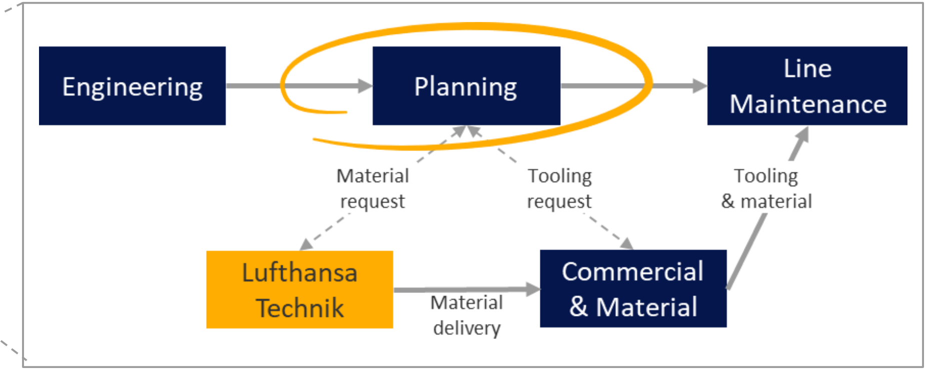
Predictive Process Management for Aircraft MRO
Final Presentation – TUM DILAB

29.07.2020

The planning and scheduling of work orders at Lufthansa CityLine GmbH (CLH) is done in the planning department and highly complex

Problem Statement (1/3)

Paula works in the **Planning Department** of CLH.
She plans all maintenance tasks for the CLH fleet.



Engineering

Planning

Line Maintenance

Material request

Tooling request

Tooling & material

Lufthansa Technik

Material delivery

Commercial & Material

Therefore, Paula needs to consider a lot of things:

- ✓ **Due date** of the maintenance tasks
- ✓ Availability of **infrastructure**, **tooling** and **aircraft**
- ✓ Lead time of **materials** for their on-time arrival
- ✓ Availability and qualification of the **line workers**

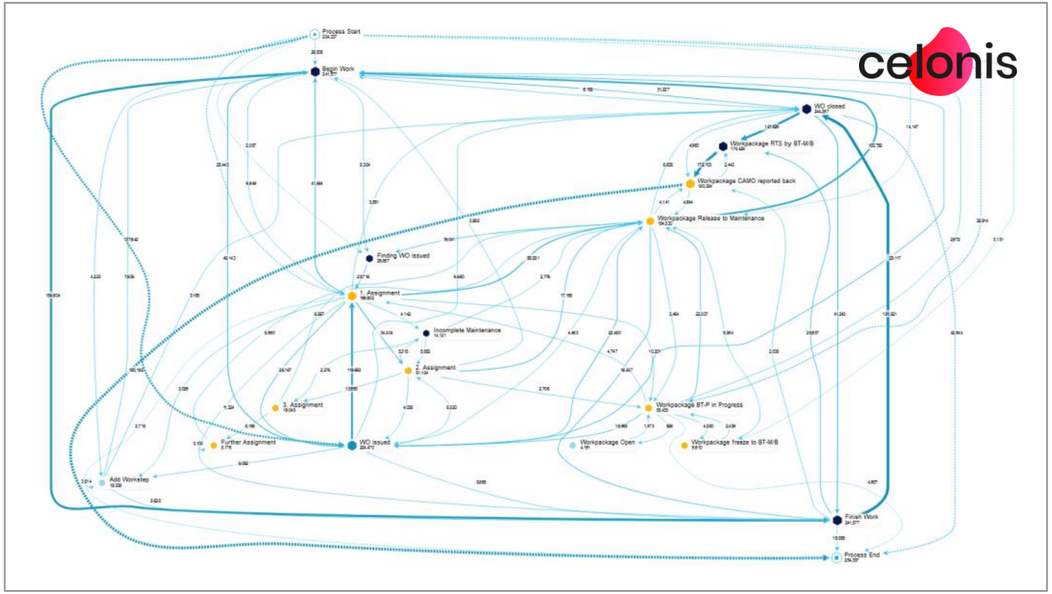
Lufthansa CityLine

Although a desired process flow is defined, the reality differs considerably from it

Problem Statement (2/3)

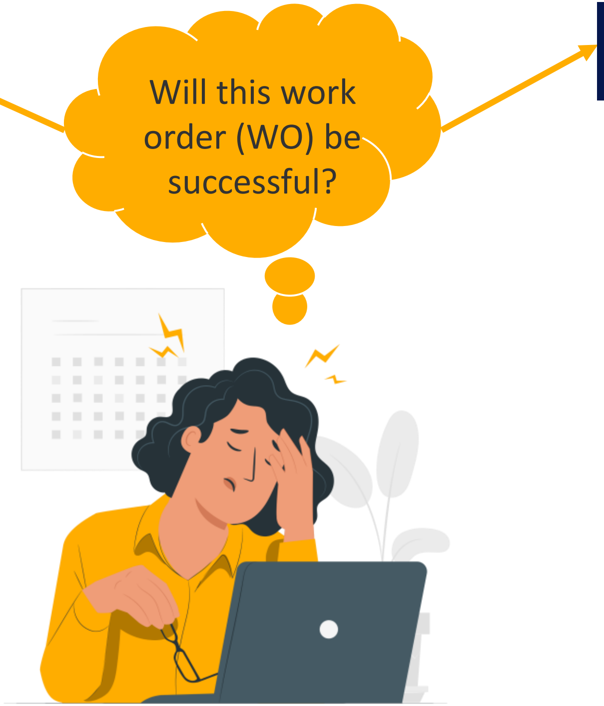
"To-be"…

… & "As-is" Maintenance Repair Overhaul Process at CLH



Paula has to deal with different challenges during her workday

Problem Statement (3/3)



Incomplete Maintenance

Work could not be conducted completely or not at all.

→ Aircraft might be grounded

→ Short term replacement needs to be found

→ Delays & cancellations occur

→ Causes Deassignments

Deassignments

She needs to find a new slot to perform this maintenance task.

→ Reworking previously assigned work orders

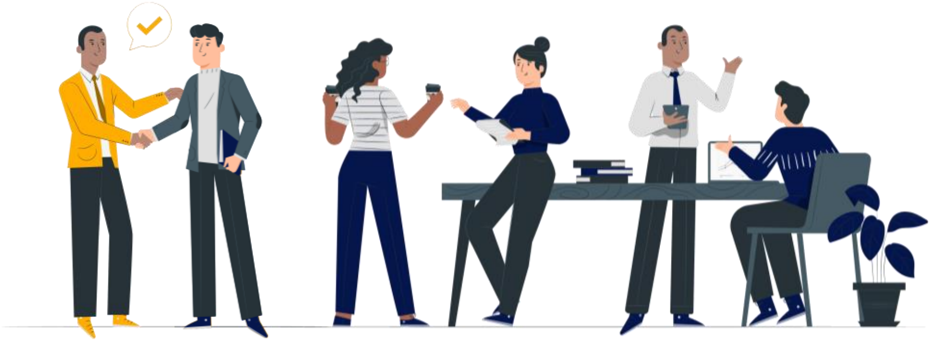
→ Waste of time that could have been spent working on more urgent matters

→ General process inefficiency

Design and implementation of a system for predictive process management to increase efficiency in the aircraft MRO process at CLH

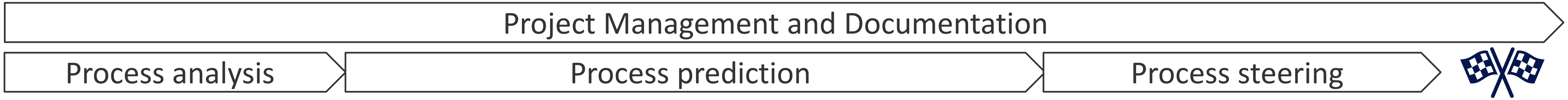
Project Plan

TUM Data Innovation Lab



Project Goal

We develop a predictive process management tool to predict and proactively steer critical cases in the aircraft MRO processes of Lufthansa CityLine until July 2020 by implementing and evaluating three different prediction models.



Project Management and Documentation

Process analysis

Process prediction

Process steering

- Defined "Project Goal"
- Business & process understanding
- First version of input dataset
- First model results
- Result interpretation & Business impact
- Final report & presentation

Agenda

1 Problem Statement

2 Implementation & Setup

- Academic Approach
- Data Exploration, Feature Engineering and Data Preparation
- ML Models
  - Logistic Regression
  - Neural Networks
  - Decision Tree
  - Relational Graph Convolutional Networks (R-GCN)

3 Results and Impact

- Comparison of Model Results
- Business Impact
- Explainability

4 Summary and Next Steps

Based on a literature review of we selected three approaches to predict process flows at CLH

Comparison and Selection of Academic Approaches

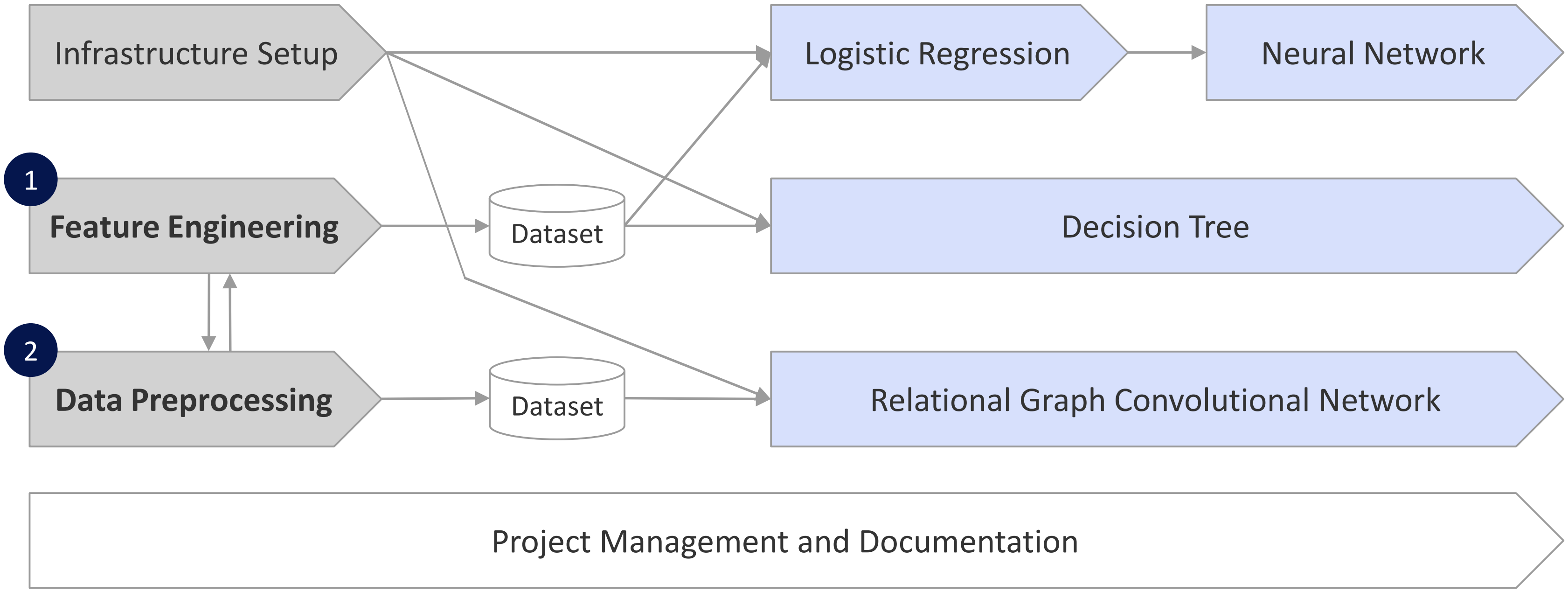
(Dis)advantages

			Feature engineering necessary?	Capturing hierarchical information?	Capturing sequential information?	Long training time?	Explain-ability?	Framework support?
Approaches	Feature Engineering	Decision Trees	Yes	No	No	Yes	Yes	Yes
		NN/Logistic Regression	Yes	No	No	No	Yes (but tricky)	Yes
		K-nearest Neighbors	Yes	No	No	No	Yes	Yes
	Relational Graph Convolutional Networks (R-GCN)		No	Yes (with limits)	No	Yes	No	Yes
	Relational Recurrent Neural Networks (R2NN)		No	Yes	Yes	Yes	No	No (but modular)

🟧 Selected approaches

In this project phase we worked on several work packages with different dependencies on each other

Work Packages and Dependencies in Process Prediction



🟦	Prediction approach
▨	Work package
→	Dependency

Lufthansa CityLine

Understanding the domain, process & data, we handcrafted the feature vector based on the data being available at the time of the assignment

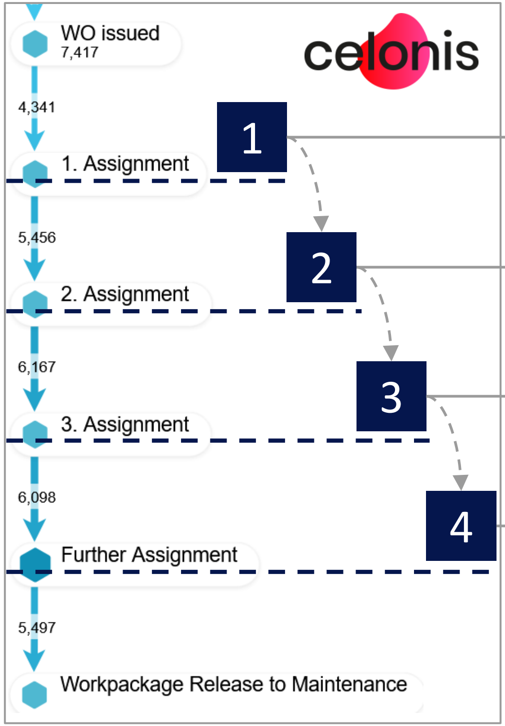
1 Feature Engineering

Input Refinement

- Planners face the **same decision** every time they re-assign a work order
- Information considered in the assignment step changes over time / the process’ lifecycle

→ **splitting the work order** at every assignment & use the **activity’s timestamp as a cut-off** criterion

Increase of available data by ≈ 33%



ID	Timestamp	Feature A	Dummy B
123_1	21.4. 6:12am	2	0
123_2	18.5. 5:24pm	2	0
123_3	22.6. 2:53pm	5	0
123_4	30.7. 9:06am	10	1

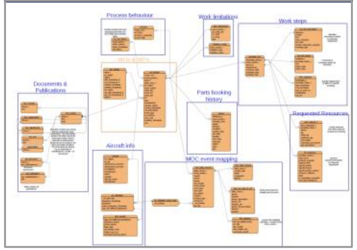
Fixed vector containing 164 features derived from 70 DB columns

Lufthansa CityLine

We transformed each workorder assignment from the database to graphs as an input for the Relational Graph Convolutional Network (R-GCN)

2 Data Preprocessing: Dataset for R-GCN

Relations from DB



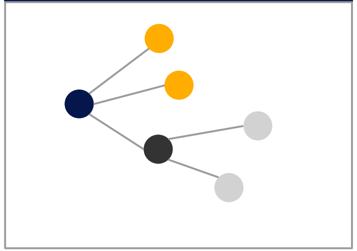
SQLAlchemy Object-Relational Mapper

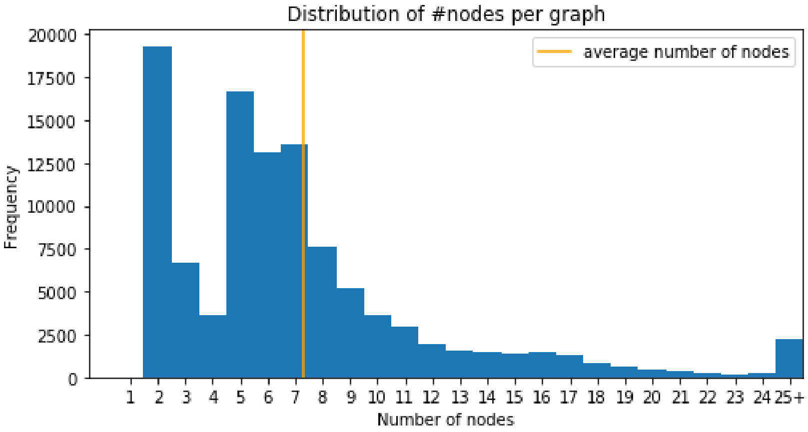
Raw Python objects

```
{
    hey: "guy",
    anumber: 243,
  - anobject: {
        whoa: "nuts",
      - anarray: [
            1,
            2,
            "thr<h1>ee"
        ],
        more: "stuff"
    },
    awesome: true,
    bogus: false,
    meaning: null,
    japanese: "明日がある。",
    link: http://jsonview.com,
    notLink: "http://jsonview.com is great"
}
```

Python **Deep Graph Library**

Graphs





- Input graphs are rather small
- On average only 7 nodes

➔ Time needed for Training and Predicting is feasible

We created 2 different datasets that are tailored towards our models while still ensuring comparability across the datasets

1/2 Feature Engineering and Data Preprocessing: Dataset statistics

	Feature Engineering	R-GCN dataset
Datatype	Handcrafted features	Raw data as graphs
Scope	70 DB columns	133 DB columns
Size	4 MB	582 MB
# Cases	133,664	

- ensure comparability
- very **imbalanced dataset**

Split	# Cases	% Incomplete Maintenance	% De-assignments	% WO Success
Train (80%)	106,931	4.0%	32.2%	63.8%
Validation (10%)	13,366	4.3%	31.8%	63.9%
Test (10%)	13,367	3.8%	32.5%	63.7%
∑ Whole dataset	**133,664**	**4.0%**	**32.2%**	**63.8%**

Lufthansa CityLine

Models

1) Logistic Regression

2) Neural Networks

3) Decision Tree

4) Relational Graph Convolutional Networks (R-GCN)

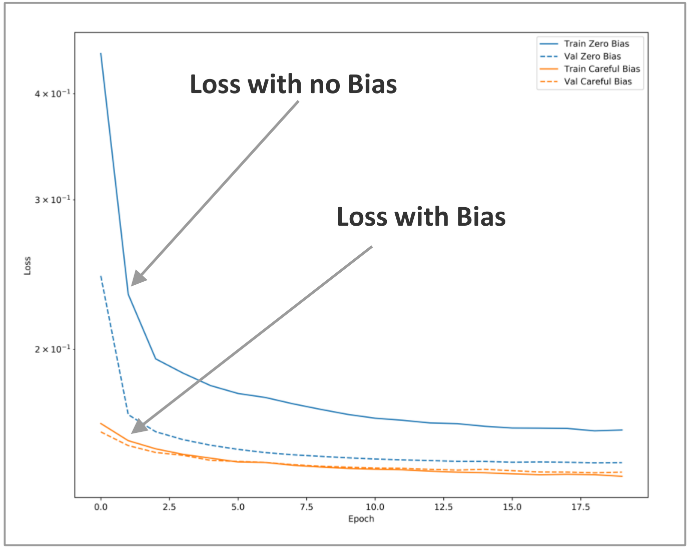
Logistic Regression is the baseline model for performance comparison in all three use cases

1 Logistic Regression

Preprocessing Techniques

1. **Bias**
2. **Class Weighting**
3. **Oversampling**
4. **Threshold Interpretation**

Explanation - Bias



Loss with no Bias

Loss with Bias

Logistic Regression is the baseline model for performance comparison in all three use cases

1 Logistic Regression

Preprocessing Techniques

1. **Bias**
2. **Class Weighting**
3. **Oversampling**
4. **Threshold Interpretation**

Explanation - Bias

Use Case - Incomplete Maintenance

Dataset:

3.96 % of class 1

96.04 % of class 0

↓

```
Weight for class 0: 0.52
Weight for class 1: 12.61
```

Lufthansa CityLine

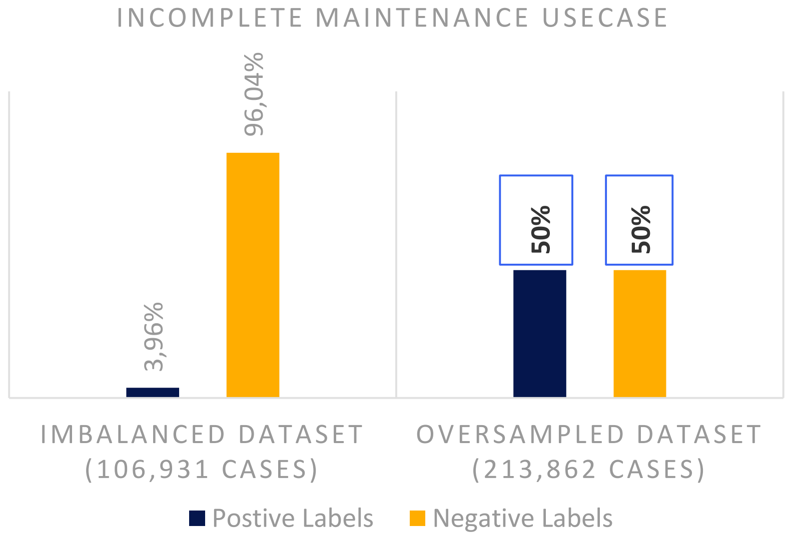
Logistic Regression is the baseline model for performance comparison in all three use cases

1 Logistic Regression

Preprocessing Techniques

1. **Bias**
2. **Class Weighting**
3. **Oversampling**
4. **Threshold Interpretation**

Explanation - Bias



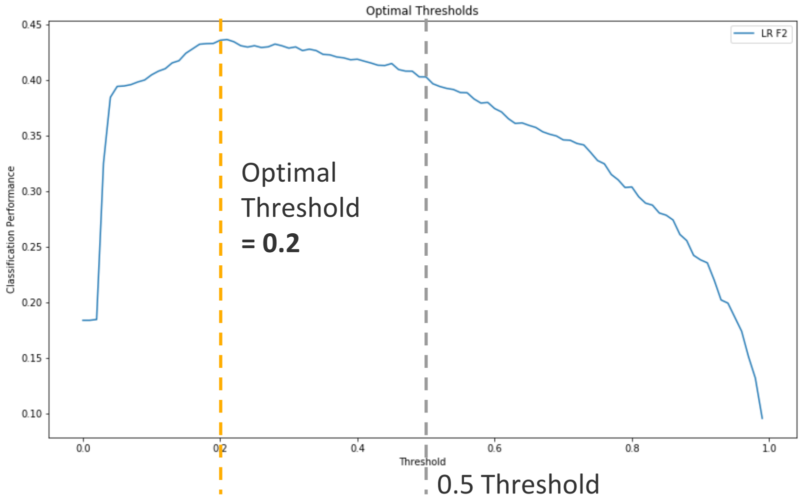
Logistic Regression is the baseline model for performance comparison in all three use cases

1 Logistic Regression

Preprocessing Techniques

1. **Bias**
2. **Class Weighting**
3. **Oversampling**
4. **Threshold Interpretation**

Explanation - Bias

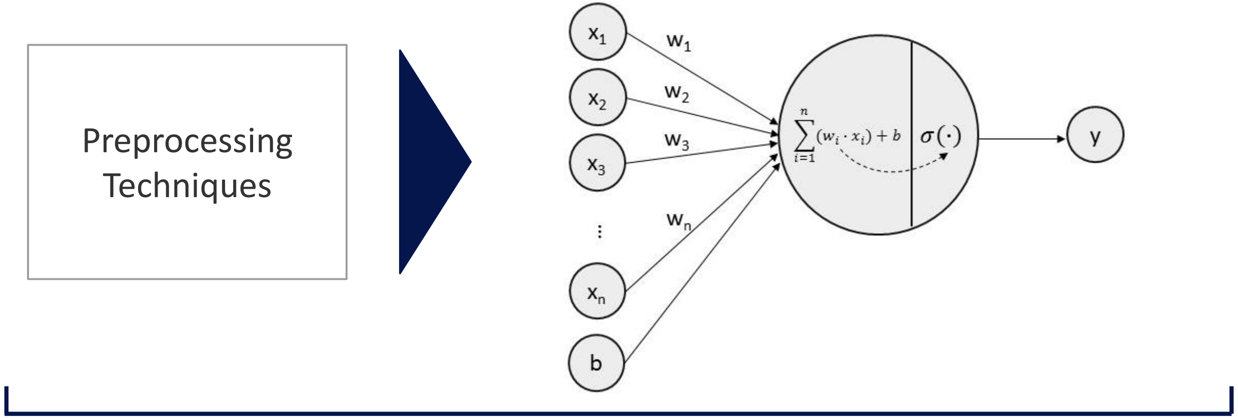


Optimal Threshold = **0.2**

0.5 Threshold

Logistic Regression is the baseline model for performance comparison in all three use cases

① Logistic Regression – The model



Preprocessing Techniques

With **Preprocessing Techniques** & **Hyperparamter Tuning**
→ **solid** and **robust** Logistic Regression **baseline model**

The Neural Network – a deeper and more advanced network build on the experience of Logistic Regression

2 Neural Network

Preprocessing Techniques & Tuning

Models

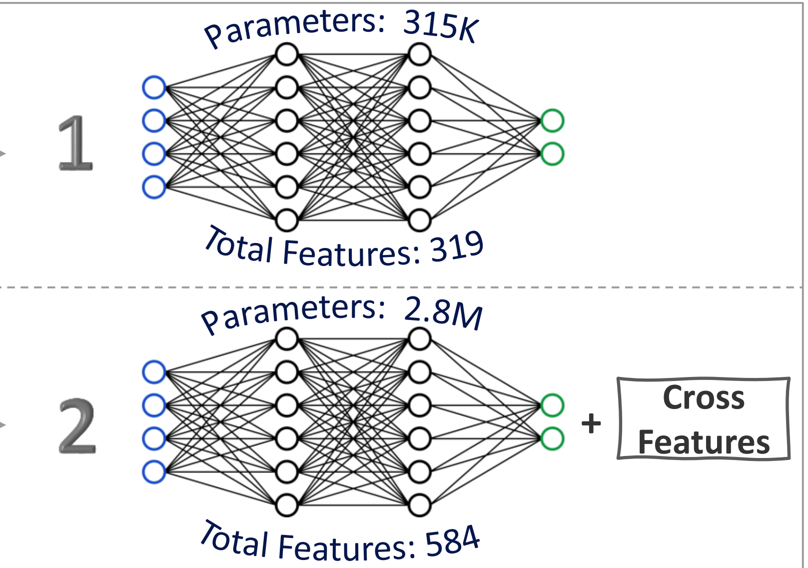
Challenges:

Feature Selection

Optimal Network Structure

Hyperparameter Tuning

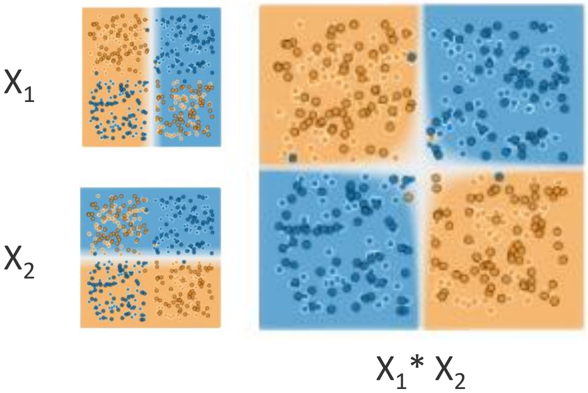
Recall/Precision Threshold & AUC discretization

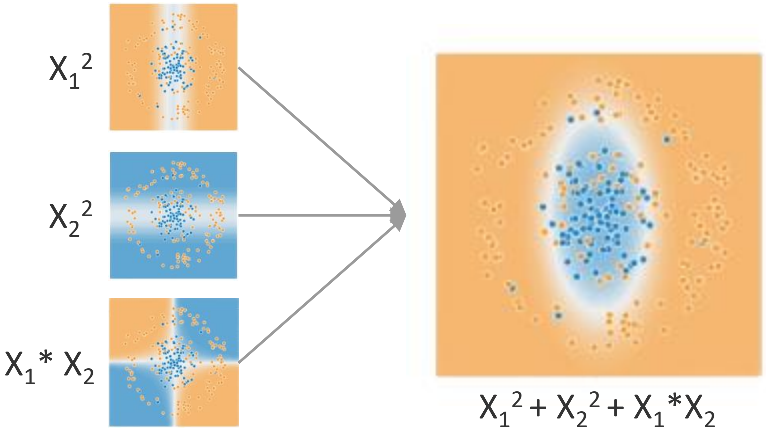


The Neural Network – a deeper and more advanced network build on the experience of Logistic Regression

2 Neural Network – Cross Features

- ☐ Formed by multiplying two or more features
- ☐ Enables a model to learn separate weights for each combination of features
- ☐ Provides better insight for the core issues of incomplete maintenance in interpretability plots





Lufthansa CityLine

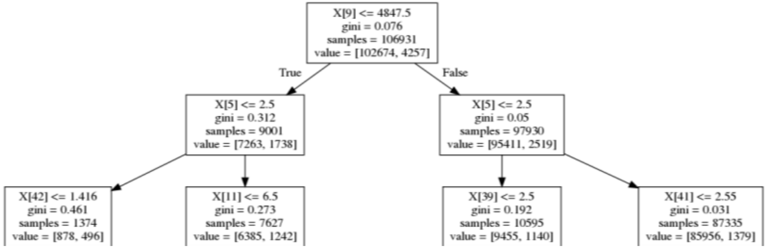
The decision tree model shows slightly better results after some adaptations and optimizations but still overfits

3 Decision Tree

Approach

Description:

- Classifier divides the observations at every node step-by-step
- At the bottom: respective the sub data set is assigned a class (e.g. incomplete maintenance)



Models and Refinement Approaches

Problem: overfitting of decision tree

Possible Solutions: (frameworks: sklearn and catboost)

- Hyperparameter optimization:
  - Depth
  - Split criterion and split method
  - Min. number of observations per leaf
- Random Forest
- Gradient Boosting
- Feature selection (incl. Interpretation)

Result: Overfitting could not be resolved
– possibly the data is too complex and imbalanced

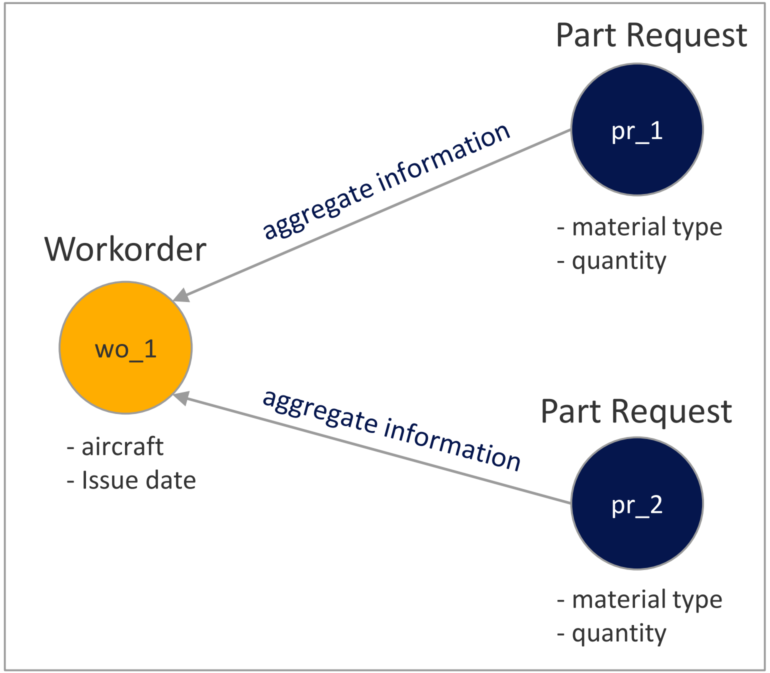
Lufthansa CityLine

The R-GCN is a powerful model but comes with some inherent challenges

4 Relational Graph Convolutional Network (R-GCN)

Challenges:

- Training time
- Feature Selection
- Overfitting on rare categories
- Imbalanced dataset



Agenda

1 Problem Statement

2 Implementation & Setup

- Academic Approach
- Data Exploration, Feature Engineering and Data Preparation
- ML Models
  - Logistic Regression
  - Neural Networks
  - Decision Tree
  - Relational Graph Convolutional Networks (R-GCN)

3 Results and Impact

- Comparison of Model Results
- Business Impact
- Explainability

4 Summary and Next Steps

We decided to compare our models on metrics that can handle imbalanced data

Evaluation Metrics

Conditions

Highly **imbalanced dataset**

Metrics

Metrics that can handle imbalanced data

- **F1 score:** equal importance of FPs and FNs
- **F2 score:** minimize FNs
- **F0.5 score:** minimize FPs
- **Decision of CLH:**
  F2 score for Incomplete Maintenance, Deassignments
  F0.5 score for WO success
- **Area Under Curve (AUC):** summarizes how well the model addresses this precision-recall tradeoff

		Actual Values	
		0	1
Pred. Values	0	TRUE NEGATIVES	FALSE NEGATIVES
	1	FALSE POSITIVES	TRUE POSITIVES

🟨 Desired to be high
☐ Desired to be low

Interpretation

The higher, the better

So far, the R-GCN shows the best performance among the different approaches

Overview: Performance of all Selected Approaches in all Use Cases

Performance/Metrics

		Incomplete Maintenance		Deassignments		WO Success	
		F2-score	AUC	F2-score	AUC	F0.5-score	AUC
Approaches	Logistic Regression	34.22	14.58	59.00	56.73	73.31	81.82
	Neural Network	41.73	31.53	75.84	73.47	82.23	89.14
	Decision Tree	19.07	-	55.21	-	83.49	-
	Relational Graph Convolutional Network (R-GCN)	46.45	39.67	**78.39**	**81.00**	**84.91**	**91.72**
	R-GCN + comb. train/valid	**49.53**	**45.71**	-	-	-	-

The higher the better

🟩 Above average
🟨 Average
🟧 Below average

Two of the use cases have a direct positive impact on the business while the third one can be used for steering

Business Case Calculations – Impact of Use Cases (1/2)

	Reducing process costs		Improving process result		Total impact
UC1: Incomplete Maintenance	Reducing rework => increased efficiency	+	Less delays, less cancellations => better service	=	++
UC2: Deassignments	Reducing the rework => increased efficiency			=	+
UC3: WO Success	KPI to track success and steer implementation			=	Neutral

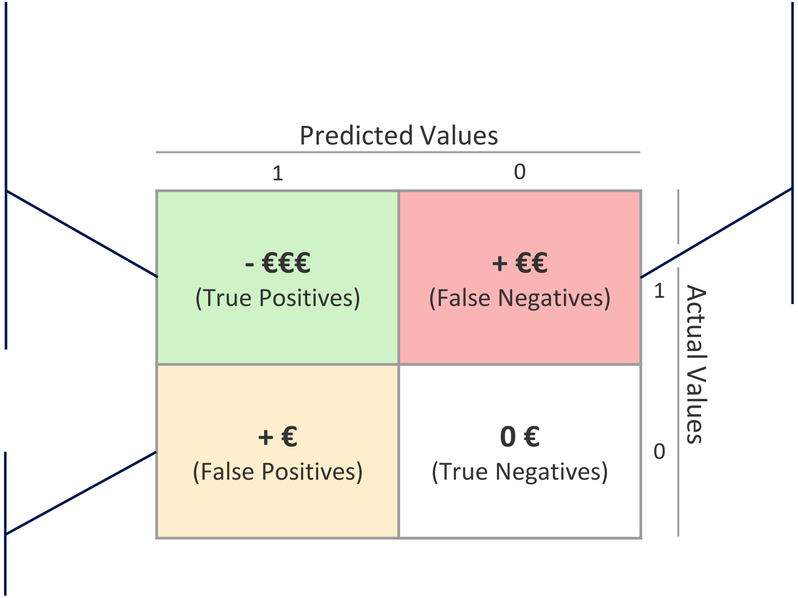
To calculate the overall impact of a ML model, we have to take the business value of all four fields of the confusion matrix into account

Business Case Calculations – Impact of Use Cases (2/2)

Deassignment:
- [Process cost of (re-)assigning a work order + cost to prevent it from happening$^{(1+2)}$]

Incomplete Maintenance:
- [Consequence cost from cancelations & delays + cost to prevent it from happening$^{(1+2)}$]

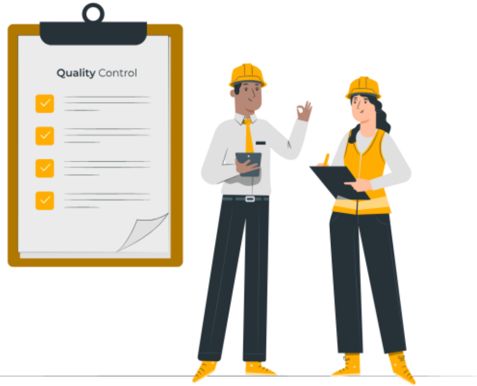
False alarm:
Process cost for the time spent to verify a prediction$^{(1)}$



		Predicted Values	
		1	0
Actual Values	1	**- €€€** (True Positives)	**+ €€** (False Negatives)
	0	**+ €** (False Positives)	**0 €** (True Negatives)

Missed Deassignment:
Process cost of (re-)assigning a work order

Missed Incomplete Maintenance:
Consequence cost (cancelations, delays) + process cost for the re-planning

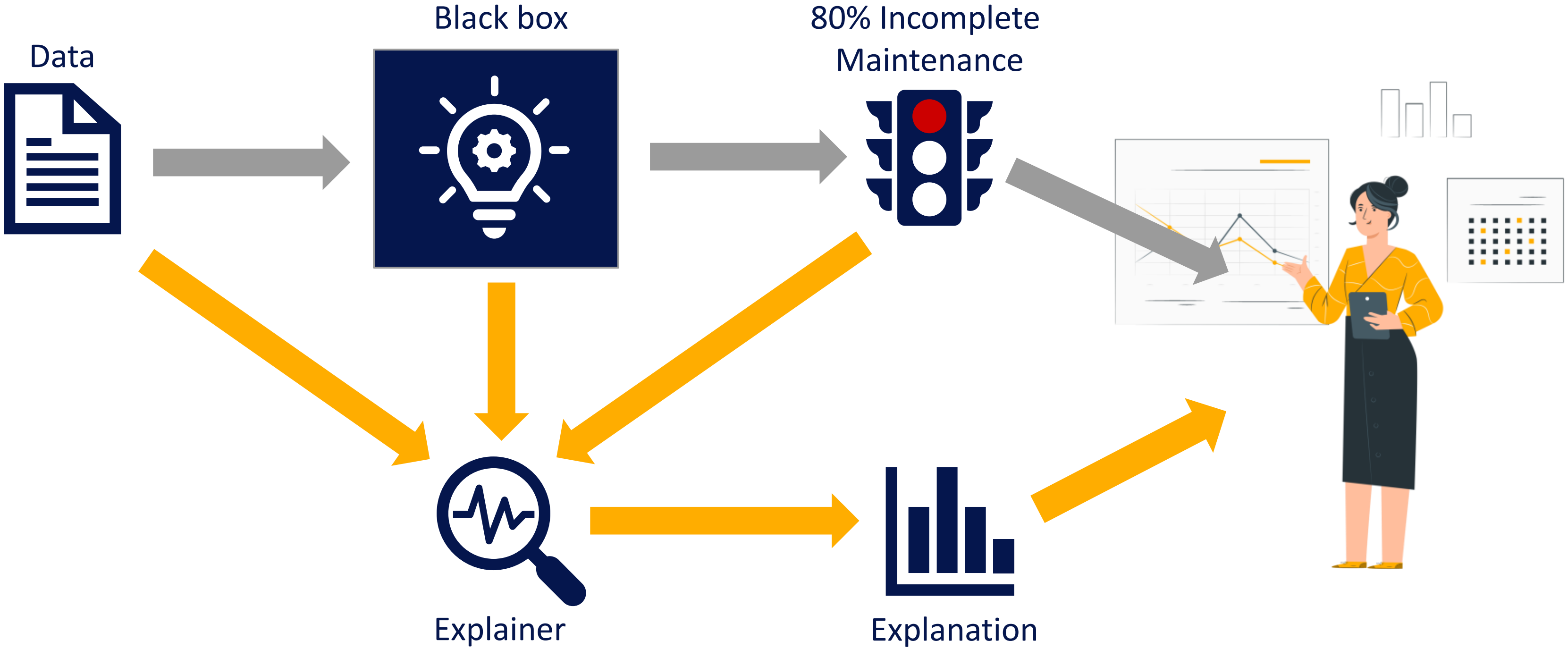


Assumptions relative to the costs for a first-time assignment

[1] verifying a prediction: 20%
[2] actively preventing it: 30%

We used explainability techniques to analyze the predictions of our models and gain further insight

Explainability (1/4)



We can either explain single predictions or the models’ general behavior

Explainability (2/4)

Explanation of a **single prediction:**

- **Why** will **this case** fail?
- Can we **trust** that prediction?

Explanation of the **model:**

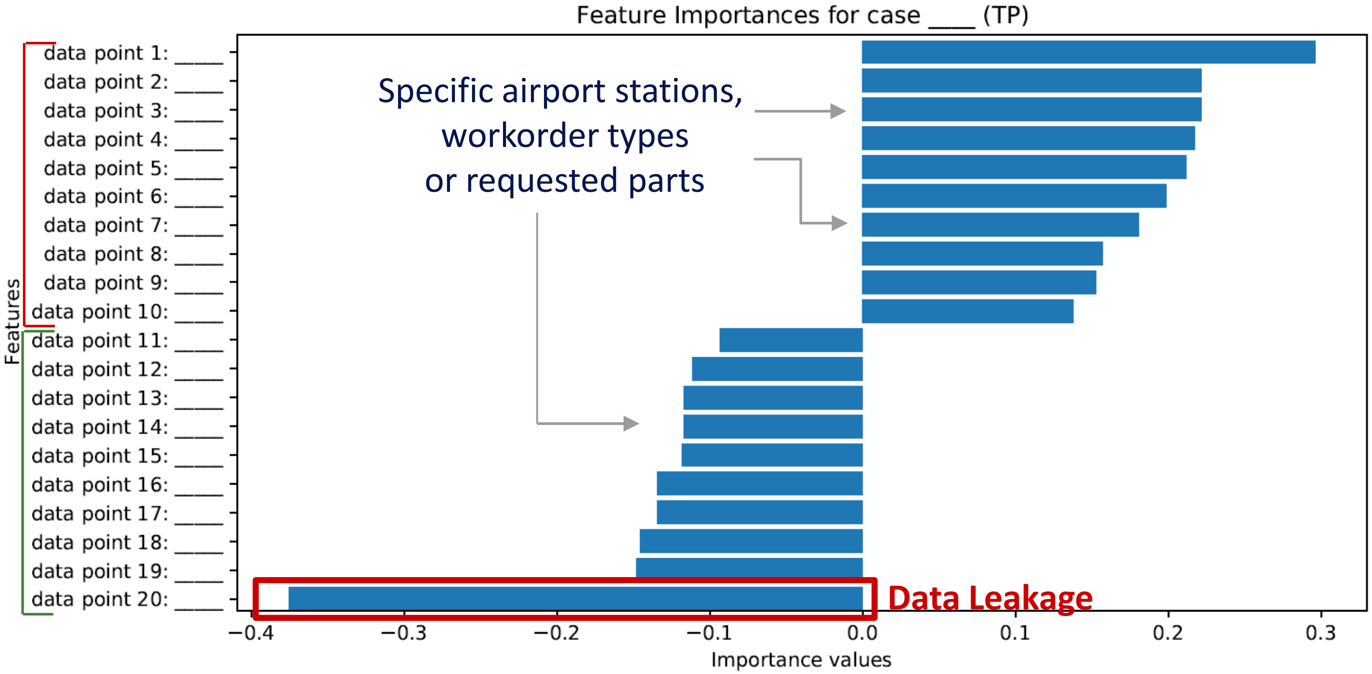
- Does the system have a **bias**?
- What **general tendencies** can we infer from the data?

We could improve the R-GCN by analyzing the explanations for single predictions

Explainability for the R-GCN (3/4)

Increase chances for Incomplete Maintenance

Decrease chances for Incomplete Maintenance



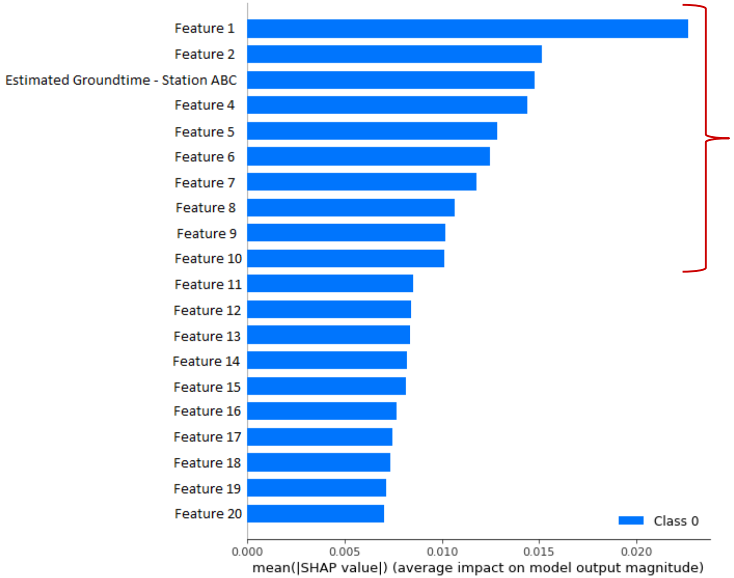
Specific airport stations, workorder types or requested parts

Data Leakage

We achieved great insights using cross features with Neural Network to identify which factors caused incomplete maintenance and deassignments

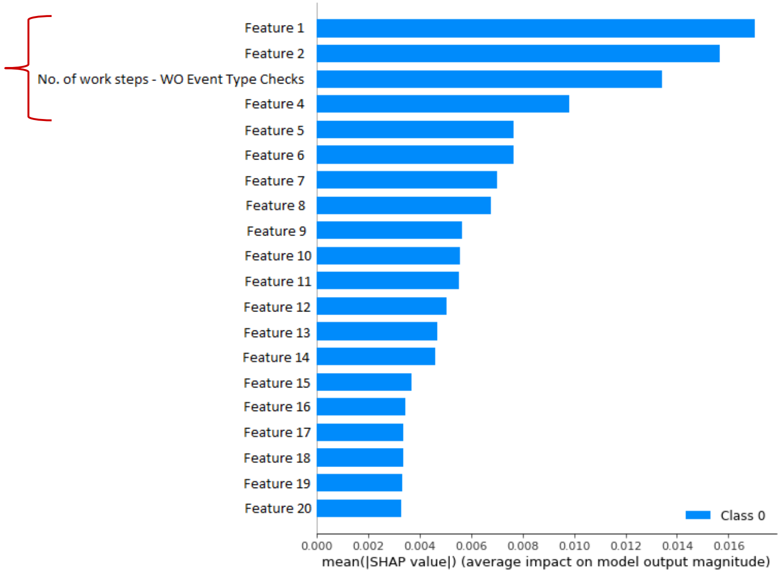
Explainability for the Neural Network (4/4)

- Analysis for 575 incomplete maintenance cases and 1000 deassignment cases



Incomplete Maintenance

Factors having higher impact on work order to be incomplete or deassigned



Deassignments

Agenda

1 Problem Statement

2 Implementation & Setup

- Academic Approach
- Data Exploration, Feature Engineering and Data Preparation
- ML Models
  - Logistic Regression
  - Neural Networks
  - Decision Tree
  - Relational Graph Convolutional Networks (R-GCN)

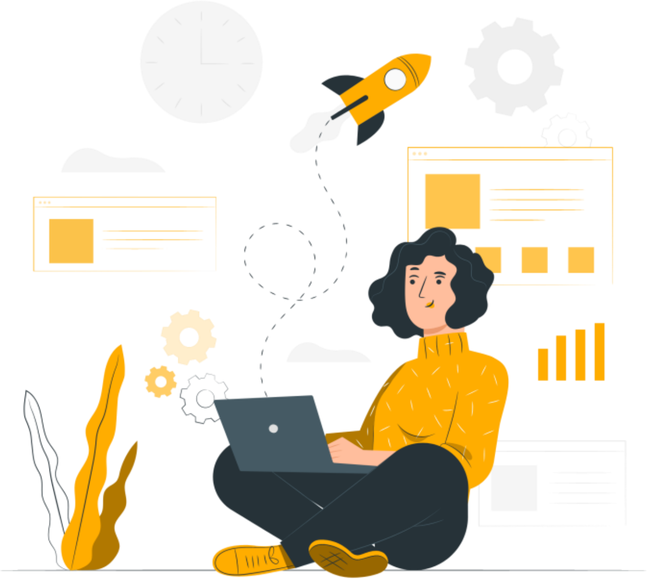
3 Results and Impact

- Comparison of Model Results
- Business Impact
- Explainability

4 Summary and Next Steps

Paula gained detailed insights into the root cause of her daily challenges

Management Summary



Problem: Incomplete maintenance and deassignments as challenges in the MRO process for planners at CLH's planning department

Approach and Solution:

- Implementation of **three different ML models** to predict three use cases (incomplete maintenance, deassignments, work order success)
- Challenging because of **high complexity** and **imbalance of the data**
- **R-GCN as best model**
- **Positive business impact** for UC1 and UC2, neutral impact of UC3
- Interpretation hints at **different important variables** e.g. required parts

- Reduced workload for Paula
- Knows how to reduce issues
- Can focus on planning new work orders

Lufthansa CityLine

What does that mean for CLH at what are the possible next steps?

Business Impact and Next Steps

Business Impact

☑ **Feasible to predict MRO process outcomes**

Overall positive value of three use cases, high potential impact

Next steps and tasks

Improving the ML model
- Adding additional data
- Data augmentation
- Improving R-GCN by integrating textual information e.g. descriptions of the work steps
- Evaluation of the feature importances with experts

Implementation in MRO Process
- Presentation of results at CLH internally (end of August)
- Evaluation of necessary adaptation in CLH workflow with regards to implementing the models

Lufthansa CityLine

Thank you very much for your attention