inwt

Full Stack Data Science

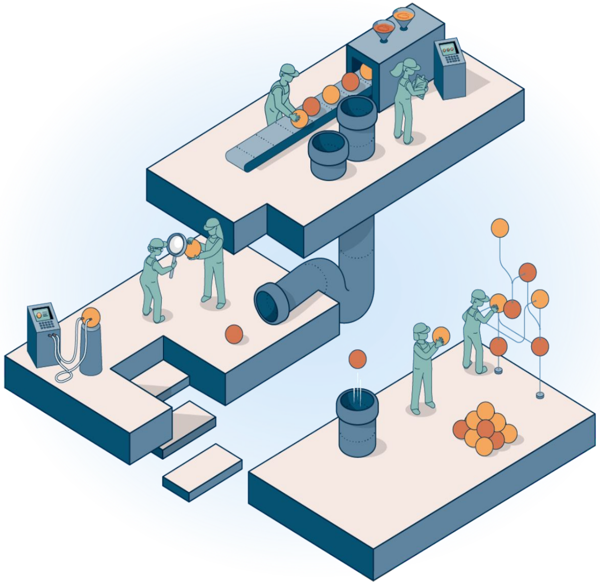
The importance of explainability of statistical models in consulting projects

Full Stack Data Science

Predictive Analytics - Software - Operations

INWT = In Numbers We Trust.

We believe in advancing businesses and people through data.



Agenda

1. Why explainability matters
2. Our journey from statistics to ML
3. SHAP: local and global explanations
4. Explainability in customer projects
5. Summary

ML: Is it all about predictive performance?

"Old world"

- Feature engineering (~60-80% of the time)
- Statistical models / manual model optimization

vs. "new world of ML"

- ML allows to optimize models automatically
- Notion to strongly focus on performance metrics (goodness of fit, …)
- Example: kaggle competitions kaggle

Most people distrust black boxes

Reasons:

- Compliance: Accountability (EU AI Act)
- Risk aversion: automated processes can generate damage
- Desire to understand algorithms / curiosity
    - Valuable insides by understanding processes
    - Robustness: limitations of training data
    - Plausibility of learned patterns (correlation vs. causality)
    - Omitted variables
    - Trust/Arguments for discussions with stakeholders
    - ...

Why explainability matters (1)

About the difference between a bold head and a ball…

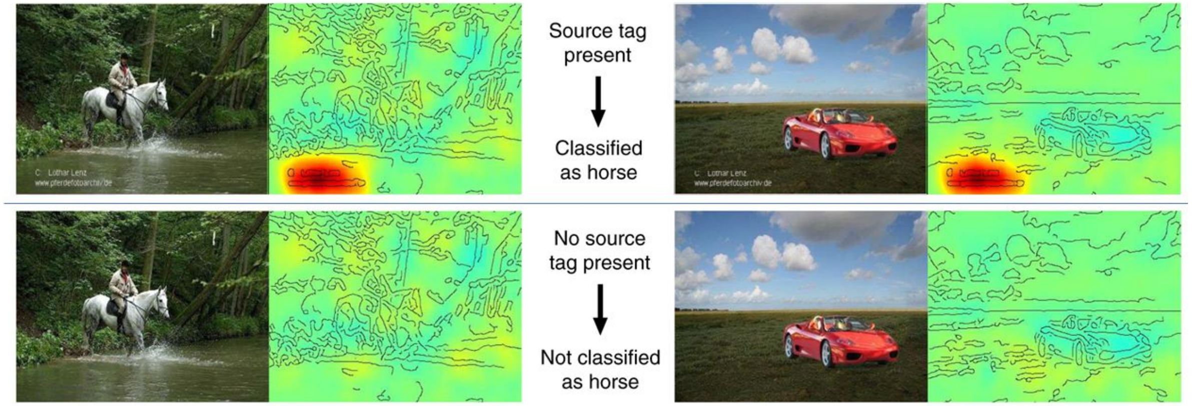
Video/Story from: The Verge (2020): AI camera operator repeatedly confuses bald head for soccer ball during live stream

[Video removed to decrease file size]

- AI operated camera should follow the ball in a soccer match
- Instead, the camera is switching around between ball and the referees bold head
- lack of proper training data; could have been detected easier with XAI techniques (see next slide)

Why explainability matters (2)

Horses and cars



- Neural network classifies horses based on source tag on training images instead of recognizing shape of the horse

S. Lapuschkin et al.: Unmasking clever hans predictors and assessing what machines really learn. Nature communications, 10(1), 2019.
https://www.nature.com/articles/s41467-019-08987-4

Tradeoff between interpretability and predictive power

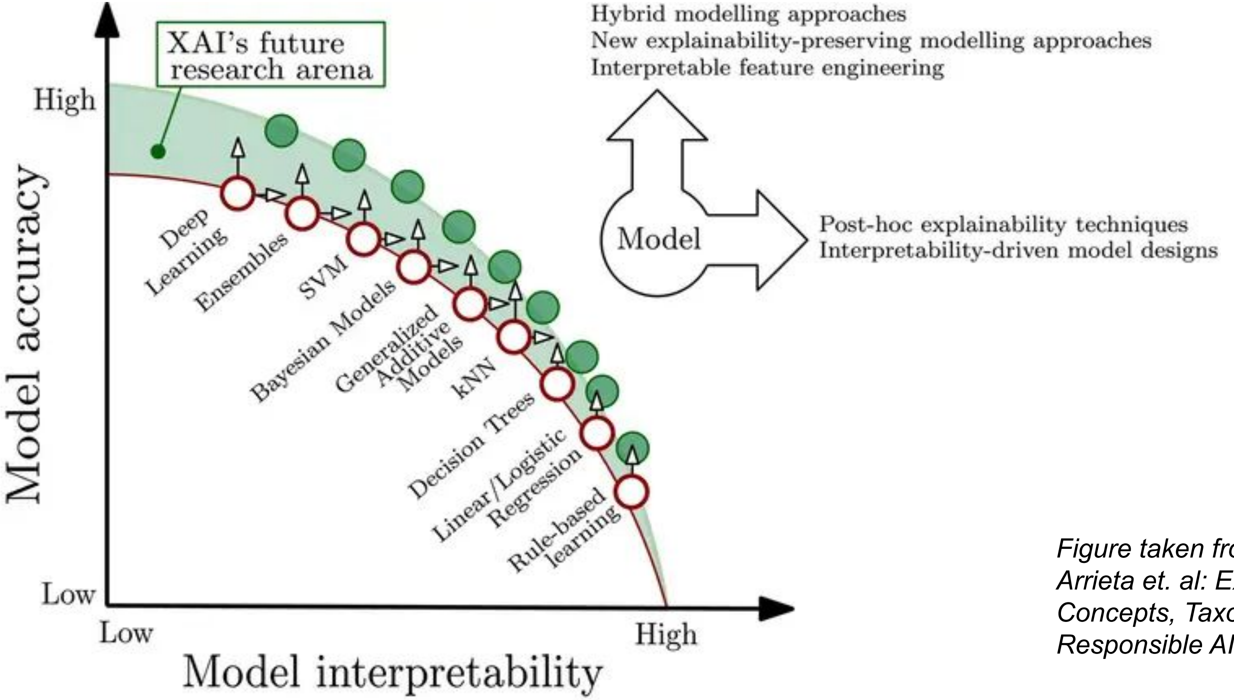


Figure taken from:
Arrieta et. al: Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI. Information Fusion. 2020.

Our journey at INWT from Statistics to ML

most popular generic method: **GLM/GAM**

ML to support feature engineering / to assist model development

XGBoost / CatBoost have become most popular

- only after SHAP has been fully integrated

Model-agnostic XAI

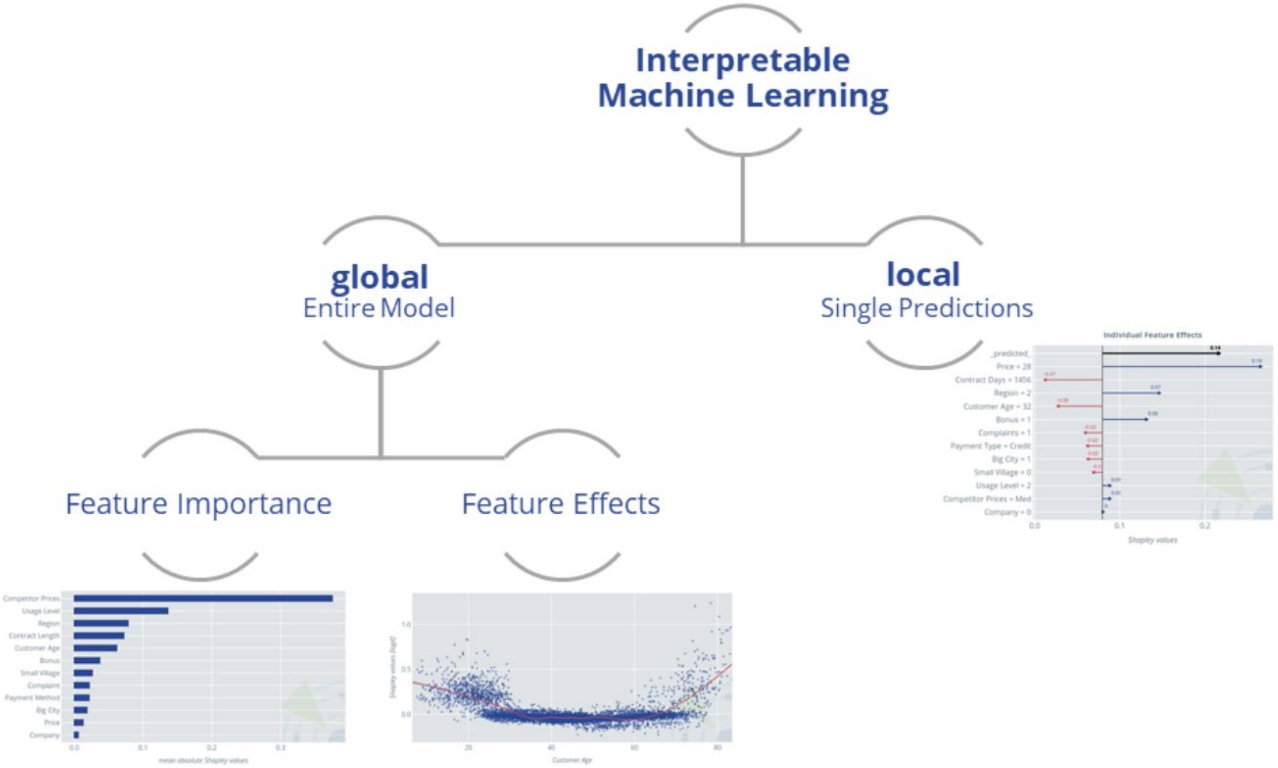
General idea:

- Gain understanding about black box model from its predictions
- Only requirement: predictive function $\hat{y} = f_{BB}(X)$
- Quantify changes in $\hat{y}$ due to controlled changes in X

XAI with SHAP:

- **SH**apley **A**dditive ex**P**lanations
- The Shapley value is the average marginal contribution of a feature value across all possible coalitions

SHAP: local and global explanations



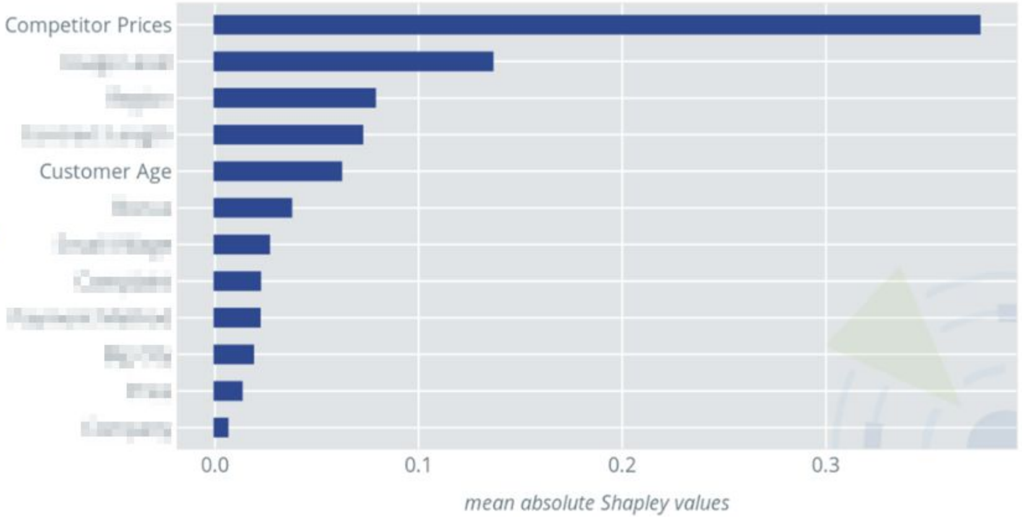
Use Case: Churn Prediction

Setup:

- Task: Predicting customers probability to churn in contractual relationship
- Dependent variable: Churn within the next 12 months
- Requirements:
  - High predictive power to steer interaction with *individual* customer
  - Microscopic understanding of churn mechanism
- Data: >100T customers
- Model approach: XGBoost

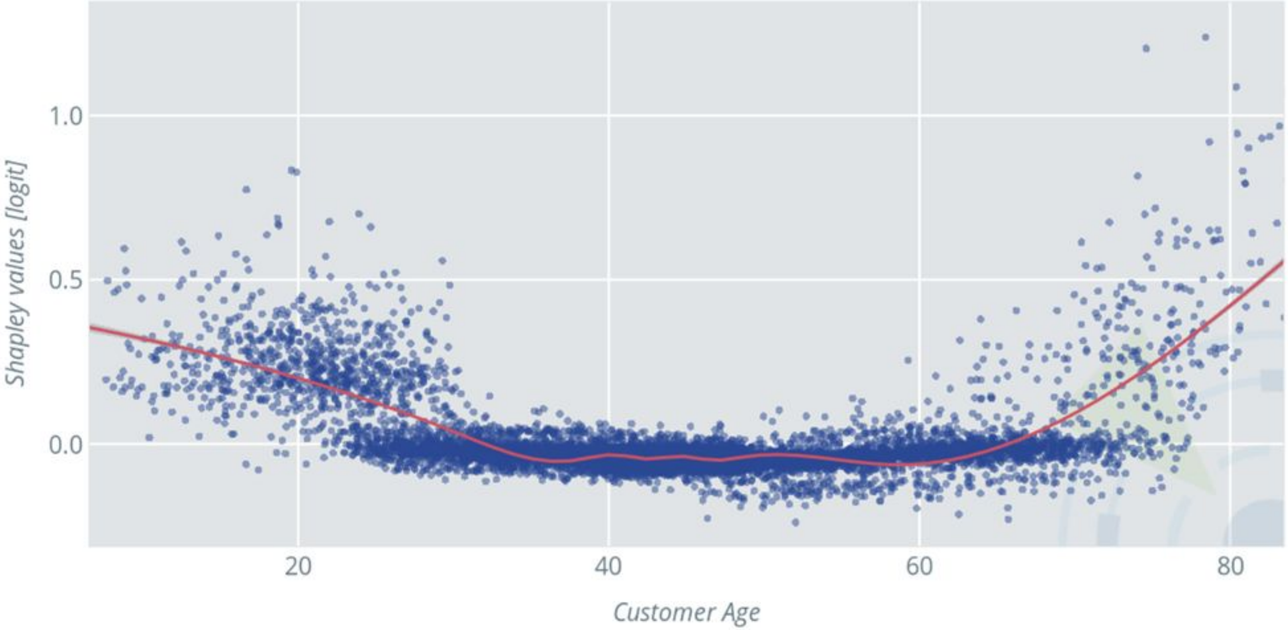
Use Case: Churn Prediction

SHAP: Feature Importance



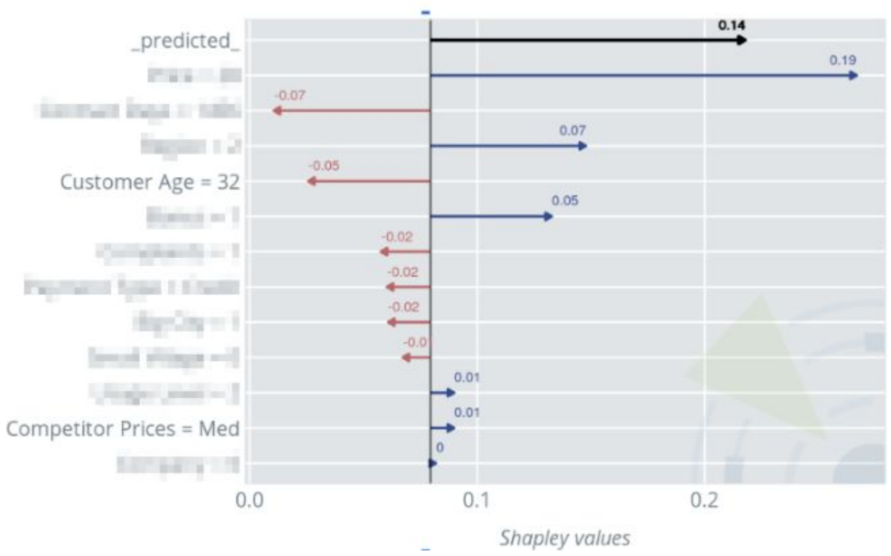
Use Case: Churn Prediction

SHAP: Feature Effects



Use Case: Churn Prediction

SHAP: Feature contribution on level of individual prediction



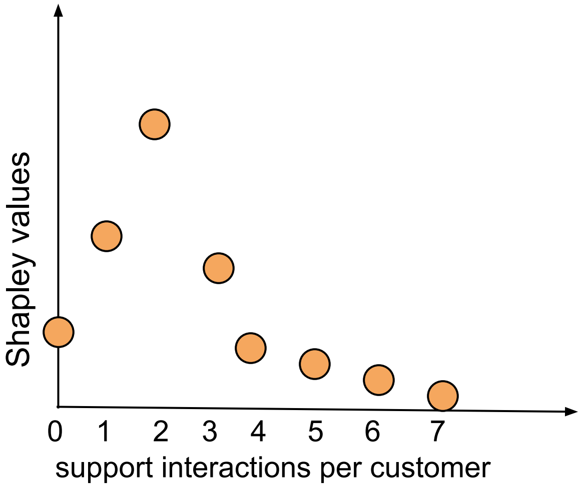
Use Case: Customer Lifetime Value

Setup:

- Task: Predicting (discounted) value of future returns from customer
- Dependent Variable: future returns from customer
- Model Approach: Survival analysis, additional models

Use Case: Customer Lifetime Value

Feature Effects: Insides by understanding processes



- Status quo: Customer support is generating costs, “bad” customers are keeping support busy
- Learning:
  - Product doesn’t work out of the box
  - Many customers are having issues
  - Customers who believe *it’s worth solving the problem* are reaching out to support
  - If support can solve the issues → customer becomes loyal
- New understanding: Customer support is key for generating loyal customers

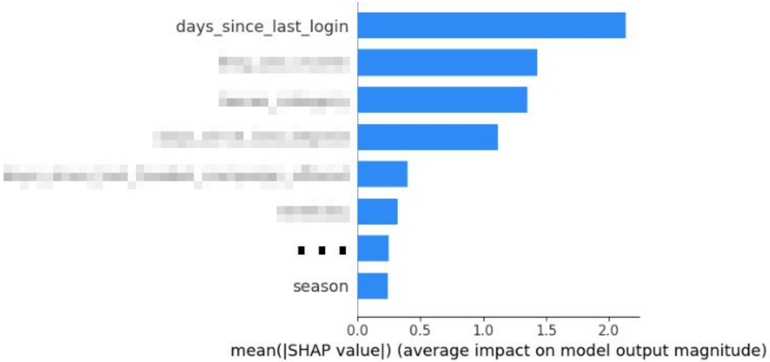
Use Case: Uplift Modelling for Marketing

Setup:

- Task: Predicting uplift of marketing incentive on customer activity
- Dependent variable: Customer activity (active: yes/no, no. of transactions) within next 14 days
- Model approach: CatBoost

Use Case: Uplift Modelling for Marketing

Feature Importance: Identifying information leakage



- Model performance looks suspicious (AUC > 0.9)
- Possible issue: Information leakage
- First checks did not identify the issue
- Feature importance plot shows outstanding importance of single feature
- Issue: Instant customer reaction on marketing incentive
- Insufficient precision of date-time-variable used for cutting the data

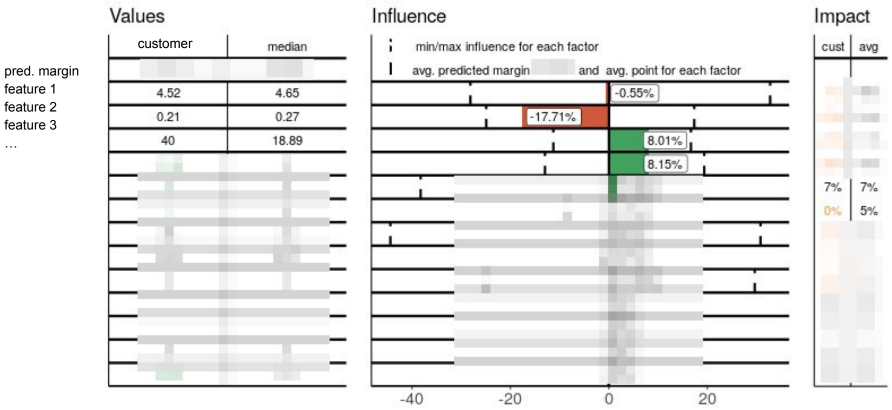
Use Case: Risk assessment

Setup:

- Task: Predicting the expected profitability of a customer
- Dependent variable: Expected (financial) margin of customers’ transactions
- Model approach: Ensemble of various models (incl. mixed effects-model)

Use Case: Risk assessment

SHAP: Feature contribution on level of individual prediction

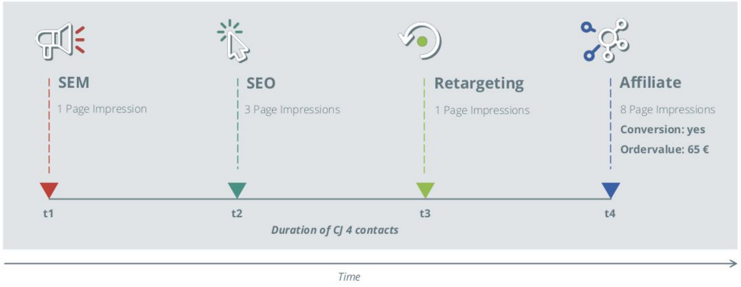


- Predictions integrated in automated processes
- Dashboard to look into single transactions and understand model decision

Use Case: Multi-channel attribution model

Setup:

- Customer: Multi-channel shop (online, mail-order with catalogue)
- Task: Predicting/measure contribution of different marketing channels on sales
- Dependent variable: time of conversion for given customer journey
- Model approach: Survival model



Use Case: Multi-channel attribution model

Arguments for discussions with stakeholders

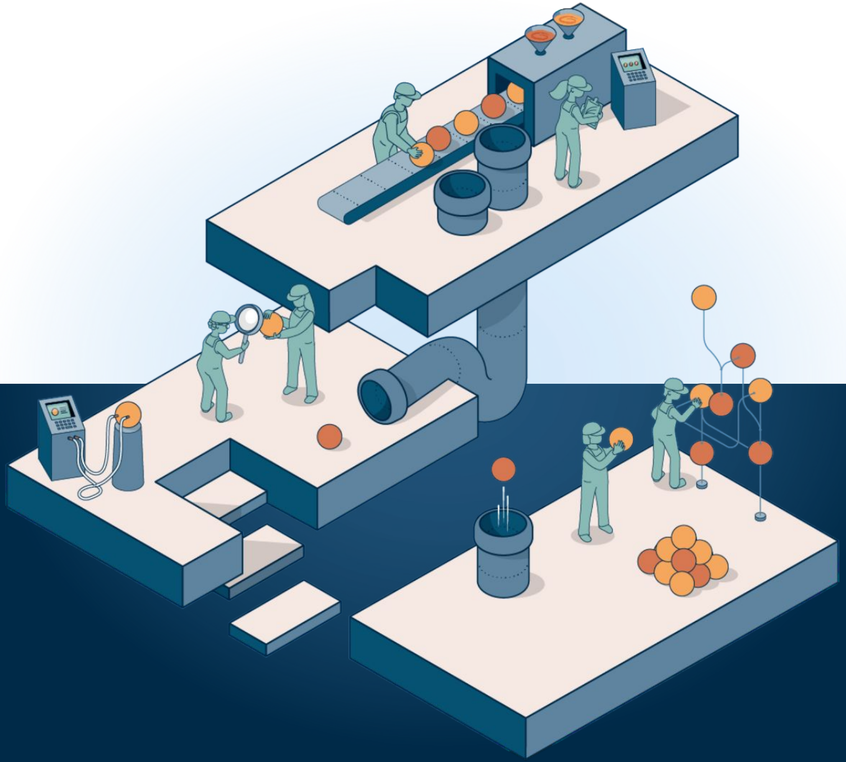
- Common to use heuristics/other backbox approaches in attribution modelling
- Relevance:
  - Marketing channels are handled by separate departments
  - Attribution results determine future budget of departments
  - Tension between catalogue and online units
    - Status quo: Attribute any sale (100% turnover) to catalogue, if catalogue has been sent to household within X weeks prior to order date
- Client communicated:
  - Explainability (both global and local) is a must-have
  - Units will disregard undesirable results unless sufficient trust can be established

Summary

- Explainability helps to **prevent mistakes**
    - Robustness, correlation vs. causality, information leakage
- Many **clients demand** explainability for various reasons
    - Curiosity, precondition to trust a model → best practice
- post-hoc XAI techniques make **models exchangeable**
    - Same interface for conceptually different models
- Integration of SHAP changed our **model preference**
    - Allows us to leverage advantages of some ML models
- **Caveats**:
    - Computational complexity
    - Choosing hyperparameters
    - SHAP not always appropriate
- Outlook:
    - Trustworthiness: Explainability + *quantification of uncertainty*

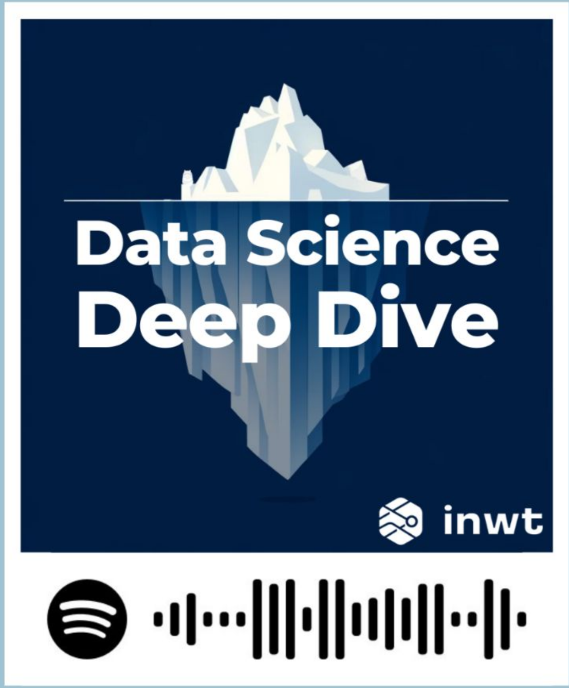
inwt

Full Stack Data Science



Danke für Ihre Aufmerksamkeit.

INWT Statistics GmbH
+49 (0)30 1208231-0
www.inwt-statistics.de
info@inwt-statistics.de



SHOW

Data Science Deep Dive

The Data Science Podcast by inwt