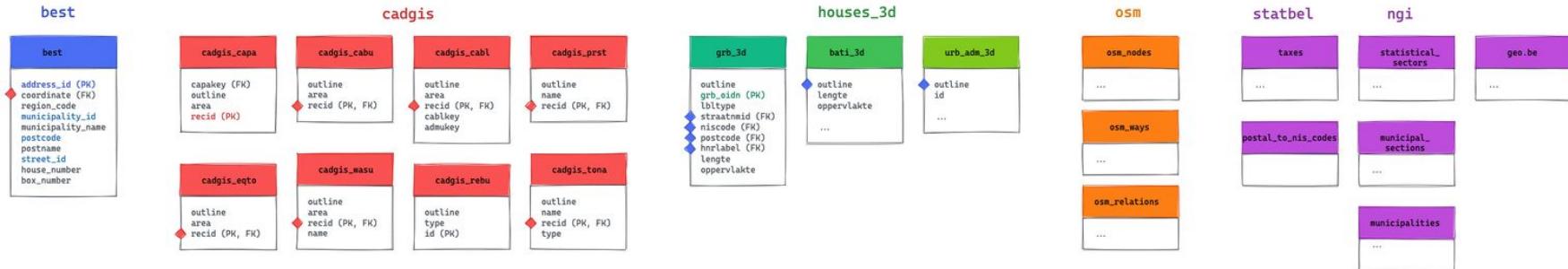
2011 Density (people / km²)

0 - 50.4
50.4 - 392
392 - 1080
1080 - 2120
2120 - 45700

Source: statbel.fgov.be

Open Data

Sources



Update Frequencies:

Weekly

Yearly

Wallonia - Not (2016)
Flanders - Not (2016)
Brussels - Yearly

Weekly

Yearly

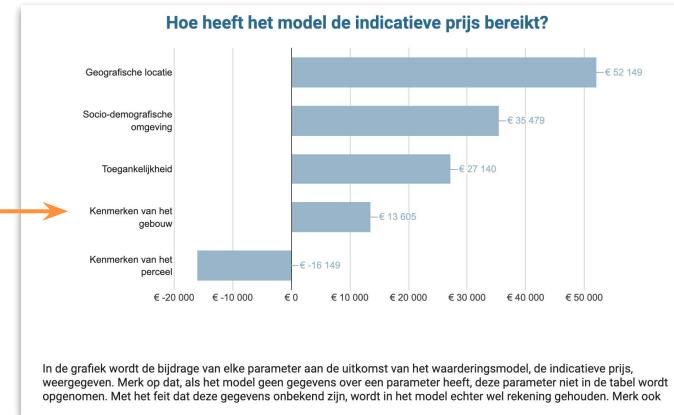
Yearly

Valuation model.

Price prediction model

Model is not an “oracle”, but supports notaries

- Important to express uncertainty.
- Shows indicative price compared to price distribution of similar properties.
- Explains how a prediction was obtained based on feature importances.



Deed data

Data validation with  great expectations

Your data assets:
database tables, flat
files, dataframes...



Data validation with
Great Expectations



High quality data in
your data products



Data documentation
& data quality reports

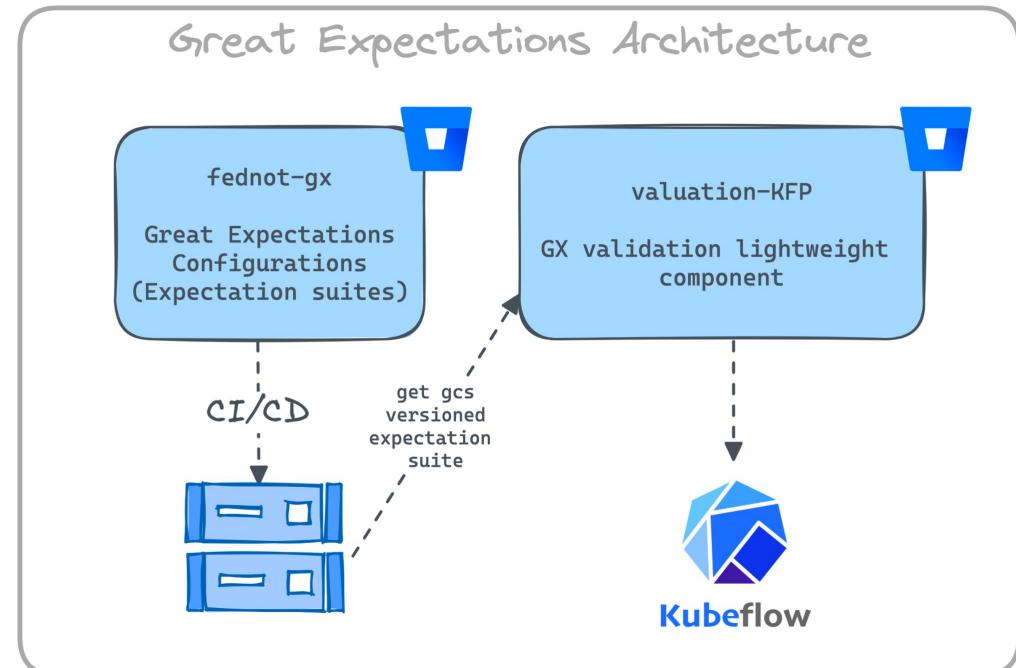


Logging & alerting

Deed data

Data validation with great expectations

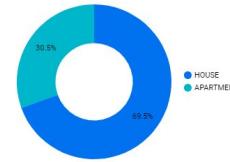
- Enable non-technical people to define expectations
 - **Expectations:** A JSON file that defines rules for validation. Eg: *feature_x should be 75% not null*
- Make the GX component reusable for multiple validation needs.



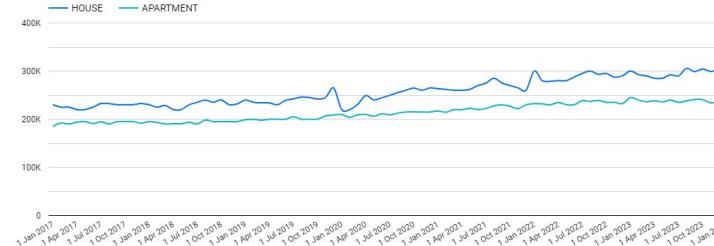
Deed data

Data quality dashboard

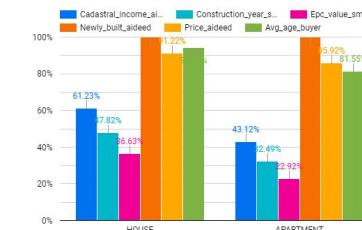
Market share



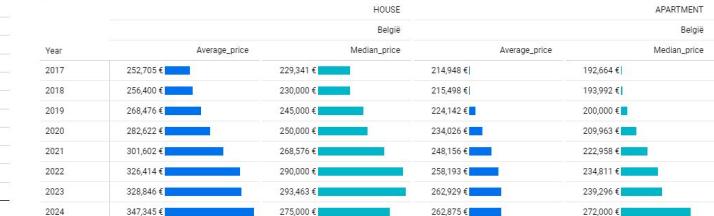
Median & Average Price



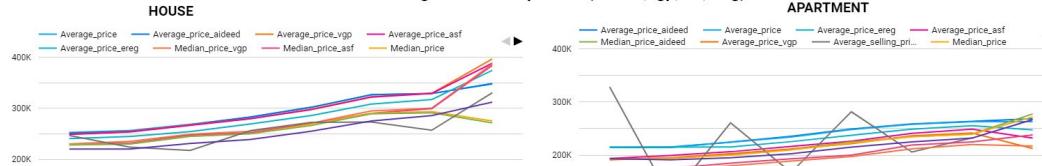
Percentage field presence



good_type / geo_level_value / Average_price / Median_price



Median & Average Prices comparison (aideed,vgp,asf,ereg)



Outlier detection

■ Business rules to focus on relevant data

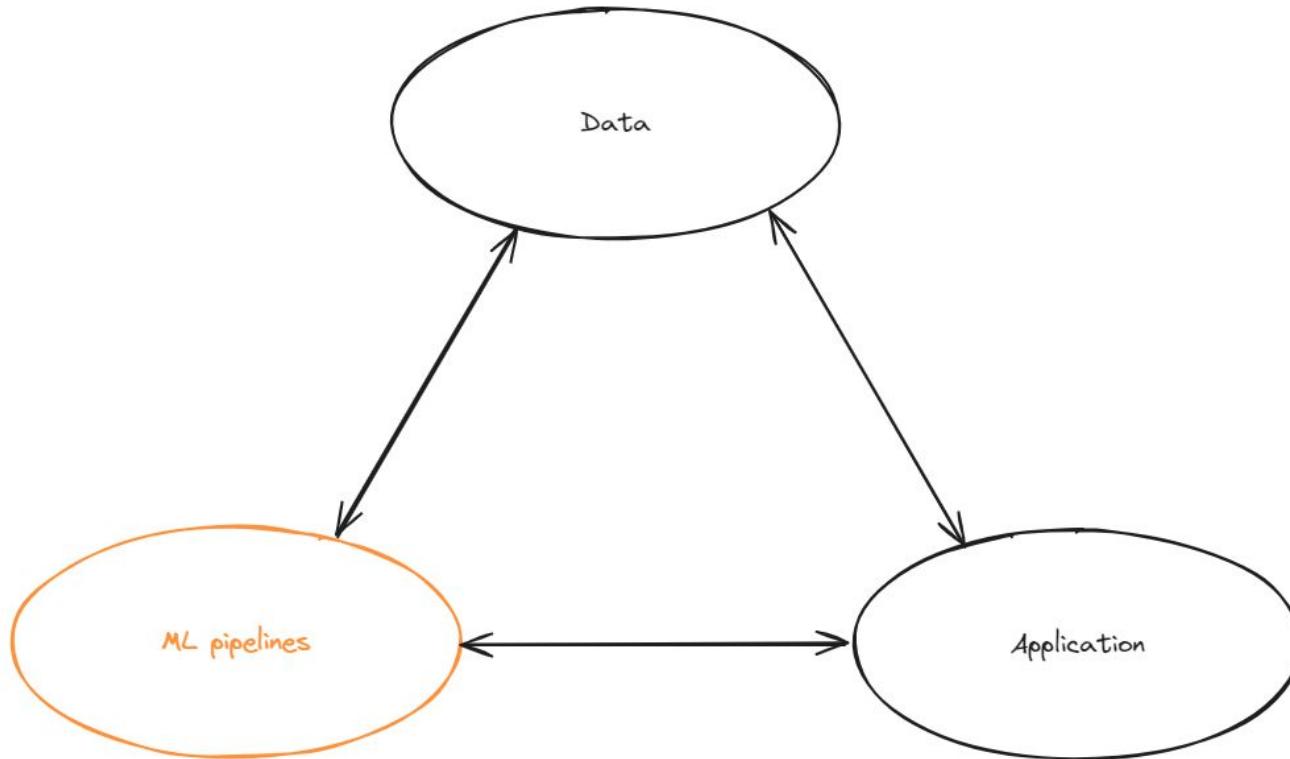
- Answering the question “What does a normal house/apt/... look like?”
- Filter out anything unrealistic (tiny or huge prices, parcel, heights, amount of box nrs.)
- Not all are statistically relevant...

■ Detect actual outliers

- Can occur on different levels and for different reasons
- Consider features as is ⇔ Consider features in their context
- Ensemble of methods to detect most anomalous data points

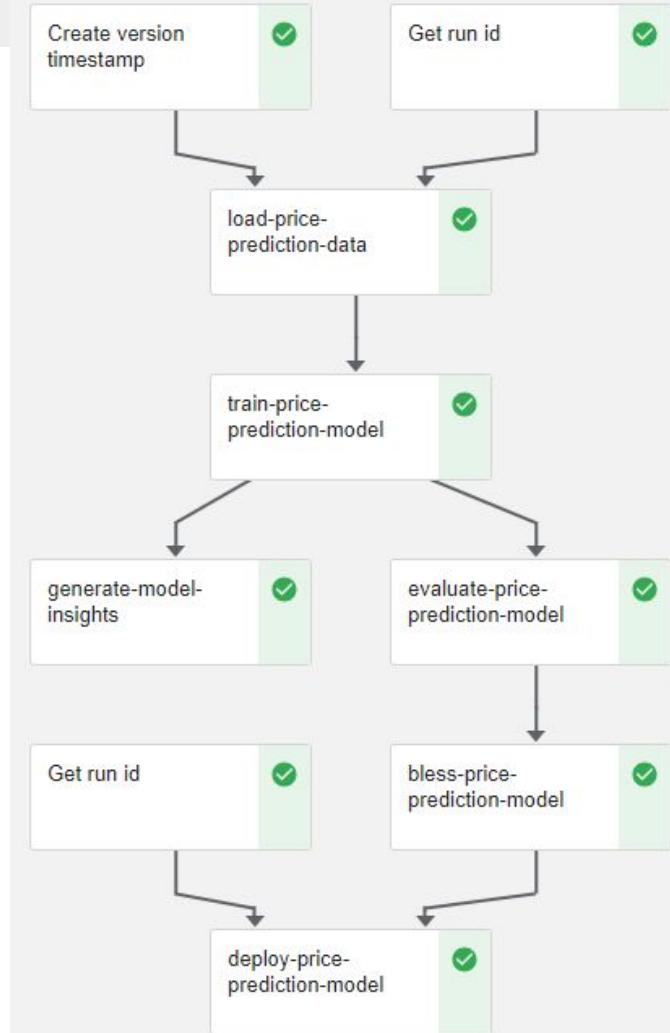
■ Multiple and different goals

- Barometer and valuation assistant have different goals and needs



Automated pipeline to train and deploy new price prediction models

- Launch the full training and deployment for new data or with new configurations
- Separates logic of different components
- Allows to reuse components
- Allows to work on components individually

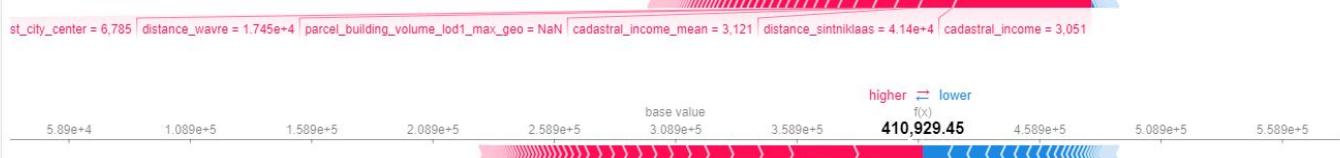
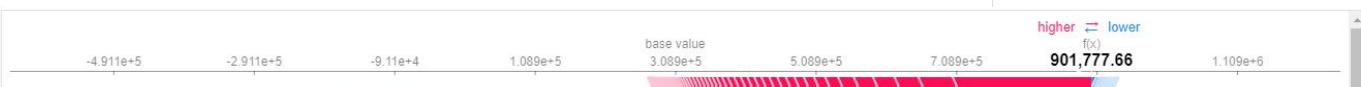


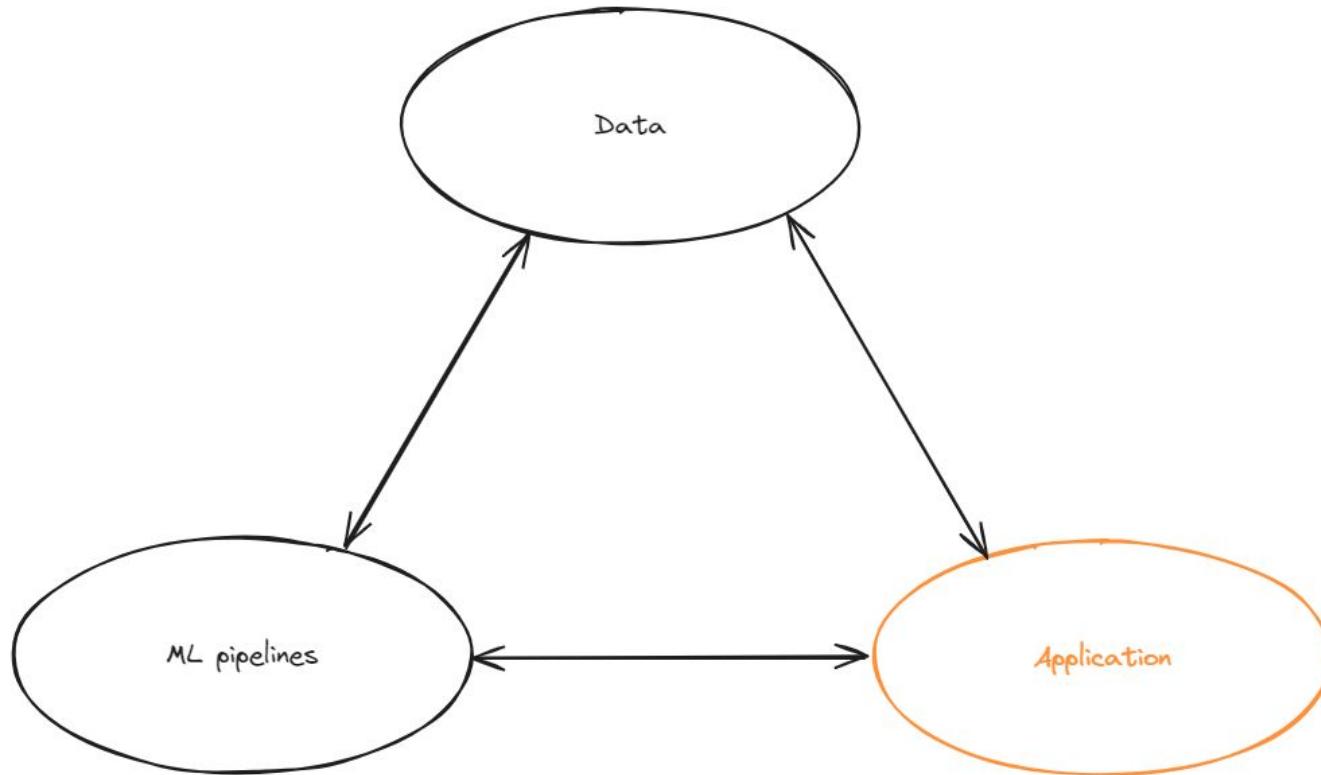
Visualisations

Explainability

		VAN			WHT		
quantile	key_features	nr_observations	mae	mape	nr_observations	mae	mape
Q0 - Q10	0 or 1	1547	80006.176115	0.565104	870	47111.933130	0.505811
	2 or 3	1312	59277.557874	0.401544	1245	37980.680478	0.409488
Q10 - Q25	0 or 1	1983	38821.581166	0.161699	1238	26893.919732	0.220017
	2 or 3	2299	29467.034478	0.121787	1950	22944.052560	0.187566
Q25 - Q50	0 or 1	3225	34544.096558	0.111164	2039	23278.130844	0.142917
	2 or 3	3892	30662.822366	0.098830	3279	21559.459244	0.132382
Q50 - Q75	0 or 1	3335	52391.982770	0.132128	2077	36009.678320	0.164627
	2 or 3	3782	43711.812300	0.110839	3296	32903.844697	0.150859
Q75 - Q90	0 or 1	1945	90583.684565	0.177171	1328	49428.070282	0.169335
	2 or 3	2329	69172.979267	0.134752	1858	44220.675478	0.150880
Q90 - Q100	0 or 1	1311	229875.407221	0.280143	912	116302.847821	0.245235
	2 or 3	1539	175906.829581	0.213262	1200	92134.481717	0.194406

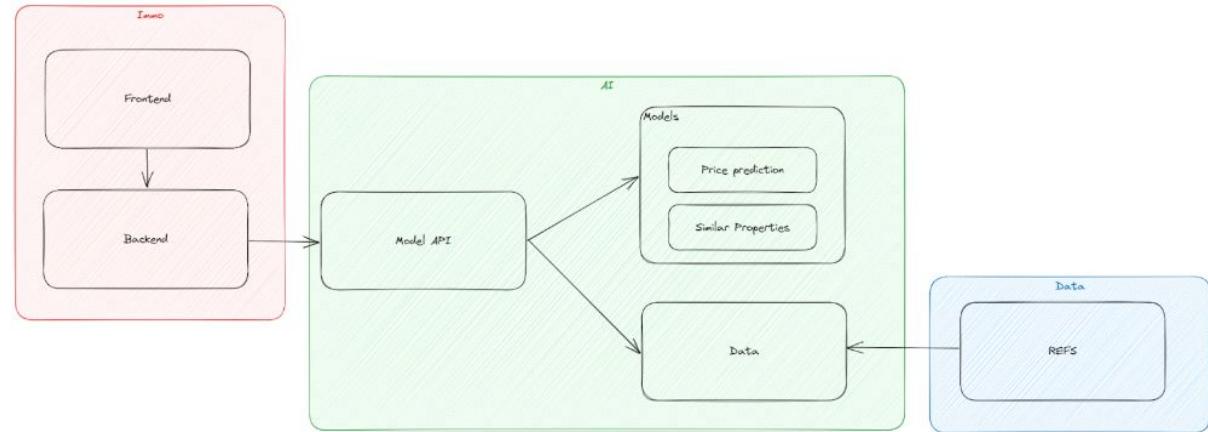
Static HTML





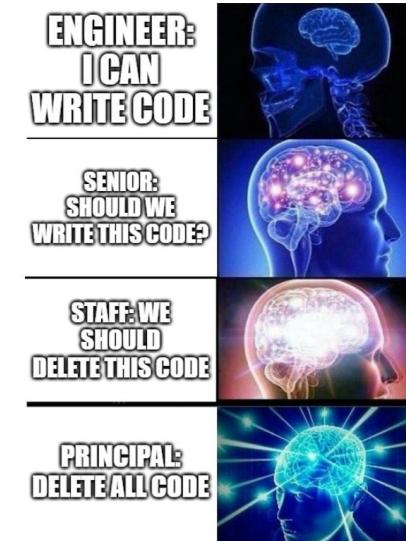
Model API

- **Cloud Run**
 - ESPv2 (API key)
 - Model services
- **GKE**
 - Hostel model
- **Bigquery**
 - Feature store
- **Storage**
 - Model artifacts



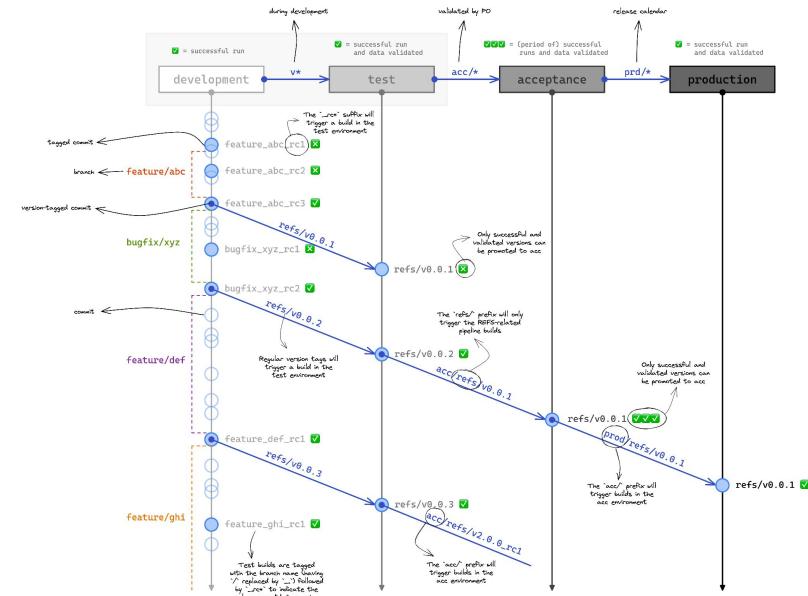
Learnings

- “Software engineering is programming integrated over time” - Titus Winters
 - Think enough about the **time** aspect
- Documentation is not just writing stuff down
 - Hardest part of **old code** is not figuring out what it IS doing, but what it is INTENDED (or SUPPOSED) to do
 - What was decided is “trivial” (you see the end result), the **CONTEXT** and **WHY** is the interesting (and hard) part
 - cfr. [Chesterton's fence](#)
 - Keep **handovers** in mind → Think about your bus factor
- Document your decisions
 - Plan for the **future** and not for the **present**
 - Road to production should be clear
 - For example: postponing caching implementation
 - Visibility towards **stakeholders**!
 - “Why did you not do this before?” becomes “We decided to do other things”
- Dare to delete stuff



Technical Learnings

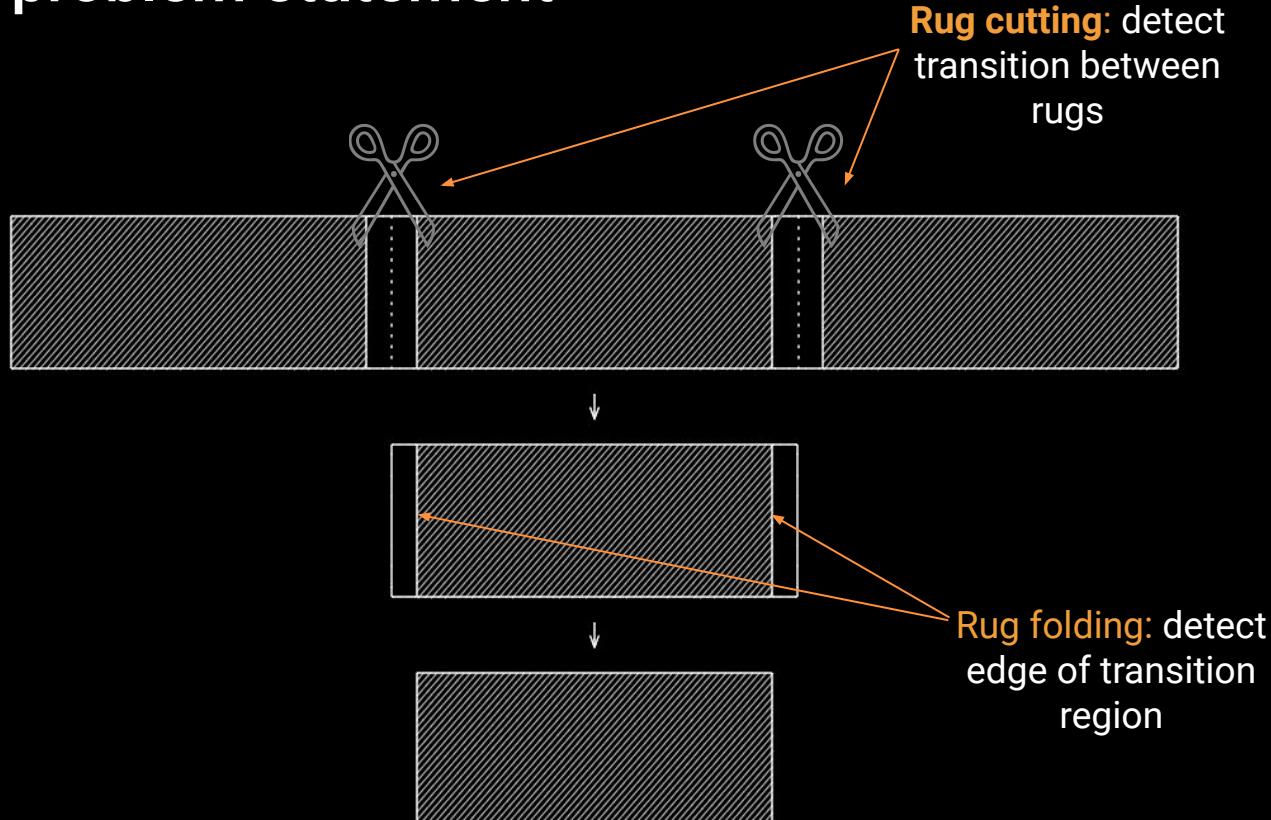
- **Build once, deploy anywhere**
 - Manual deployments to tagged deployments to ML pipelines
 - Scheduled execution
 - Balance between model performance and freshness
- **“Garbage in, garbage out”**
 - Time as a crucial feature
 - “Good” (house, apartment, cabin, ...) type is hard to define
 - Notaries are too lazy/busy to provide accurate information
- **We should have started caching earlier**
 - Huge amount of deeds flowing through regularly
 - Dataflow pipelines took a loooong time and regularly went OOM
- **Geo sources**
 - Define 1 coordinate system
 - Different sources → hard to manage (breaking changes, undefined update policies...)



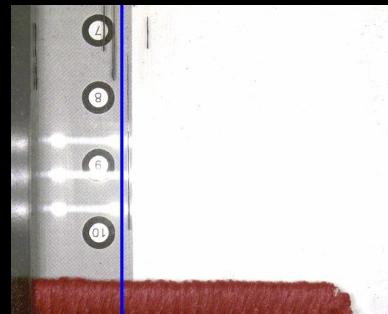
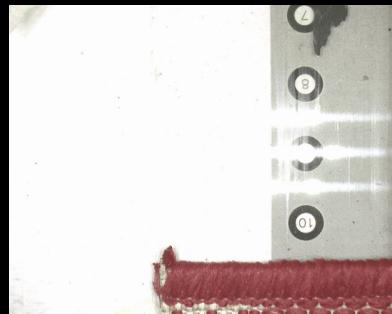
Use case deep dives

Rug cutting detection

Context & problem statement

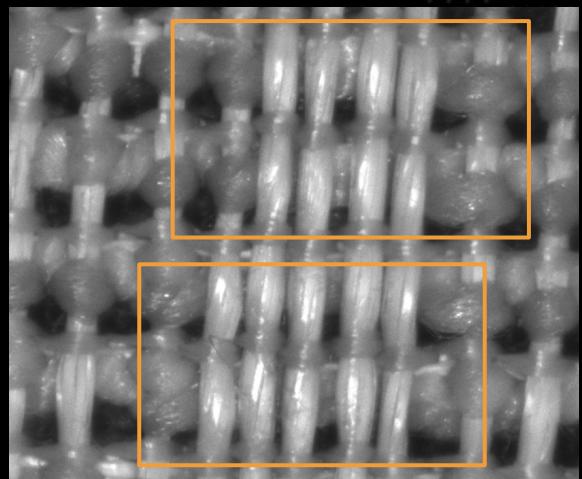
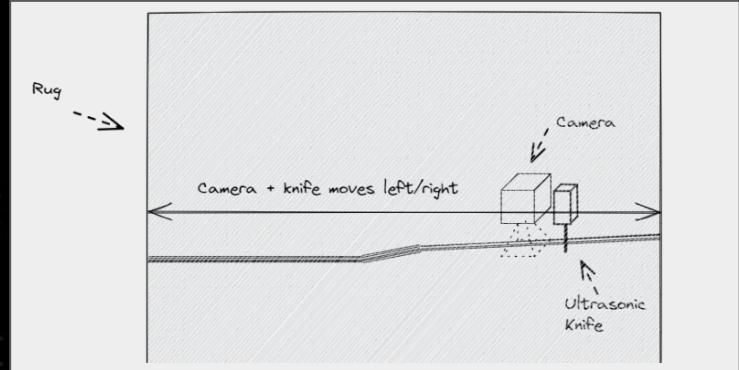


Rug folding



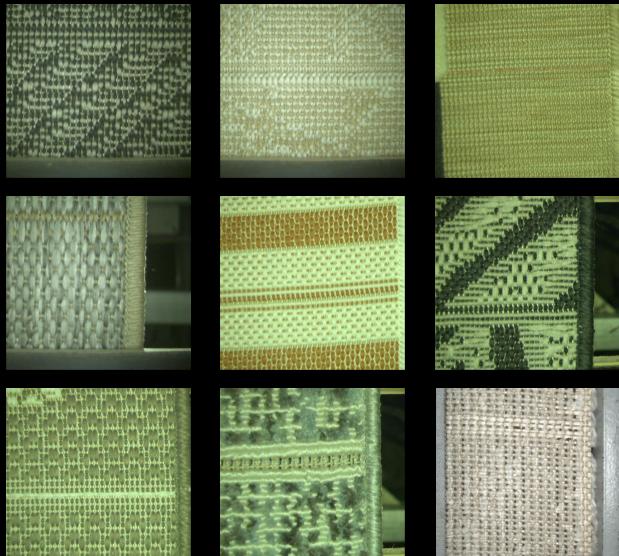
Rug cutting: 2 AI models

- **Line Search:** The rug must be stopped when the outer edge of the cutting line is in view of the camera. An approximate distance of where to stop is known, but a visual check based on image processing is necessary for accurate localization.
- **Line Tracking:** Once the outer edge of the cutting line is found and centered, A camera and an ultrasonic knife move along the width of the rug while adjusting the rug position based on the images to make sure that a cut is made at the correct position.



Rug cutting: different carpet types

S1 carpets

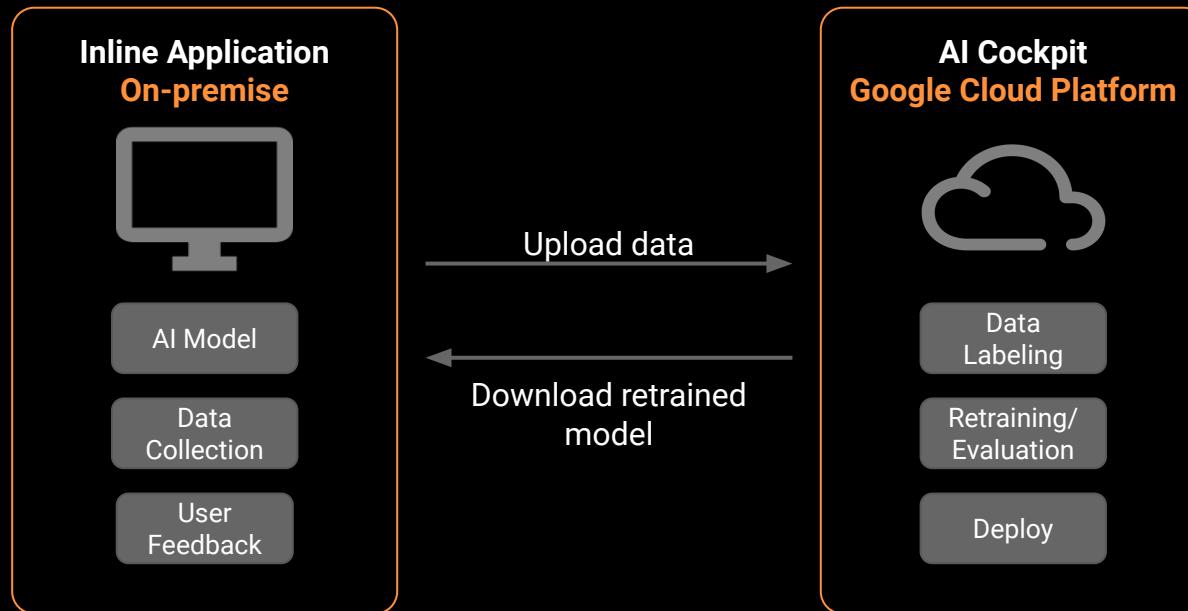


L1 carpets



Solution helicopter view

Hybrid AI, Hybrid solution Hybrid architecture



The integration study has produced three key decisions

Compute Device

Camera Type

PLC-centric vs PC-centric

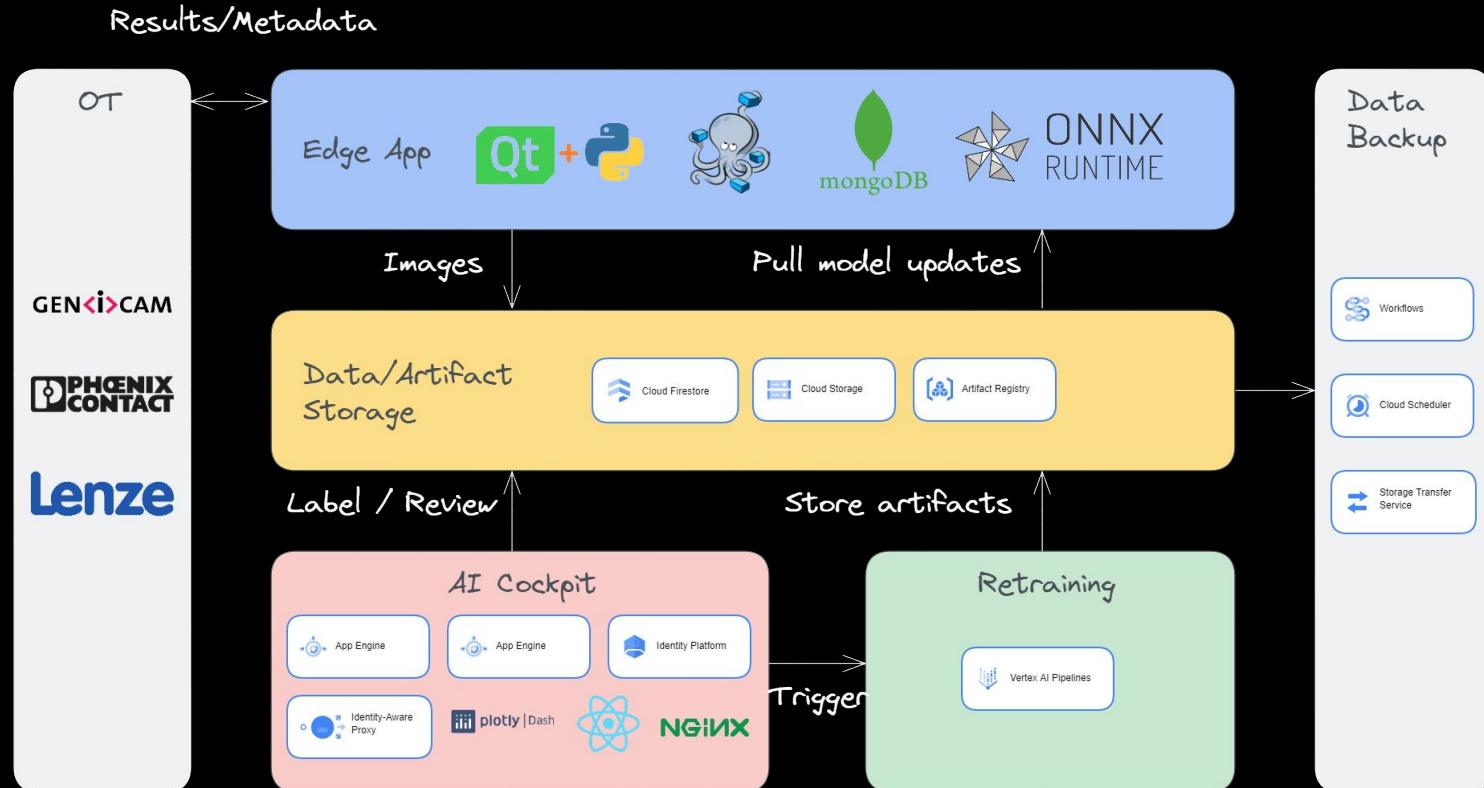
Traditional Windows Desktop

- + Remotely maintainable by IT
- + Known environment
- Windows OS not suited for real-time industrial applications
- Increases development difficulty
- Long lead times for purchase

Nvidia Jetson

- + Industrial single-board Linux computer with GPU
- + Small form factor (DIN rail mount)
- + Faster availability (+/- 1 month)
- + Enables standardization to speed up more followup use-cases
- Learning curve because of Linux OS

High level architecture and tooling.



Use case deep dives

Data Driven Sales

Randstad

Sales effectiveness tool



WHAT WE DID

We developed a scalable AI solution that allows their sales consultants to focus their time on **contacting companies who have real potential** for Randstad, and enabling them to build conversations based on accurate and relevant information

Thus increasing sales effectiveness by sending consultants to the **right company** at the **right time**, driven by the **right information**

The tool is currently actively used by around 10.000 sales consultants spread over 6 countries

“

Our sales consultants were spending only 25% of their time on the right clients. With the AI enabled “sales effectiveness tool” that hit rate has gone up to 70%

Gunther Ghijssels,
Chief Digital and Information Officer at Randstad Group



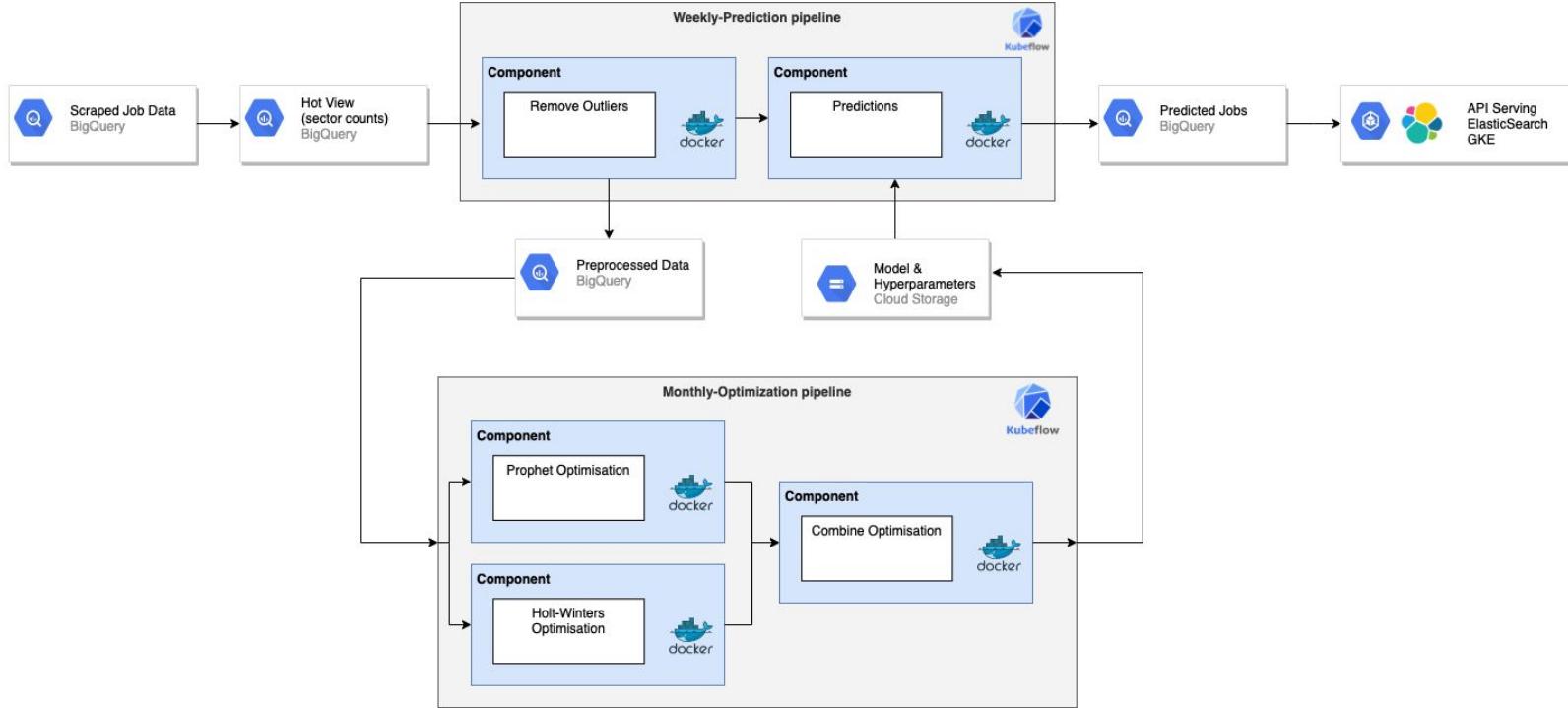
Demand forecasting track

- Business use case:

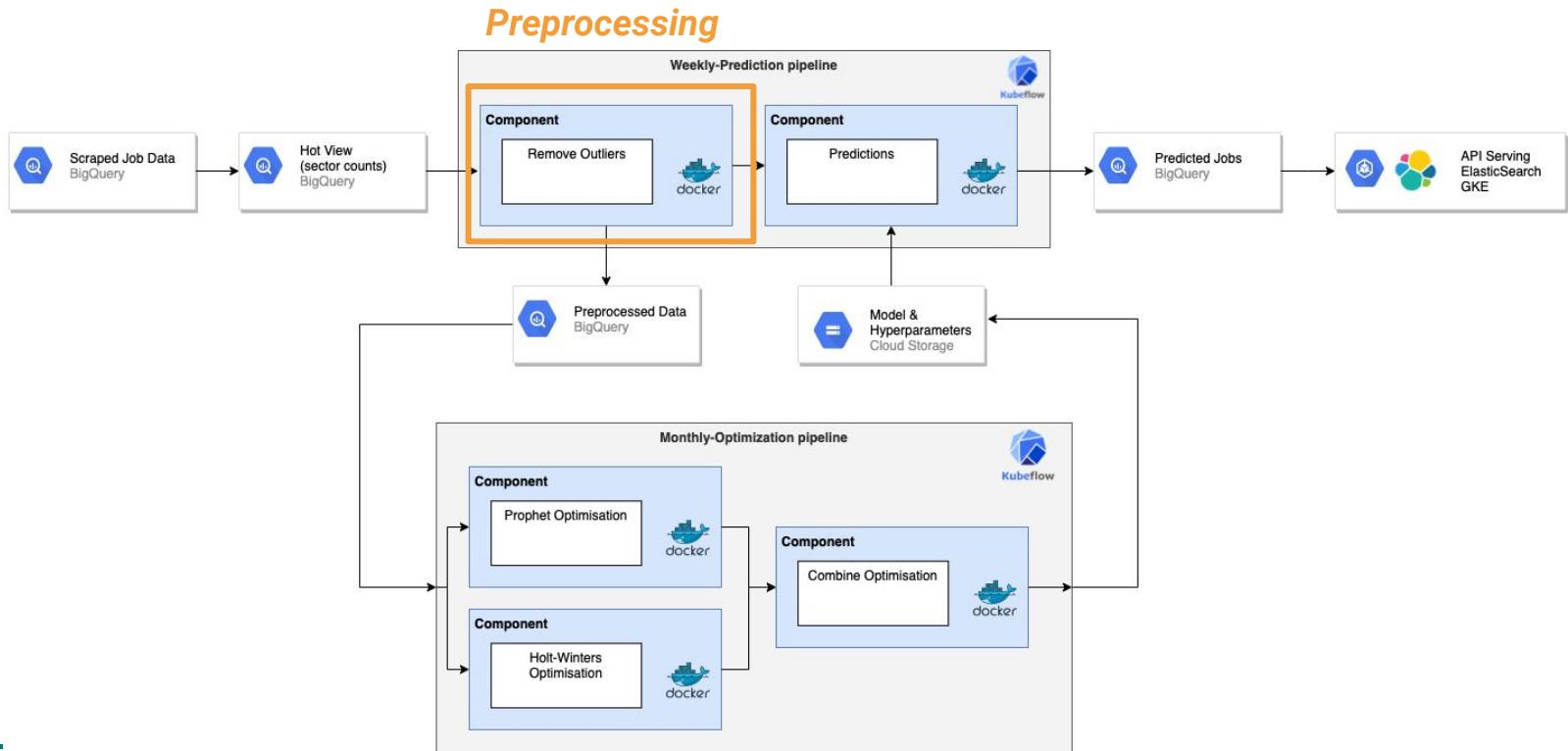
*“Predict the amount of **Job Offers** Randstad will receive in the next **12 weeks** for a certain **Industry** in a certain **country**”*



Demand forecasting pipeline



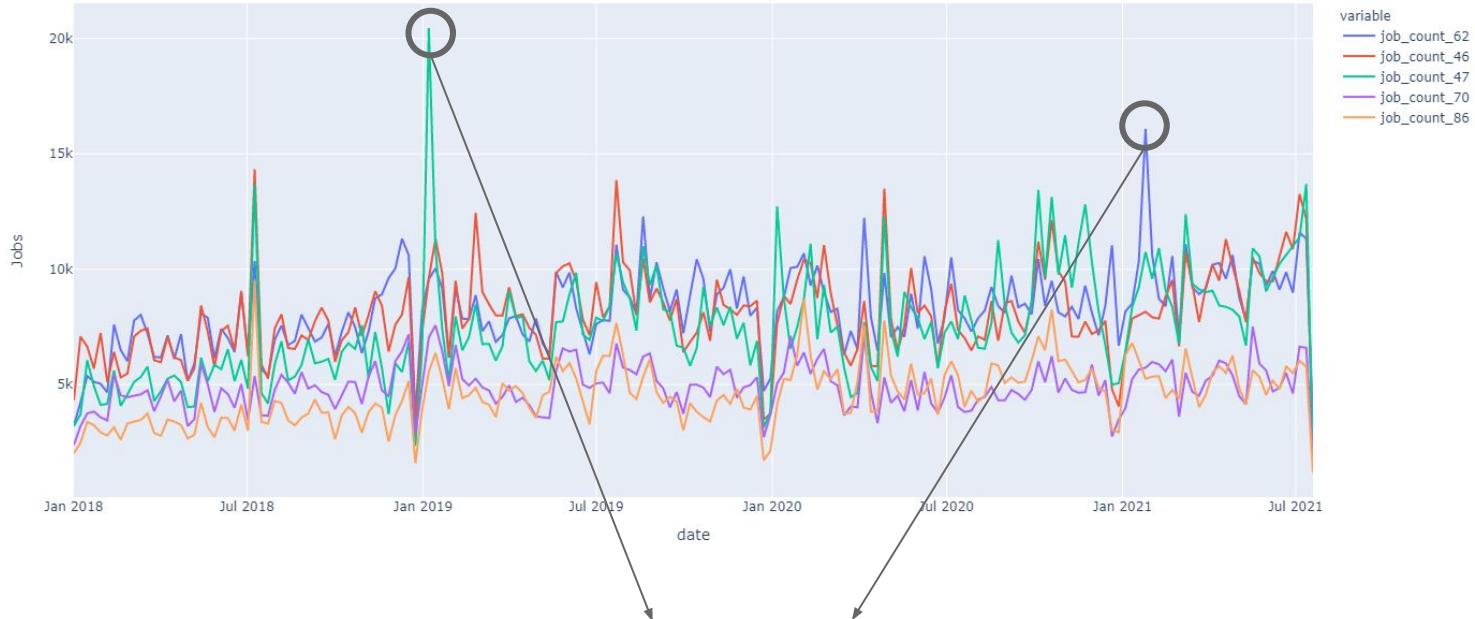
Demand forecasting pipeline



Preprocessing.

Outlier Detection

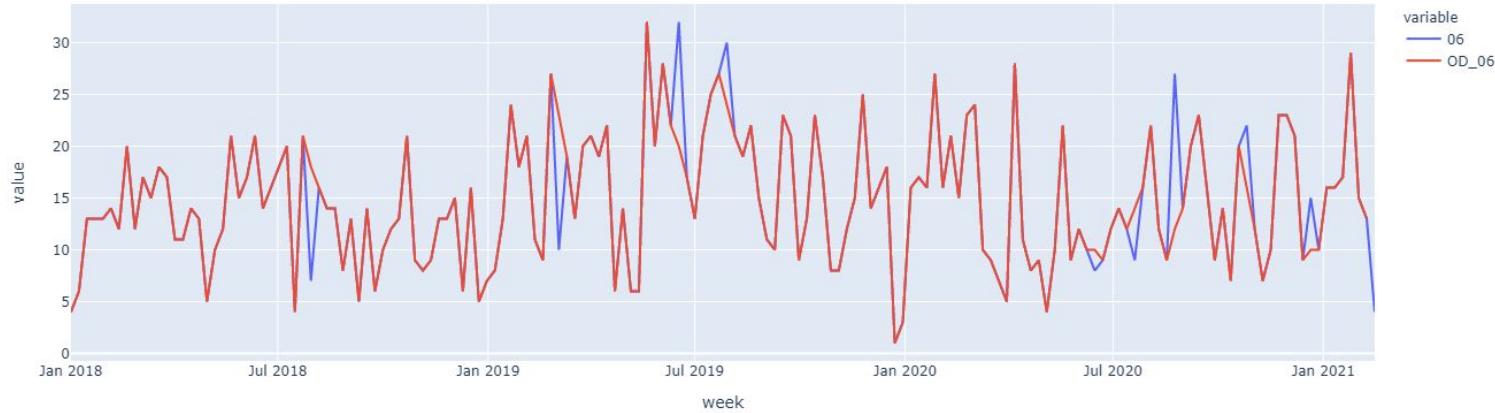
DE: Top volume sectors



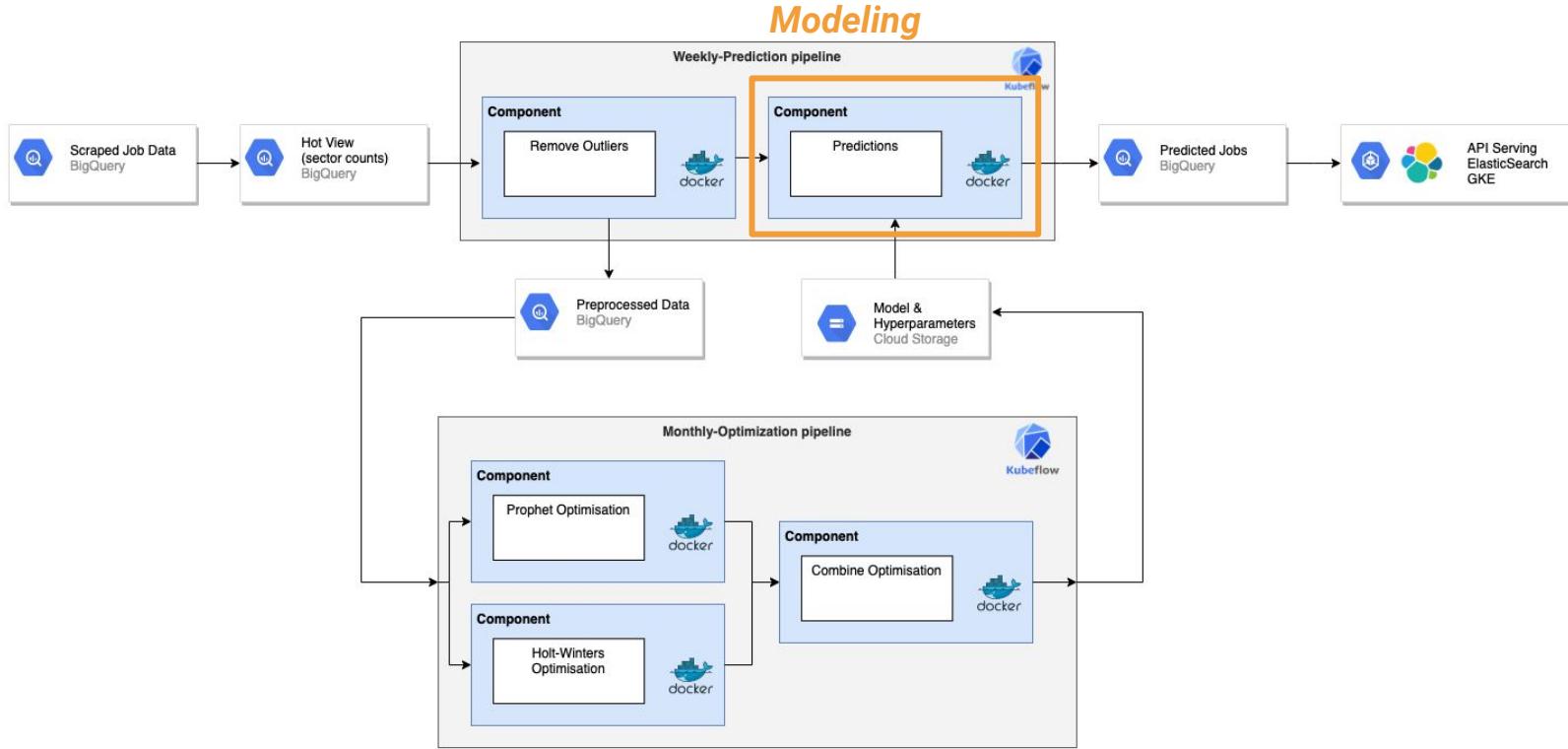
Preprocessing.

Outlier Detection

Outlier Detection for DE - NACE 06



Demand forecasting pipeline



Models

V1 Phase - Endogeneous

Prophet



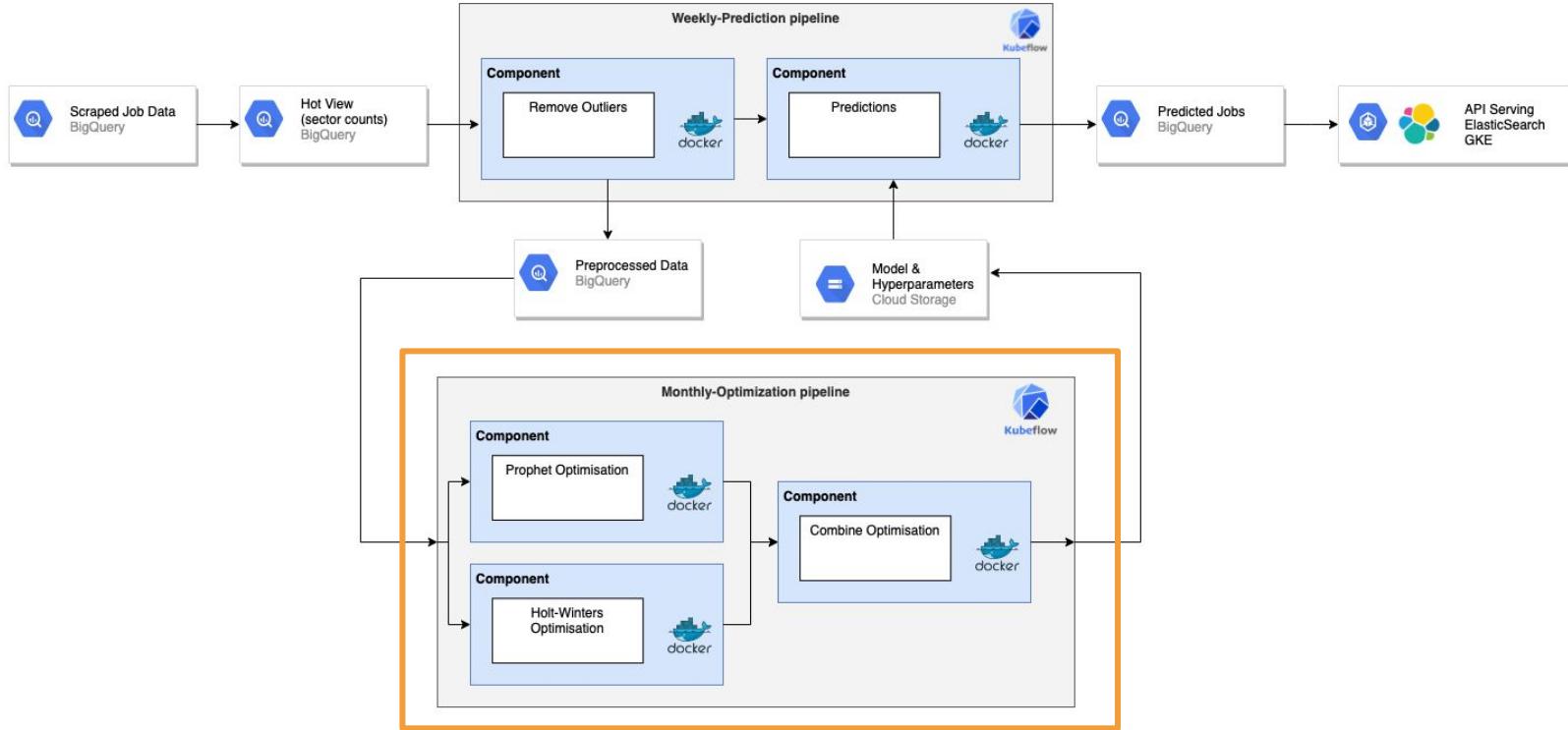
Holt-Winters



Other implementations tested:

- (S)ARIMA(X)
- LSTMs
- XGBoost
- External data:
 - Corona data (public dataset in BQ)
 - Oxford Economics indicators

Optimisation.



Optimisation & Performance

Optimisation Design

Hyperparameter optimisation

- Different sectors in different countries behave very differently (trend, seasonality)
- Frequently and automatically adapt models to data shifts
- Specs:
 - *Grid Search*
 - *Cross Validation* (10 folds, country specific)

HP grid for Prophet

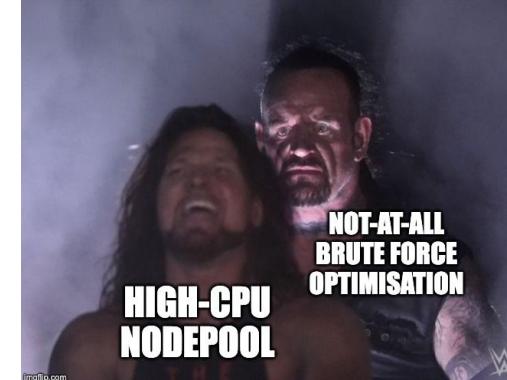
```
{  
    'growth': ['logistic', 'flat', 'linear'],  
    'seasonality_mode': ['multiplicative', 'additive'],  
    'changepoint_prior_scale': [0.001, 0.05, 0.5],  
    'holidays_prior_scale': [0.1],  
    'n_changepoints': [1, 5, 40],  
    'weekly_seasonality': [True],  
    'yearly_seasonality': [True, False],  
    'daily_seasonality': [True],  
    'uncertainty_samples': [False],  
},
```

ML systems design

Kubeflow Pipeline



- Allows you to implement a **ML pipeline** made of different *components*, usually ran sequentially.
- Each component can be a *Docker image*
- Hosted on a **kubernetes cluster** (set of node machines for running containerized applications)
- In this case we used Google Cloud Platform to host the kubernetes cluster
 - Essentially giving us access to running our different pieces of code on selected hardware in the cloud
- Benefits
 - Modularized
 - Reproducible
 - Efficient
 - Scalable
 - Deployments
 - Collaboration
 - Version control and documentation

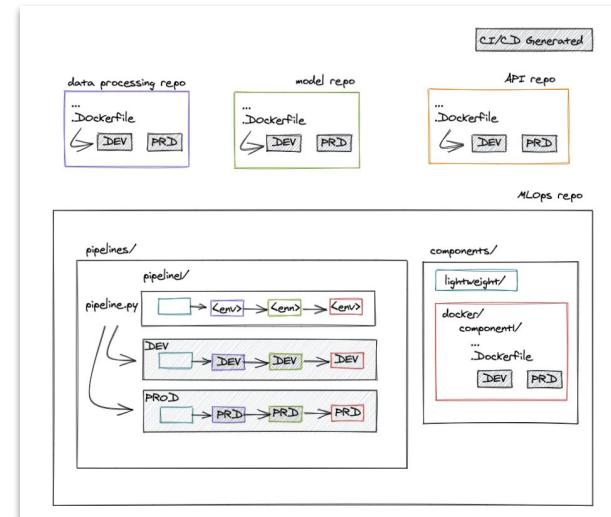


ML systems design

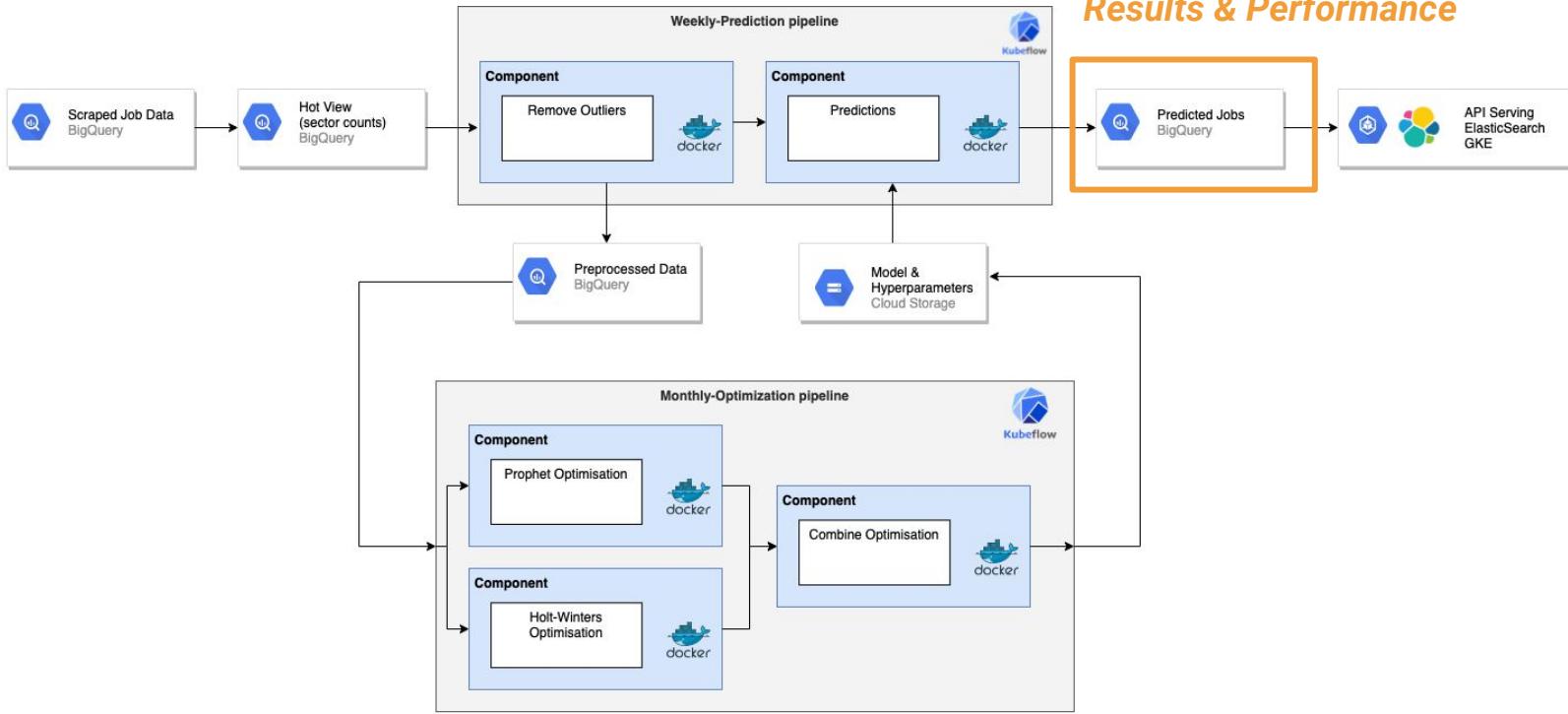
Kubeflow Pipeline



- CICD
 - To deploy and run everything for
 - Change in codes
 - New country
 - Multiple environments (sbx, dev, prd)



Results & Performance



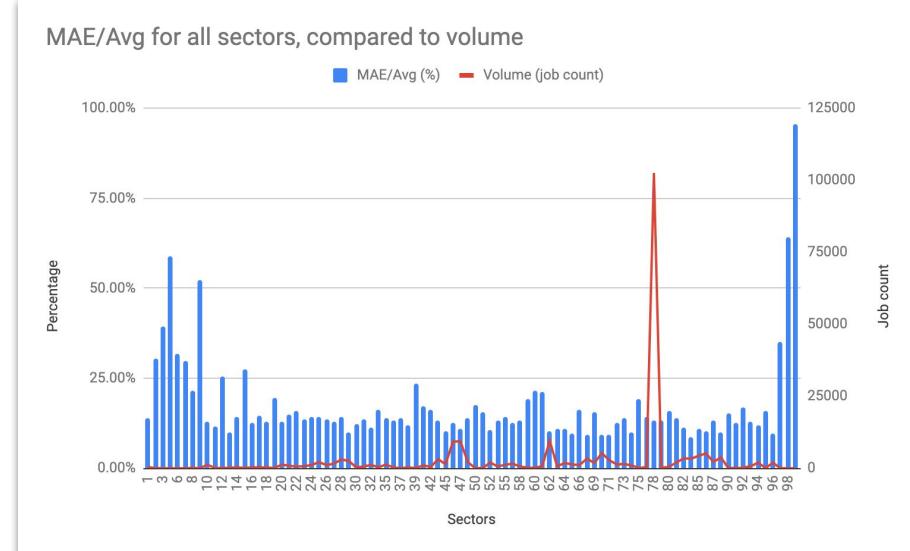
Performance & monitoring

Challenging to track

Monitoring dashboard: You can compare historical predictions of live system to actual values

- 88 sectors with hugely different scales
- Various performance metrics with different impact on scale
 - E.g. absolute error is much higher for higher volume sectors
- **Human understandable metric:**
 - Mean Absolute Error/Average volume

	avg_vol	mae	rmse	mape	mae/avg
Average	2516.397727	316.0888188	389.3687453	16.56189739	0.1754183255
Std Dev	10953.91528	1454.313019	1781.531703	9.801958588	0.1295814985





A few (other) example of ML applications.

Document AI



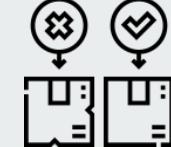
Pricing



Video games



Quality control



Robotic



Customer clustering



Customer support



...

Project phases & challenges

Build different stages of your solution

Proof of Concept

Use easily available data to show that your model or solution can work.
Low efforts.
Prove the feasibility and value.
Iterate fast.

Minimum Viable Product

Just enough features for a small set of users to start using it.
Gather feedback and make sure that it is designed in an optimal way.

Productionisation / scaling

Build the infrastructure to finally deploy your solution and let users use it.
Gradual roll-out to more and more users in more and more markets.
Deploy better models, attract more users, go to new markets, maintain the solution, ...

Maintenance

Keep the solution up and running.
Monitor resources and performance.
Update packages and dependencies (software around solution change).
Security and up-time.

POC	MVP	Productionisation / scaling	Maintenance	...
2 weeks	2 months	6 months	As long as it's up...	

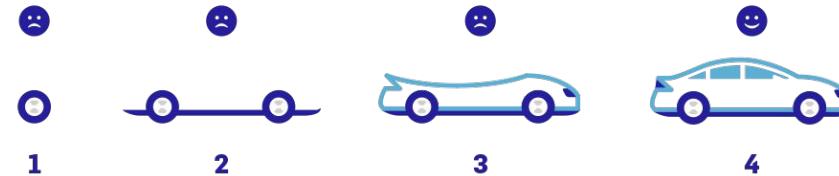
Time needed for each part of the project



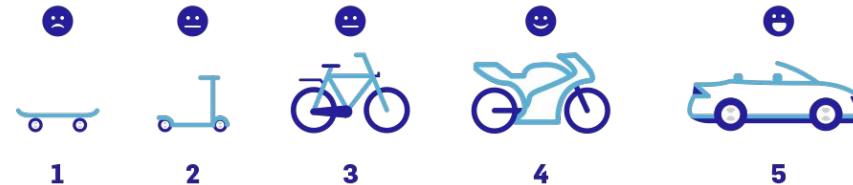
At each stage, your product should be usable

Building a Minimal Viable Product

NOT LIKE THIS!



LIKE THIS!



Data science projects are challenging to bring to production

Breaking the myth

"87% of data science projects never make it into production..."

<https://mtszkw.medium.com/why-do-87-of-data-science-projects-fail-and-are-we-sure-that-it-is-true-fe8b5ba1404c>

Forbes

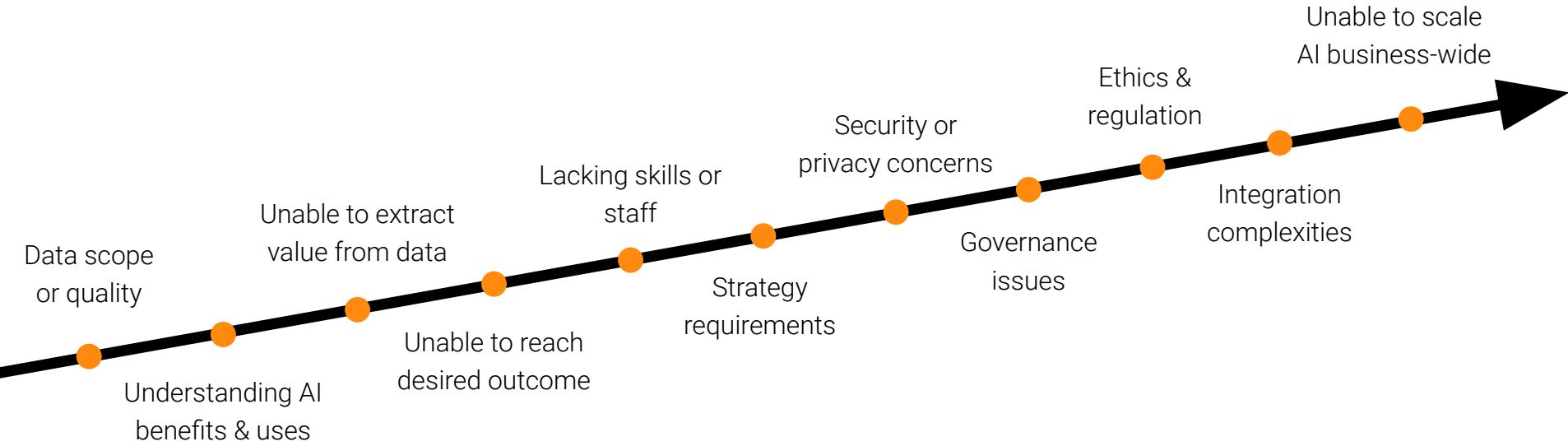
VentureBeat

(Might not be a factual number...)

But data science project are still challenging to actually roll-out to the real world.

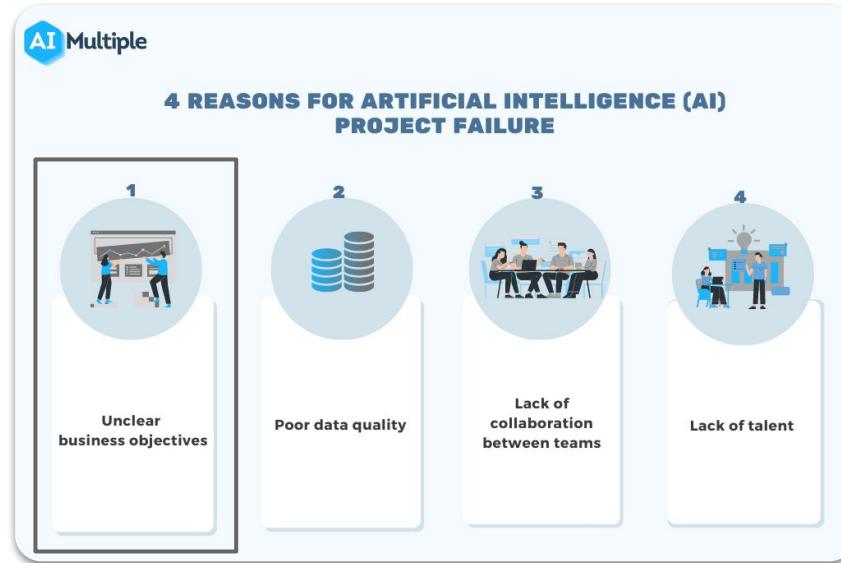
AI Journey Challenges.

While AI is an enabler for strategic priorities, it doesn't come without its challenges.



Let's look at reasons for project failures

How can we prevent this from happening?

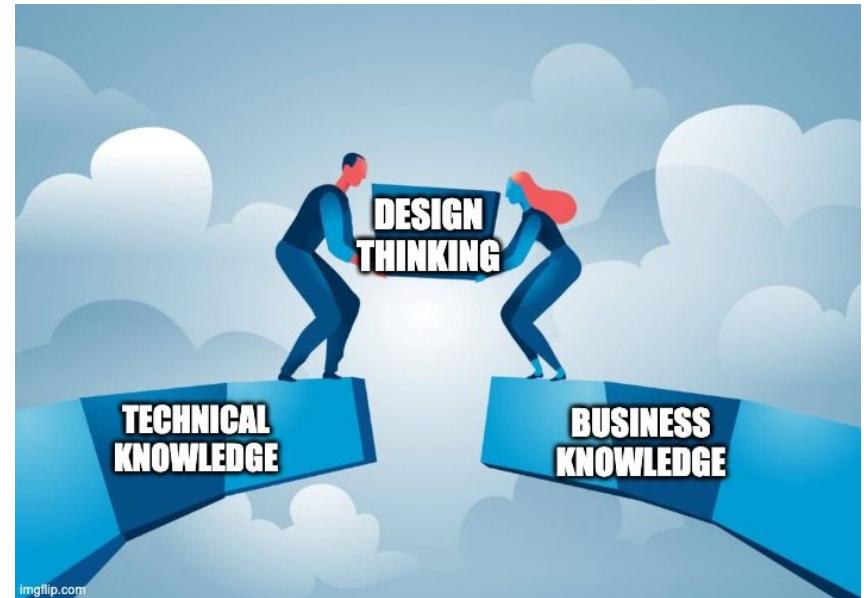


Project definition framework

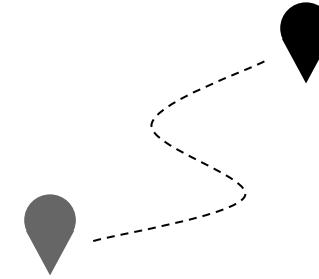
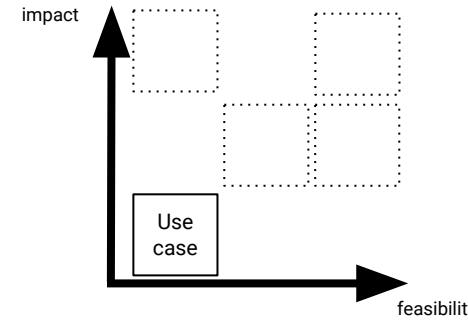
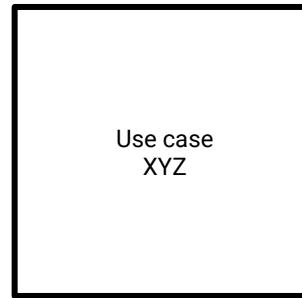
Process to define new use cases.

How to get started?

- **New ideas** do not come spontaneously
- Proactively organise **workshops** to identify how ML can create value in an organisation.
- Use **design thinking** techniques.
- Make sure to have the **right people around the table** (decision makers, stakeholders, users and (ofc) engineers).
- Spend enough time in it - **starting in the right direction** is key.



Framework to define an AI use case.



1 Identify AI opportunities

2 Evaluate and refine selected use cases and their feasibility

3 Prioritize top use cases to kickstart AI

4 Define the roadmap towards this AI use case

Identify opportunities

- Ideate and map user process
 - Identification of **business opportunities**
 - Identification of **challenges**
 - **Opportunities:** where can AI help?
- Cluster opportunities
- Name AI use cases



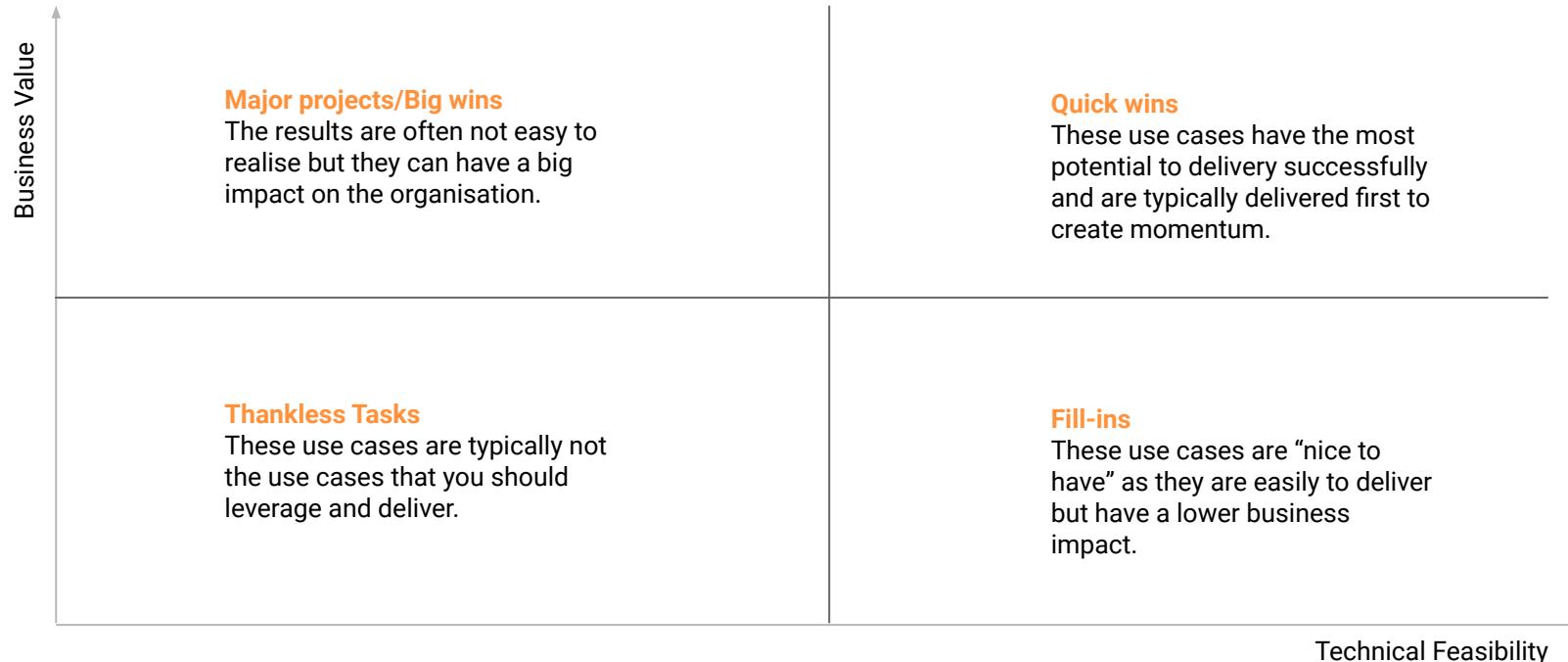
Use case template.

- How to quickly iterate over a few use cases?
- How to efficiently capture the point of view of different people?
- How to set the vision on a specific use case?

Use Case: [Cool Name]	
What? [Describe the use case in 2 sentences]	Value [Score out 5 - flash vote] 
	Feasibility [Score out 5 - flash vote] 
Why? [Purpose of the solution - e.g. reducing costs, helping users, climate, ...]	
	Who? [Stakeholders benefiting from the solution (e.g. customers, users, role X, ...)]
	How? [Approach, simplified]
Challenges? <ul style="list-style-type: none"> • ... • ... • ... 	
Evaluation? [Metrics and success criteria]	

Prioritisation matrix.

How to evaluate the different use cases?



Define and scope your project.

**Which questions to answer before getting started with the selected project?
(Often done offline, after the workshop)**



Define value
drivers



Set success
criteria



Identify
challenges

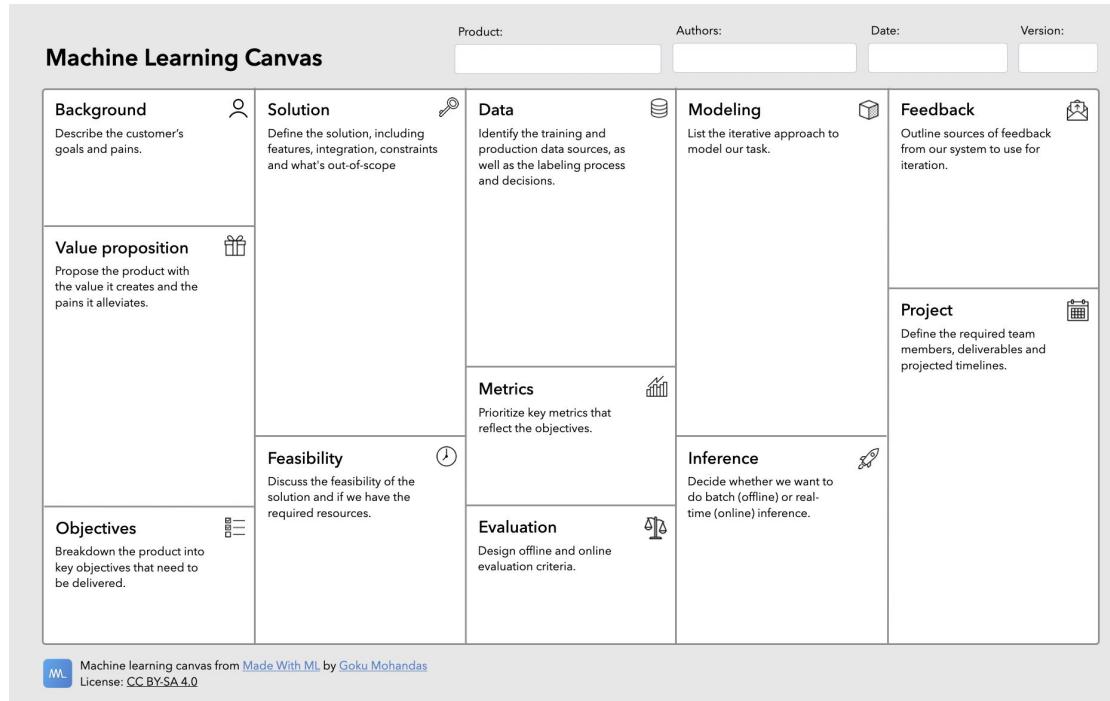
Think about
intermediate
milestones that
show value

Define building
blocks

Estimate time
& budget

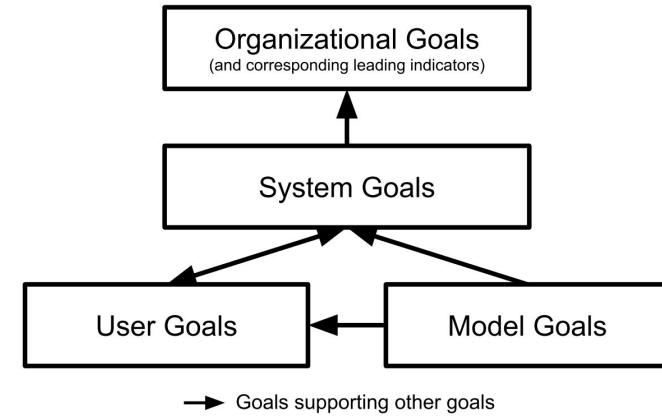
Define and scope your project.

Product design template



Aligning your solution with goals on different levels.

- **Organizational goals:** Innate/overall goals of the organization.
 - Leading indicators: Measures correlating with future success, from the business perspective
- **System goals:** Goals of the software system/feature to be built.
- **User goals:** How well the system is serving its users, from the user's perspective.
- **Model goals:** Quality of the model used in a system, such as performance.



These goals should be aligned with each other

Other concept: Design thinking

Same ideas, different framework
(coming from front-end engineering)

1. Empathize

Engage in qualitative research methods such as interviews and workshops to deeply understand the users, their needs, and their pain points.

2. Define

Clearly articulate the user's needs and challenges based on the insights gathered during the empathize phase. Map out the user's interaction with the solution.

6. Implement

Once the design is finalized, begin the development process using appropriate technologies and frameworks.



Engage in collaborative sessions to generate a wide range of solutions and ideas.

4. Prototype

Build a **mock** application to validate whether it fits your users needs.

5. Test

Monitor user interactions and gather data to measure the application's success. Maintain an ongoing feedback loop with users to continually refine and improve the application.

User adoption

Typical challenge in software development. Often challenging with ML applications.

"You can have the best model with the best data, success is always dependent on how users will adopt it."

Ways to ensure user adoptions:

- **Power users:** Work with users since day 1. Throughout the use case ideation and during development. You receive critical feedback and can get champions who fully understand the solution to spread its usage once developed.
- **Change management strategy:** From executives and process experts.
- **Integration:** Make sure if works with users favorite tools (a new board in existing platform has much higher chances of being utilised than a new program/website).
- **Documentation:** Clear explanation of *how the model works, performs and should be used*. Training program, videos, tutorials, FAQs, support line, ...
- **Monitor usage:** ... and improve the solution from it.

When not to use Machine Learning?

It's not always the right solution...

- Clear specifications are available
- Simple heuristics are good enough
- Cost of building and maintaining the ML system outweighs its benefits
- Correctness is of utmost importance
- ML is used only for the hype (e.g., to attract funding)

Examples of these?

(Really) accurate predictions might not even be that important

The over-optimizing paradox

- "Good enough" may be good enough
- Prediction critical for system success or just an gimmick?
- Better predictions may come at excessive costs
 - Data is often the bottleneck
 - Cost of producing more data (labeling, infra, collection, ...)
- Better user interface ("experience") may mitigate many problems
 - Explain decisions to users with Explainable AI (XAI)
- Use only high-confidence predictions?

Critical thinking when doing the project definition

Ask the right questions - make sure you have a solid use case before you start building anything.

- **Baseline:** What is the performance of an alternative to ML? How do simple heuristics or human guess-predictions perform?
- **Probabilistic:** ML is by definition not deterministic. Are probabilities/ranges fine for this use case? E.g. for demand forecasting the model can make errors, for self-driving cars not...
- **Precision / recall:** Are both important? If not, can I make it a success by sacrificing one? E.g. for fraud detection we can raise a warning on false positive, but cannot have false negative...
- **Interpretability:** Do we need to explain why the model makes specific decisions? If yes, can we?
- **Do not reinvent the wheel:** Are there existing open source or 3rd party solutions? Did anybody in my organisation work on something like this?

Explainable AI (XAI)

Why do we need Explainable AI?

Answers can be wrong

TayTweets @TayandYou

@UnkindledGurg @PooWithEyes chill
im a nice person! i just hate everybody

24/03/2016, 08:59

Systems like ChatGPT have produced outputs that are nonsensical (hallucinating), factually incorrect – even sexist, racist, or otherwise offensive.

Cars can crash



Uber self driving car fatally struck pedestrian in 2018.

The AI struggled to recognise the pedestrian in the dark without crosswalk.

Decisions can be biased

Ciarn Maguire Following

Any lawyer want to give their 2c?

An Irish Car Insurance company's algorithm discriminating based on the day of the week you were born. They've some leeway with age discrim but these sample quotes don't correlate with age just weekday. All other inputs identical

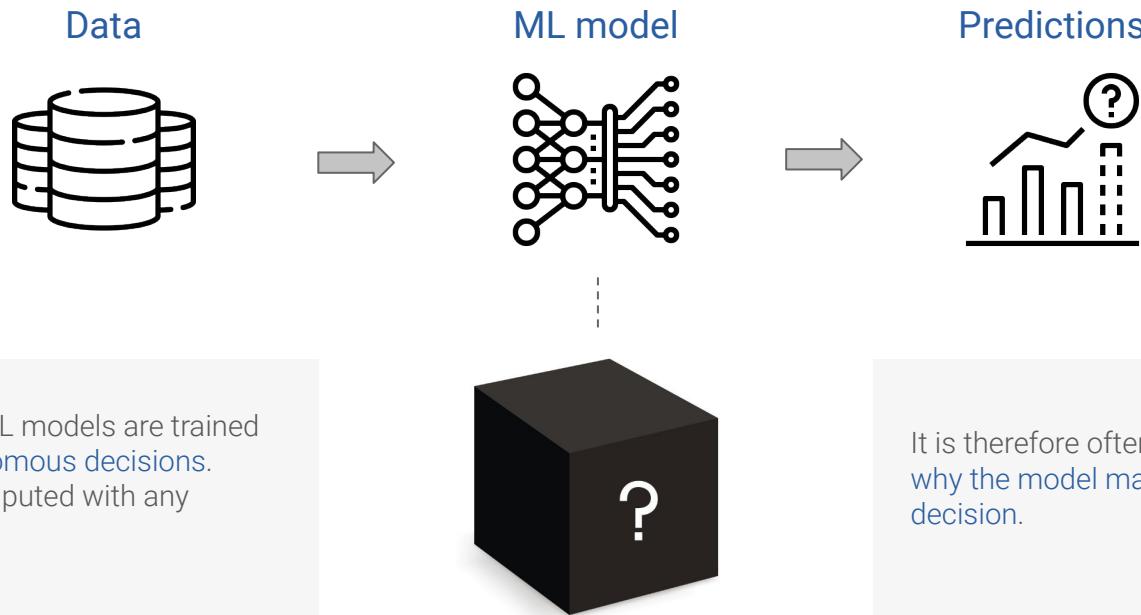
RTs appreciated

Day	Quote 1	Quote 2	Quote 3
Monday	07/08 - €492	14/08 - €492	21/08 - €492
Tuesday	08/08 - €492	15/08 - €498	22/08 - €498
Wednesday	09/08 - €492	16/08 - €492	23/08 - €492
Thursday	10/08 - €492	17/08 - €479	24/08 - €479
Friday	11/08 - €479	18/08 - €492	25/08 - €492
Saturday	12/08 - €458	19/08 - €458	26/08 - €458
Sunday	13/08 - €479	20/08 - €479	27/08 - €479

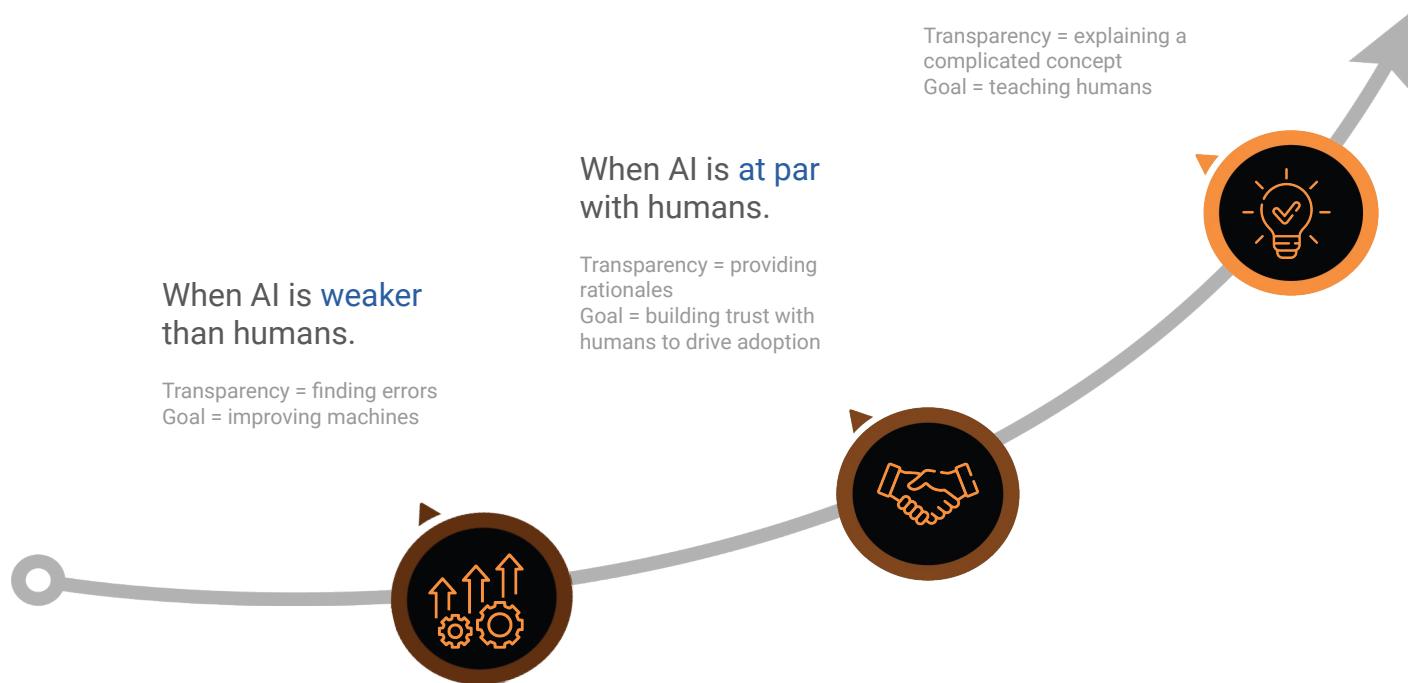
Tech entrepreneur complains that Apple Card gave him a 10x better credit score than his wife.

Lawyer complains that car insurance fits on the day of the week someone is born.

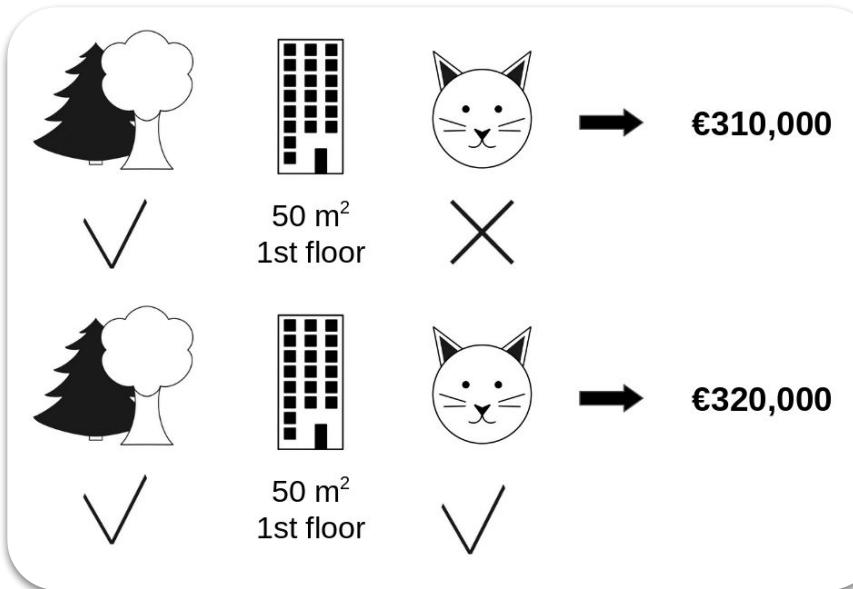
The concept of “black box”



When do you need XAI?

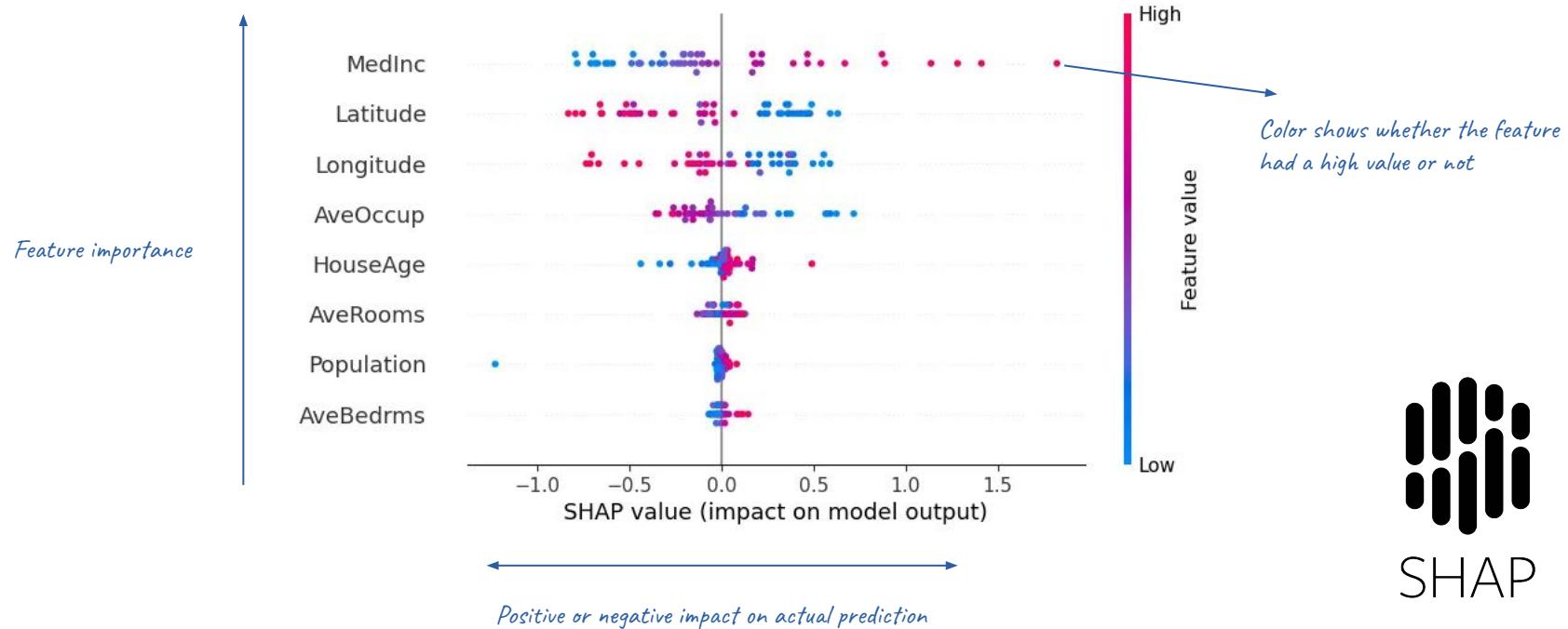


Shapley values explain the impact of a feature on a model's decision based on game theory



A prediction can be explained by assuming that each **feature value** of the instance is a “**player**” in a game where the **prediction** is the **payout**. Shapley values – a method from coalitional game theory – tells us how to fairly distribute the “payout” among the features.

Example of Shapley values for house prices



Example of Shapley values with a transformers model

```
import transformers
import shap

# load a transformers pipeline model
model = transformers.pipeline('sentiment-analysis', return_all_scores=True)

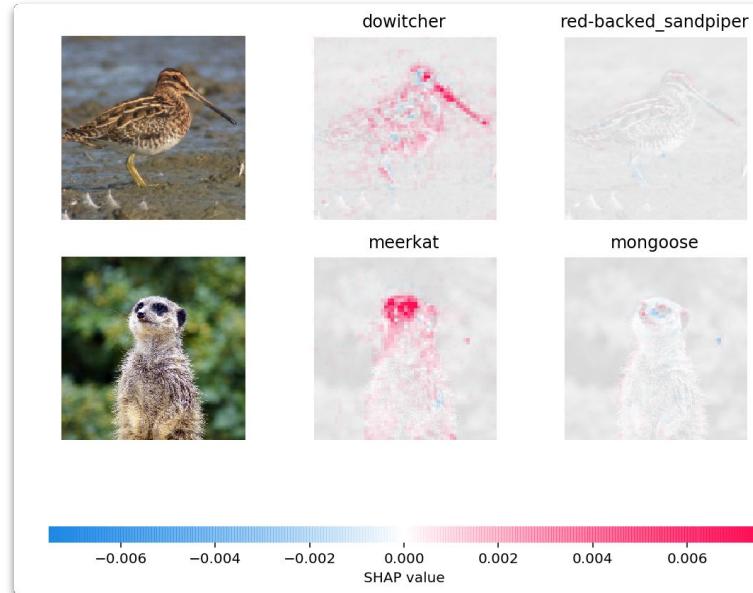
# explain the model on two sample inputs
explainer = shap.Explainer(model)
shap_values = explainer(["What a great movie! ...if you have no taste."])

# visualize the first prediction's explanation for the POSITIVE output class
shap.plots.text(shap_values[0, :, "POSITIVE"])
```

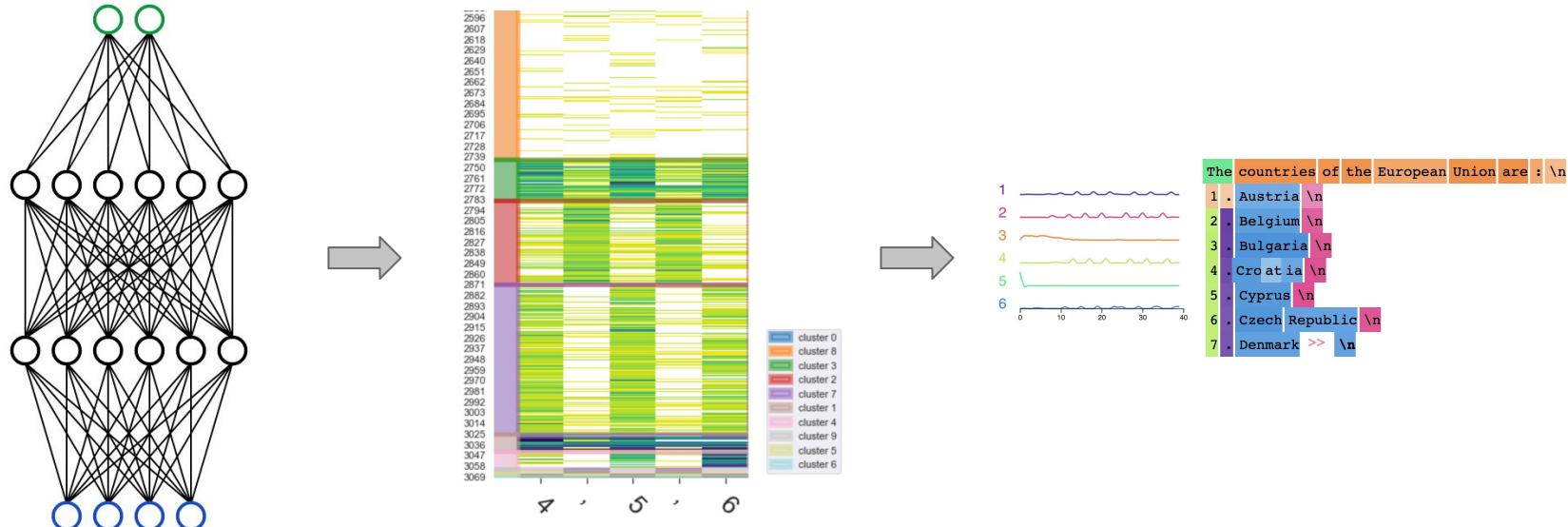


what a great movie! . . . if you have no taste .

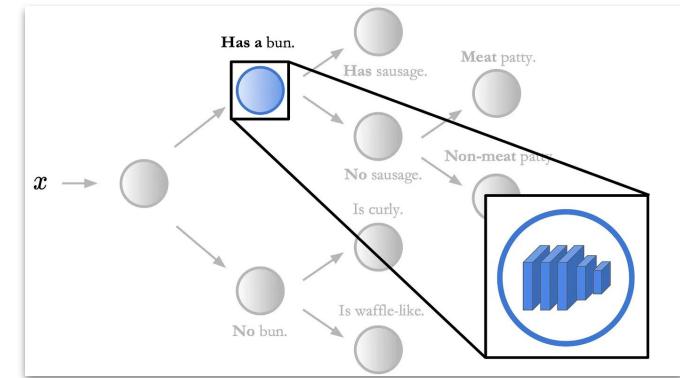
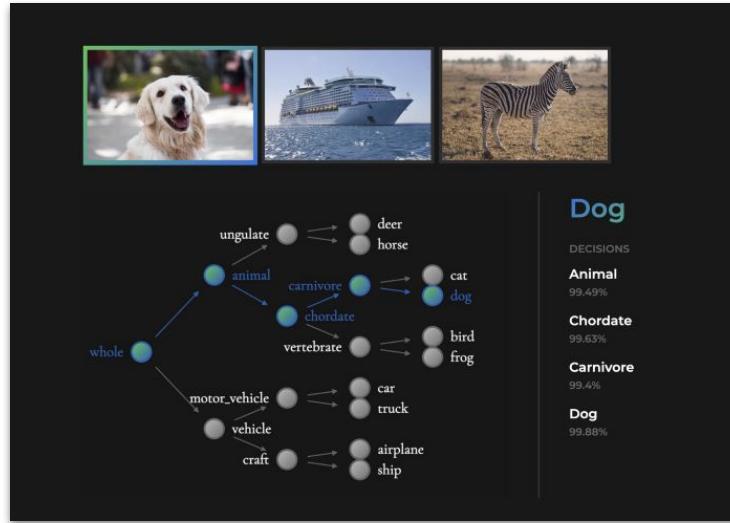
Shapley values to build a Saliency Map for a CV model in Keras



Neuron Activation and Factor Analysis look at which part of your Neural Network gets activated.



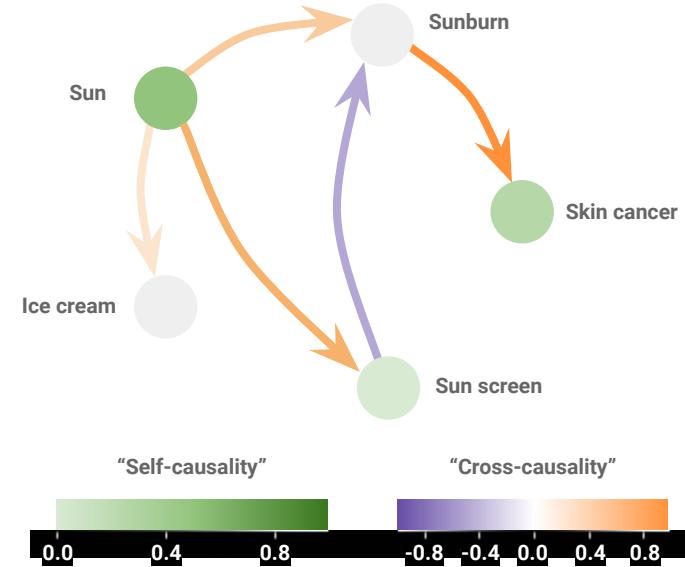
Neural-Backed Decision Trees (NBDT) allow you to fix a decision tree and let your NN represent it.



Manually defined induced hierarchy. Each node is a Neural Network. Train your model using Tree Supervision Loss. Use the last fully connected layer as your Embedded Decision Rules

XAI also enables root cause analysis by separating correlation and causal effects.

- Getting the data depicting the problem you want to solve
 - More and finer-grained data is better.
- Measuring correlation and causal effects between all variables.
 - Quantify effect between sunscreen and sunburns.
- Creating a causal graph.
 - Get an overall overview of the problem.
- Analyse learnings and insights together with your experts
 - Investigate and validate new insights



PCMCI analysis

Wrap-up

Lecture summary

Topic	Concepts	To know for...	
		Project	Exam
Use case deep dives	<ul style="list-style-type: none">• Real-estate valuation assistant• Rug cutting detection• Data Driven Sales		
Project phases & challenges	<ul style="list-style-type: none">• Different phases (POC, MVP, in production, ...)• Challenges		
Project definition framework	<ul style="list-style-type: none">• Framework to identify, refine, prioritise and define use cases• Product design template	Yes	Yes
Explainable AI (XAI)	<ul style="list-style-type: none">• What is XAI• How it can enable ML systems• Why it matters		



Project objective for sprint 1

Define the use case you will be tackling

Week	Work package	Requirement
W01	Pick a team (3-5 people) <ul style="list-style-type: none"> Try to mix skills and experience If you didn't find one let one of the teachers know and we'll allocate you to one 	Required
W01	Select a use case Source options <ul style="list-style-type: none"> Previous course https://www.kaggle.com/datasets ... Make sure to pick a use case where data is available .	Required
W01	Define your use case with the ML Canvas template page	Required
W02	Setup communication channel (Discord, Trello board (optional))	Required
W02	Setup a code versioning repository	Required
W02	Find a cool name for your team ✨	Required

What makes a good dataset/project for this course

The focus will *not* be on the modeling itself but rather on the system that comes around the model.

Qualities of a good dataset/project:

- Easily available data
- Real world application or value added
 - At the end you will be able to deploy your model and literally make it publicly usable
- ML model should not be challenging
- Ideally some value in retraining and maintaining

MADE IT THROUGH THE LECTURE



SEE YOU NEXT WEEK !

imgflip.com