ADVECTAS

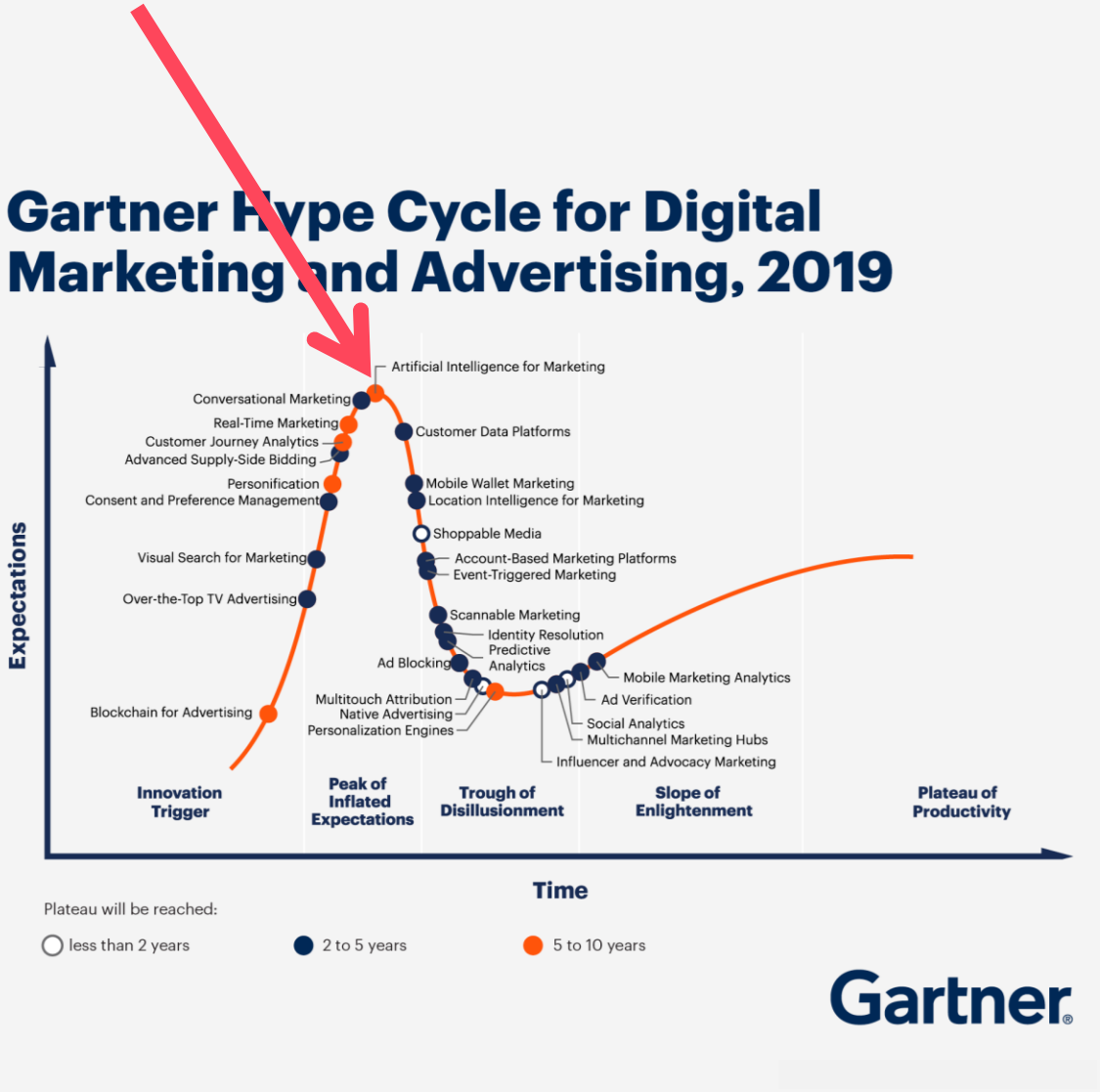
Excellence in your decisions

victor.backman@advectas.se

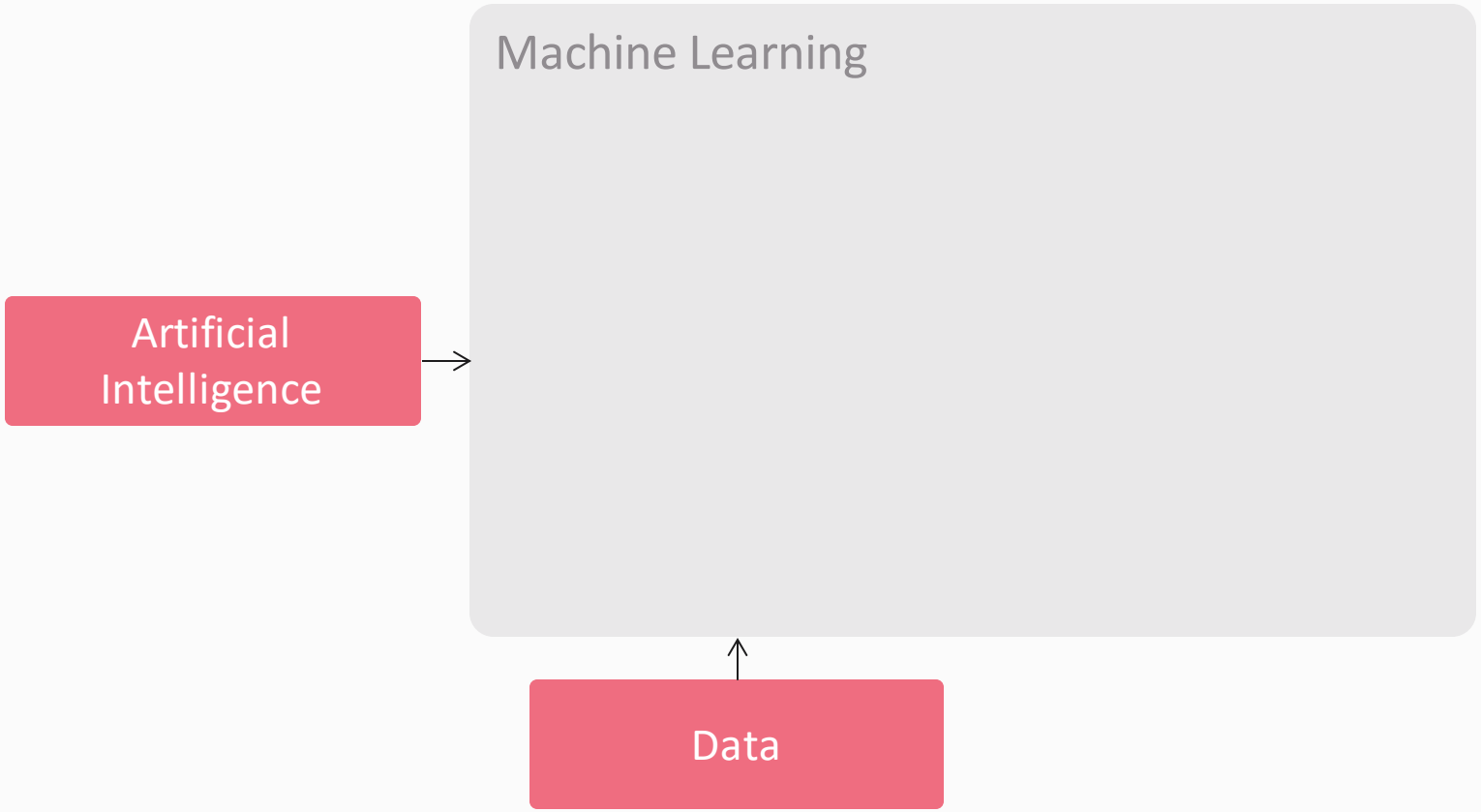
We do Analytics. Nothing else.

AI HYPE

- AI is often predicted to 'revolutionize' marketing.
- According to Gartner's latest report on digital marketing[3], AI has since last year reached the peak of inflated expectations.



[3] Gartner, 2019

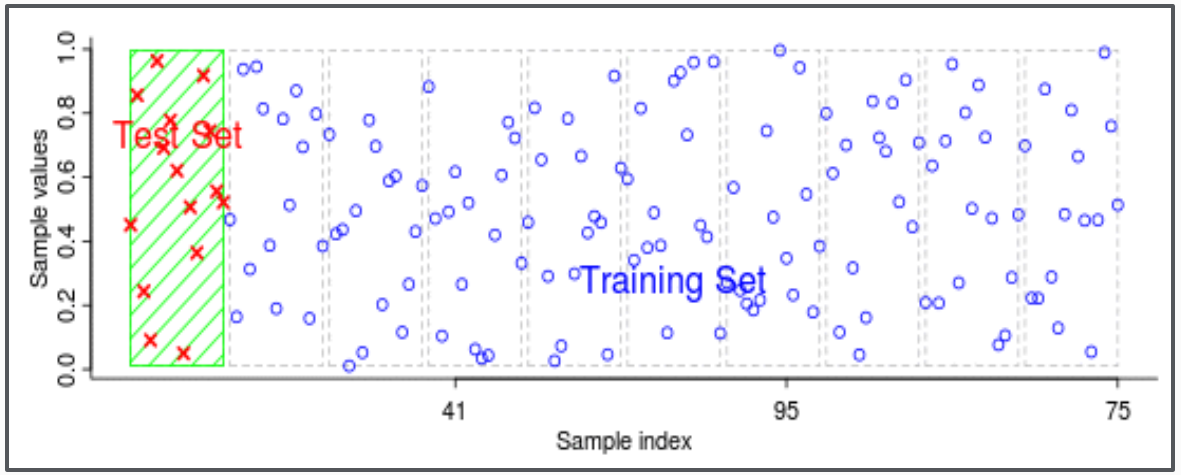


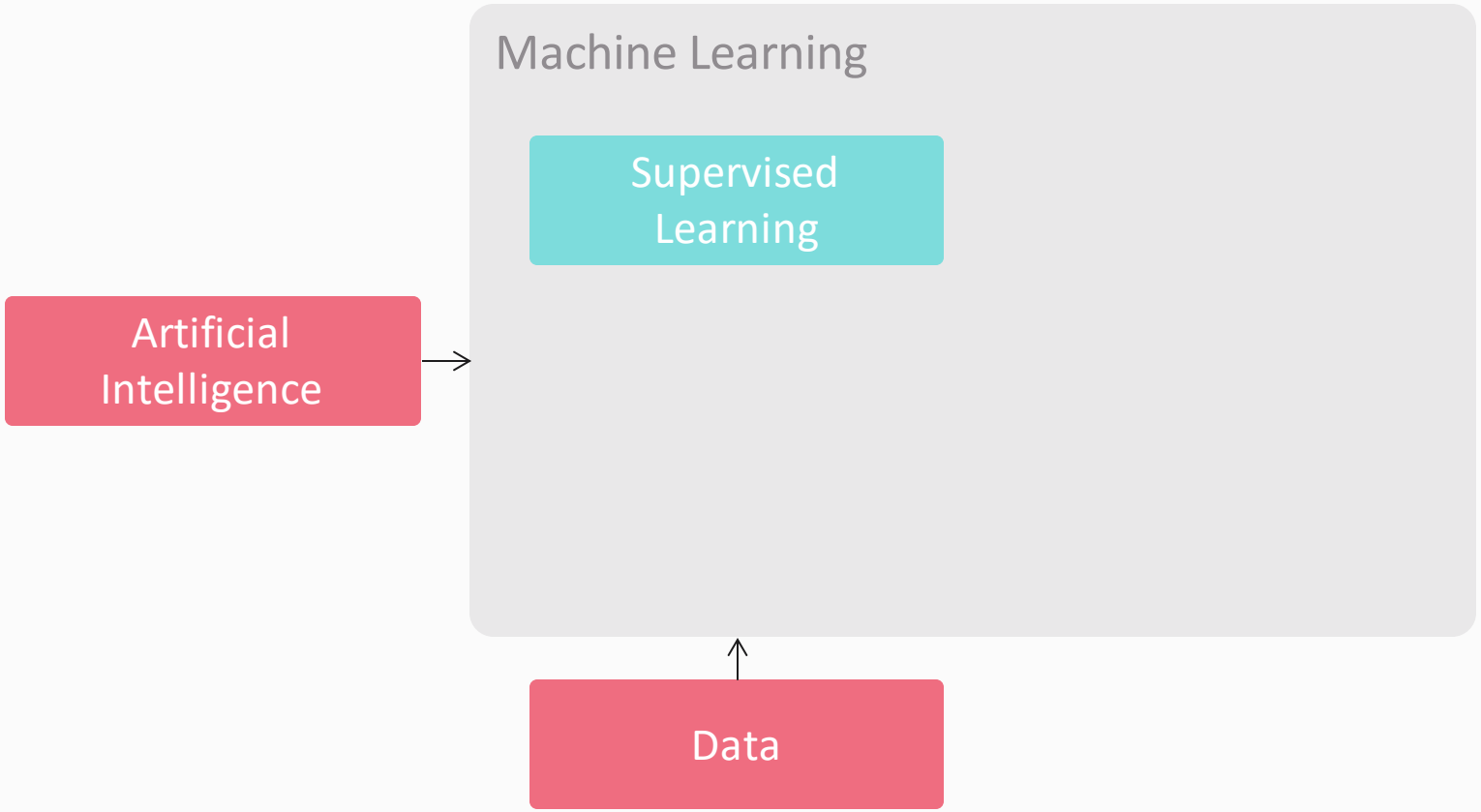
victor.backman@advectas.se

MACHINE LEARNING IS ”TRAINED” ON DATA

Split the data into a train set and a test set
(much like school exercises and exams)

Count number of correct answers produced by the test set and give it a “grade” (accuracy, precision, recall, etc.)



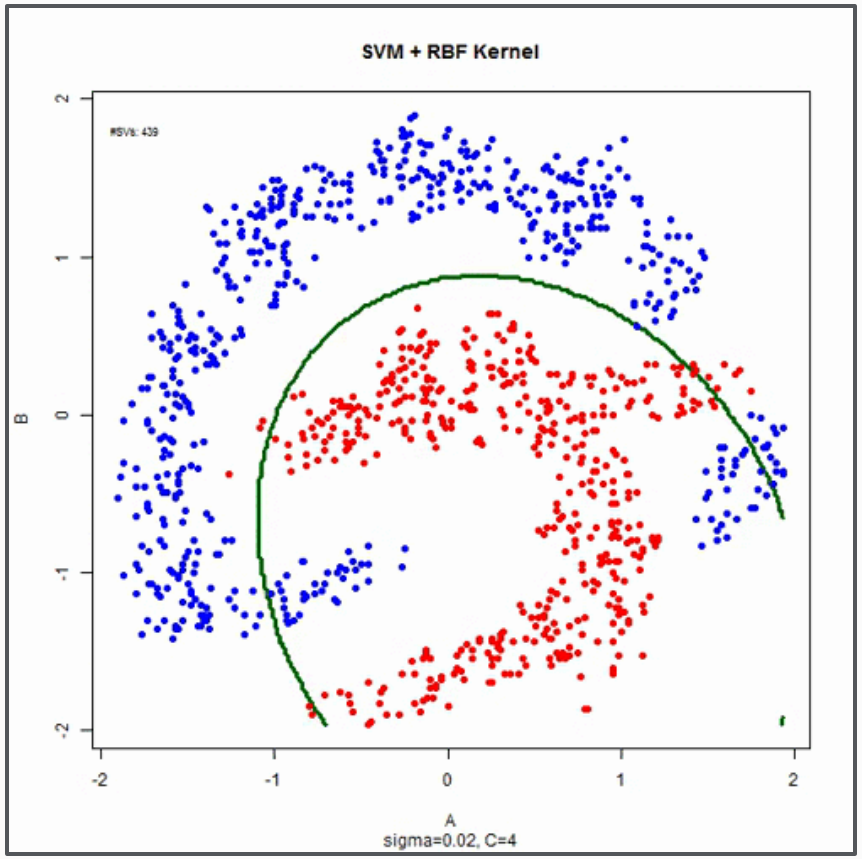


victor.backman@advectas.se

SUPERVISED LEARNING: CLASSIFICATION

Needs historical data labelled as different classes (e.g. fraud=1, not fraud=0)

Generate one output (predicted classification) for every data point



PREDICT CHURN

Challenge

Some customers are more likely than others to switch supplier (price hunters). The customer commits to a new supplier, which makes win-backs tough.

Solution

Predict whether a customer will terminate its contract. The prediction is based on factors such as type of agreement, consumption patterns, and interactions.

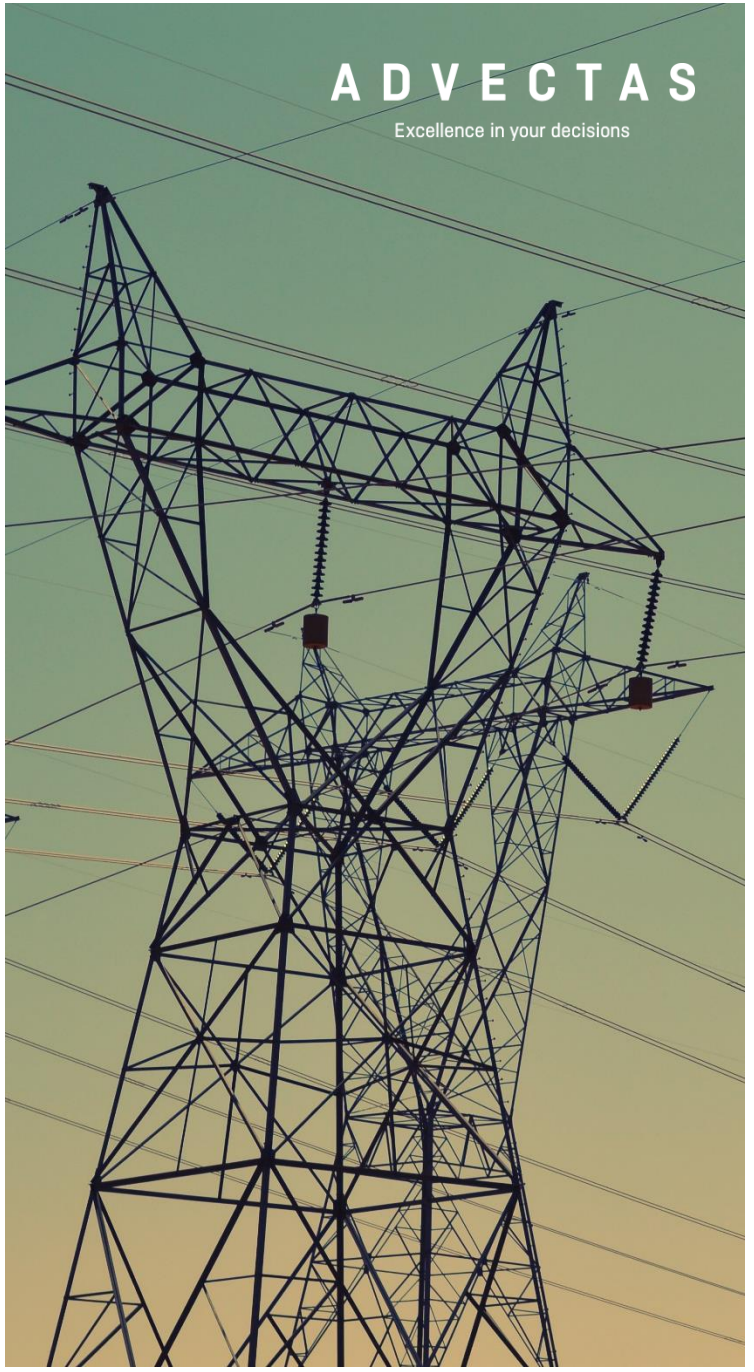
Tech

Classification with a XGBoost model based on a snapshot of recent interactions with the company, type of agreement, consumption, etc.

Business value

The company can now prioritize its marketing better for retaining the customer base, rather than winning back they who have already left. They have also made themselves more relevant to customers by providing customized offers at the right time.

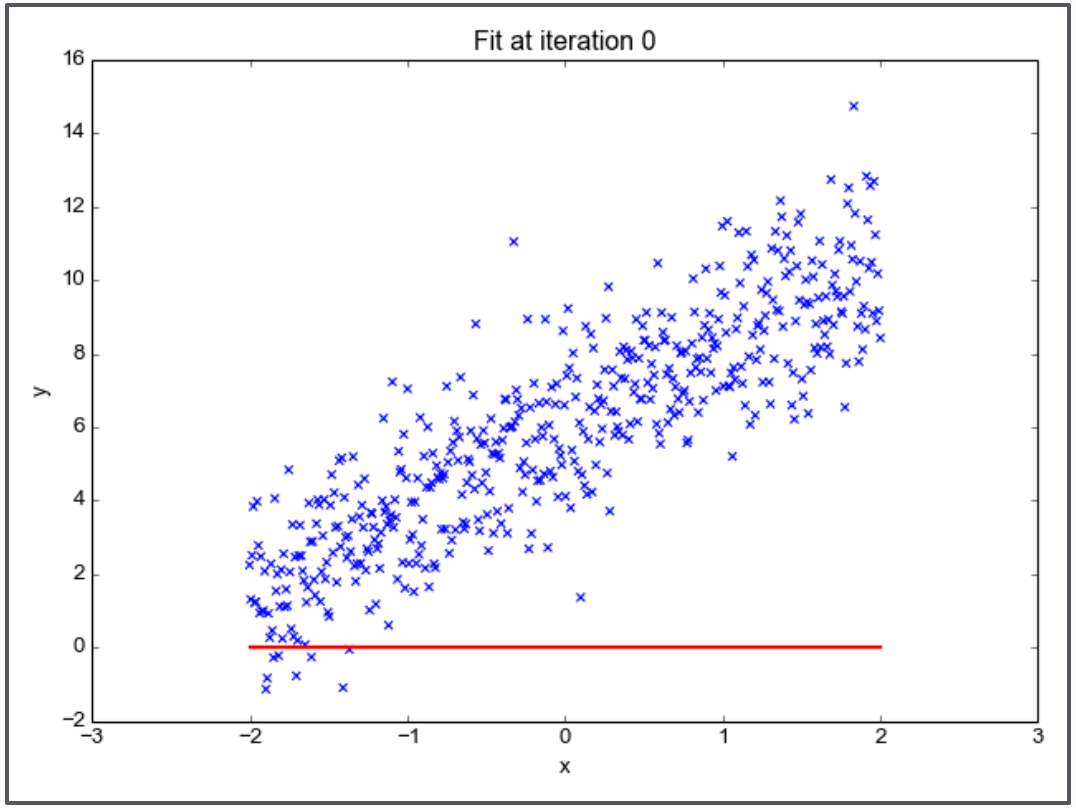
< Energy company >



SUPERVISED LEARNING: REGRESSION

Needs historical data with some target value

Predict one output (predicted value) for every data point



PREDICT CROSS-SALES

Challenge

Some customers only buy a few types of the bank's products / services. Without knowing the customer's preferences, it is difficult to be relevant in their marketing.

Solution

Twin-based model based on customer data, historical transactions and activities. Consumption of different product groups per customer was predicted as a sales-potential and then compared with actual sales to find unutilized potential.

Tech

LightGBM model written in Python and deployed on a Microsoft SQL-server.

Business value

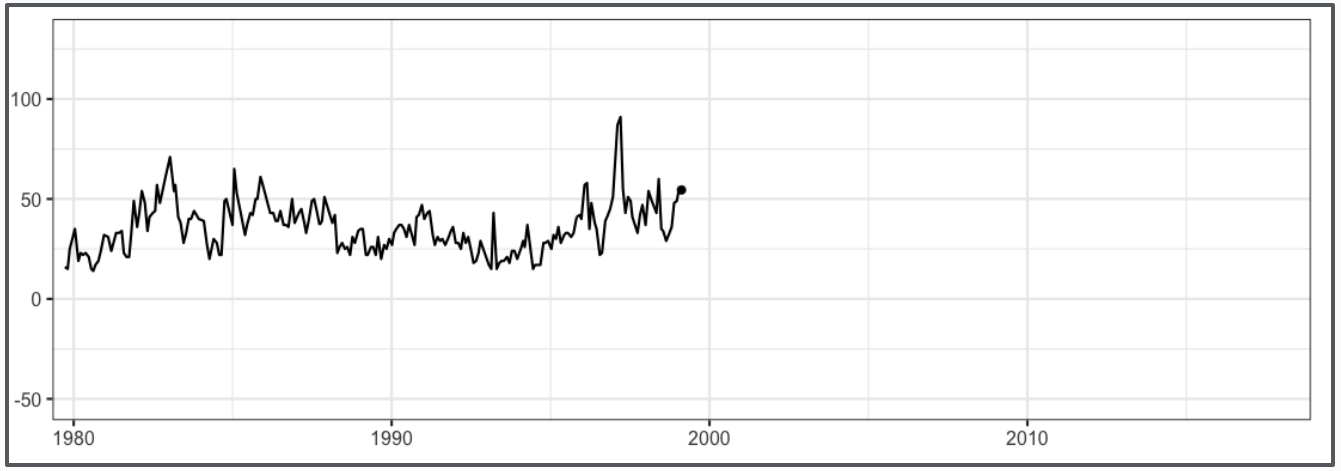
Targeted campaigns and offers to customers with high purchasing potential per product / service instead of broad campaigns in order to become more relevant to customers with increased sales as a result.

< Bank & Insurance company >

SUPERVISED LEARNING: FORECASTING

Needs historical data with a timestamp

Predict multiple steps forward



FORECAST NUMBER OF VISITORS

Challenge

The number of visitors is Liseberg's most important KPI. It's strongly dependent on external factors such as weather, other events in Gothenburg, season, weekday, holidays, etc. But also internal factors such as is there an artist playing on the big stage, is there a theme in the park, opening hours, etc.

Solution

Predict the number of visitors per hour, 10 days ahead, based on historical internal data as well as external data sources (such as SMHI and last.fm).

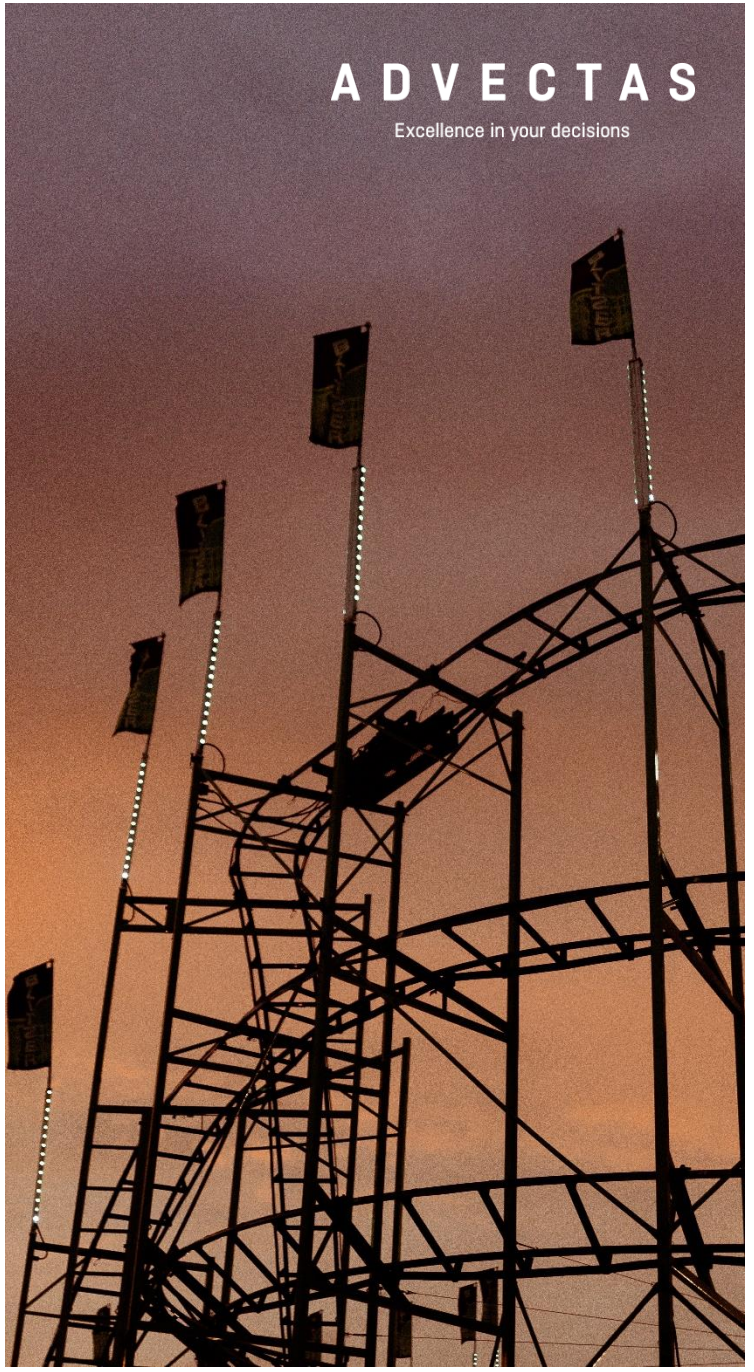
Tech

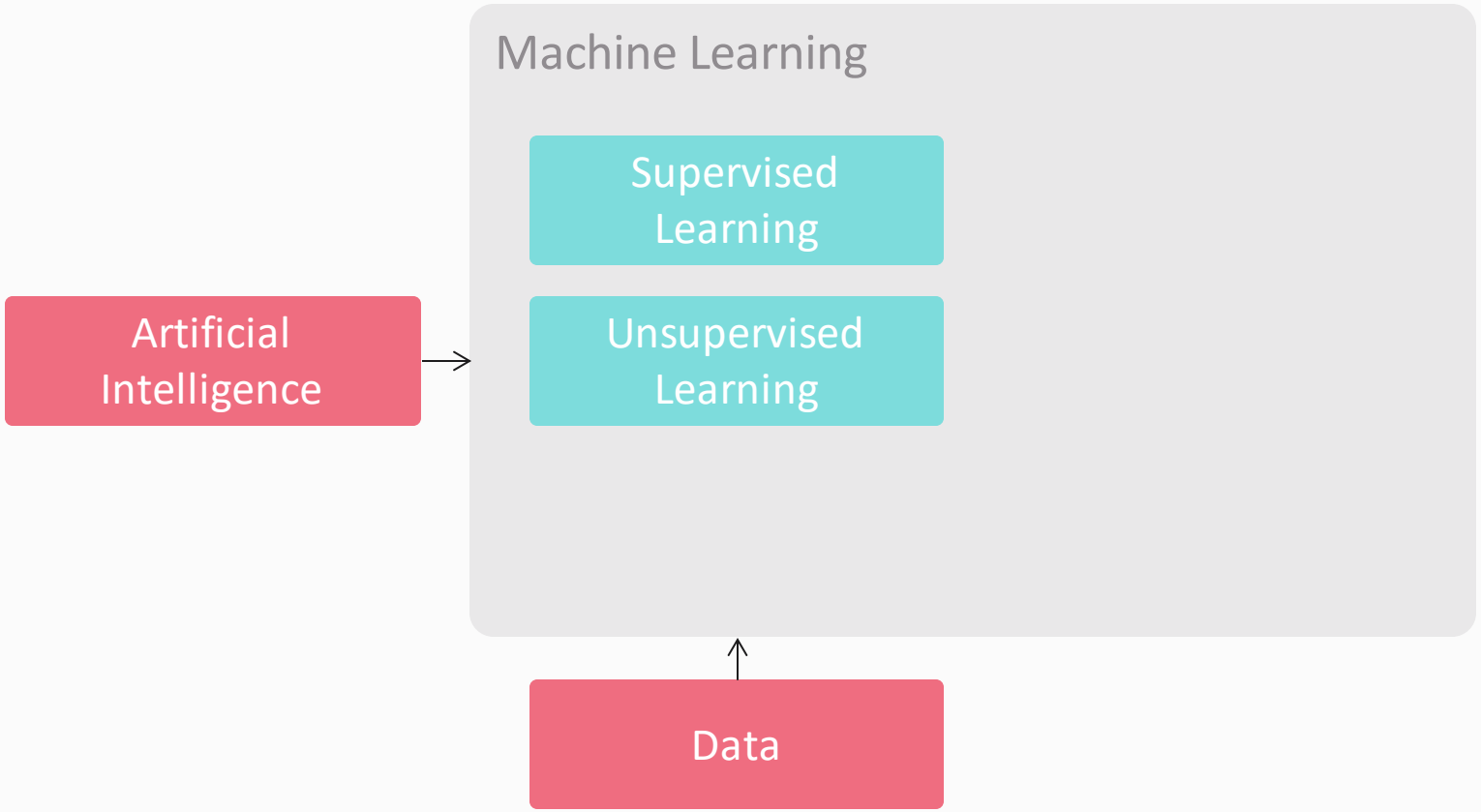
Facebook Prophet was used for day-level predictions, and a neural network in PyTorch distributed the visitors by hour. The solution was put into production in a Docker container in Azure.

Business value

The forecasts is a basis for marketing and personnel scheduling. With better planning, the Liseberg visitor will have a better experience.

Liseberg



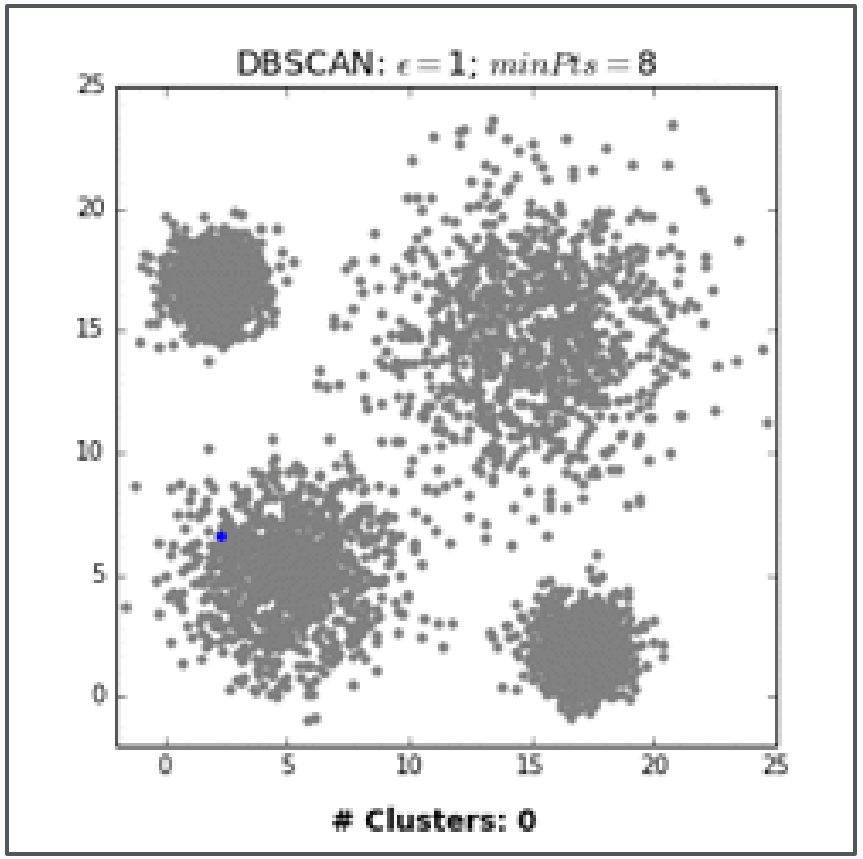


victor.backman@advectas.se

UNSUPERVISED LEARNING

Finds groups/clusters in the data.

Find anomalies (long distance from any group).



SUMMARY OF SURVEYS/REVIEWS

Challenge

Receives several thousands of survey responses every year. Analyzing these are very time-consuming and human bias might affect the result. Often, only a small selection of texts is read to try to form a holistic view.

Solution

Find recurring topics (groups of words that are often used together). This is a statistical approach which is independent of language. It's also possible to track topic magnitude over time to follow up actions.

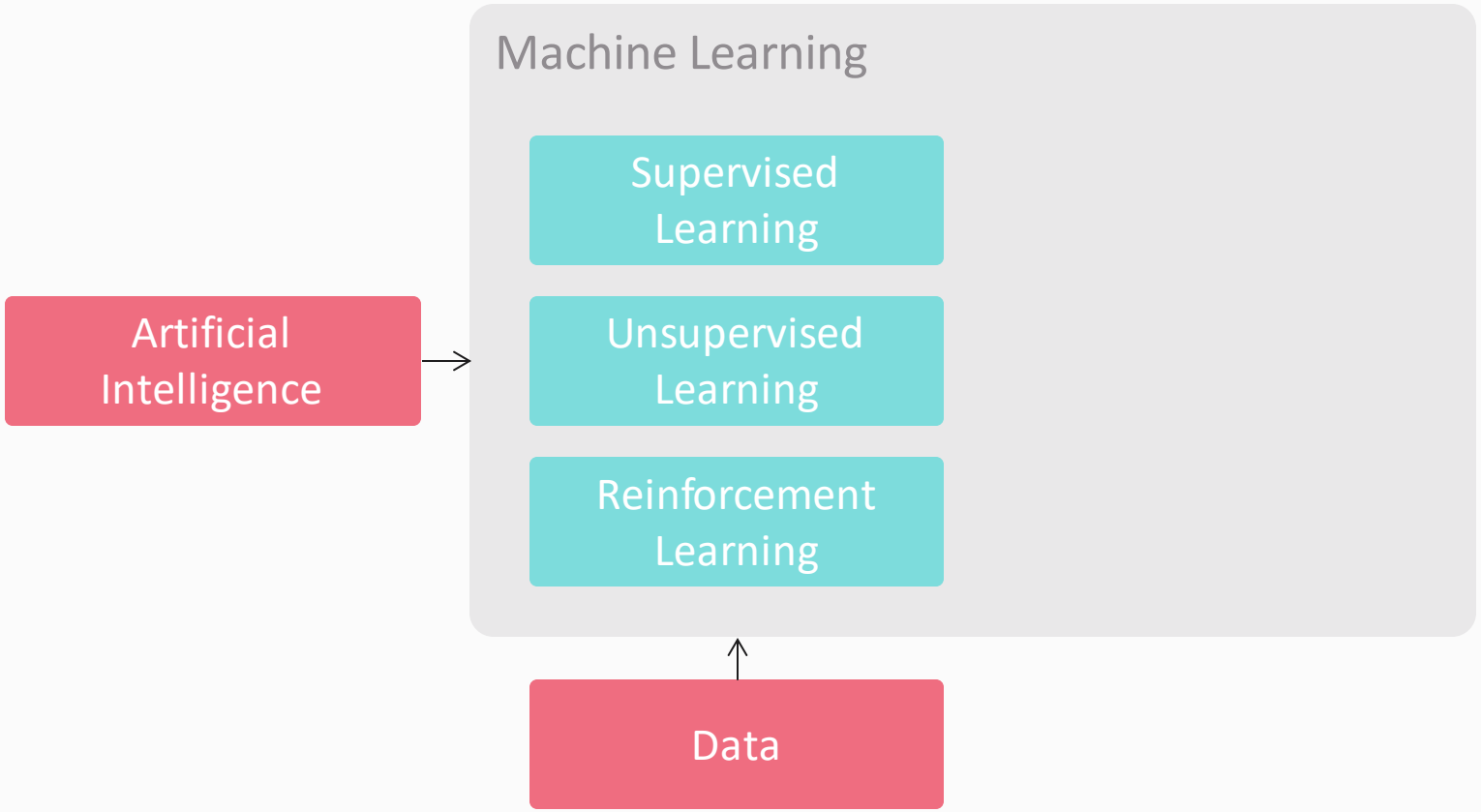
Tech

Topic modeling with LDA (Latent Dirichlet Allocation) and NMF (Non-negative matrix factorization) developed in Python. Visualized in several different BI tools as well as standalone HTML.

Business value

Impartial summary of texts. Big time savings. Ability to identify and act on insights that otherwise drown in large masses of text.

< Swedish region >

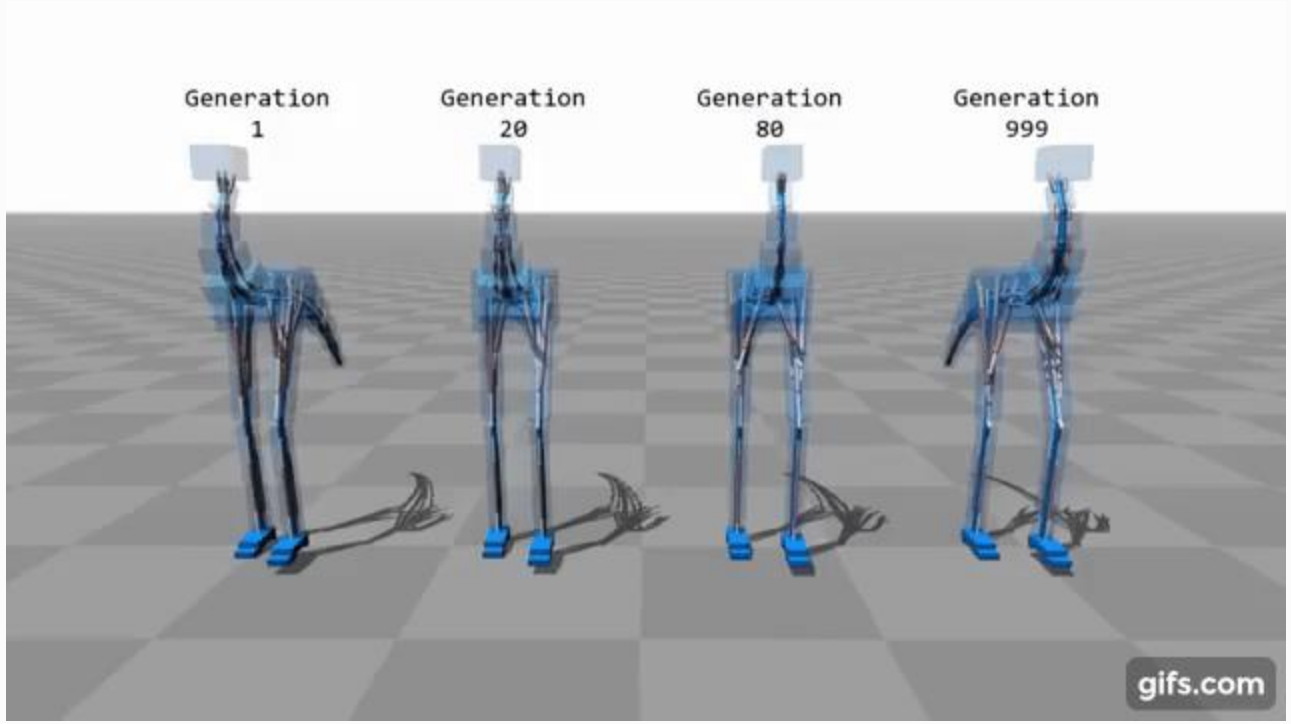


victor.backman@advectas.se

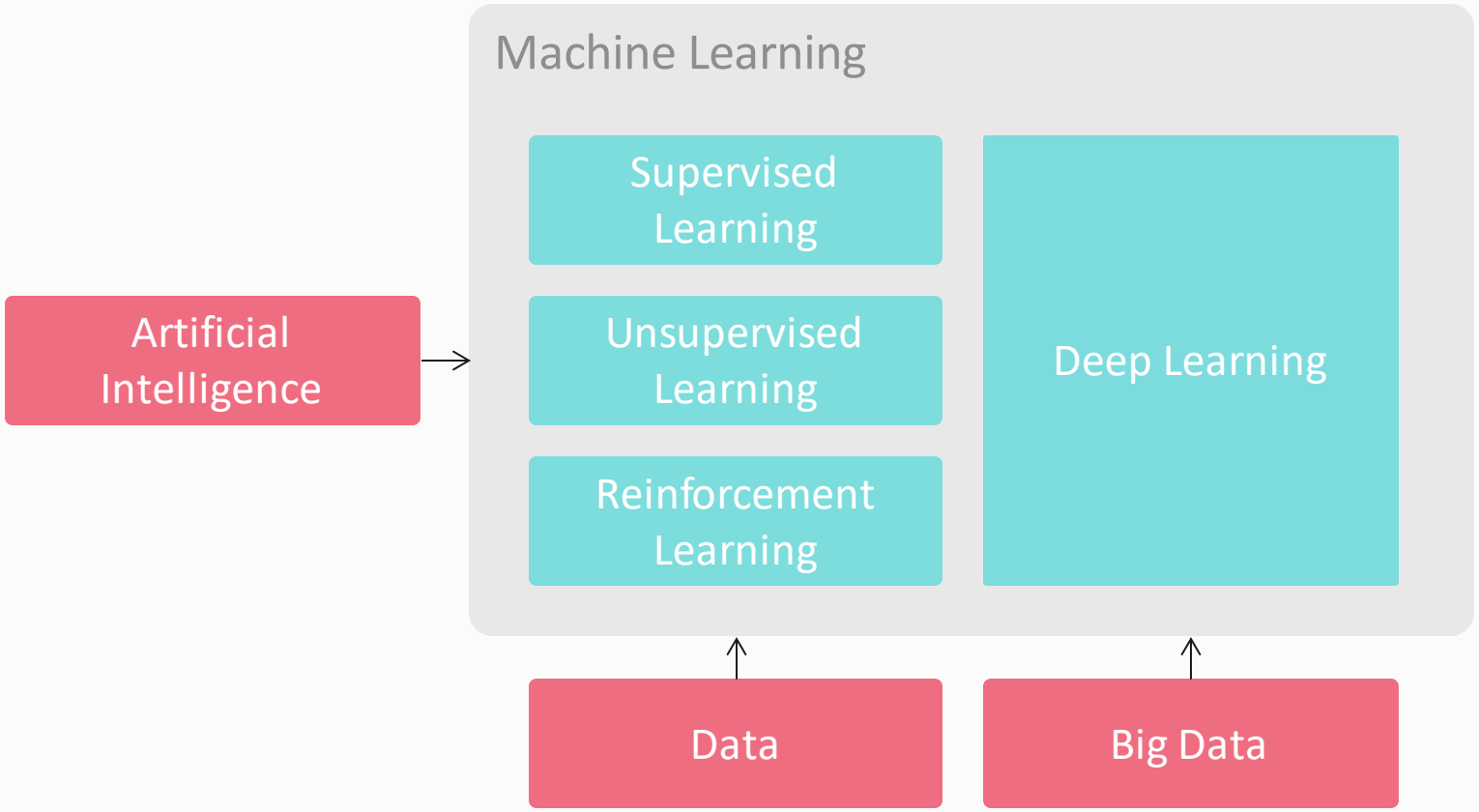
REINFORCEMENT LEARNING

Learn iteratively.

Define goal / punishment function.



victor.backman@advectas.se



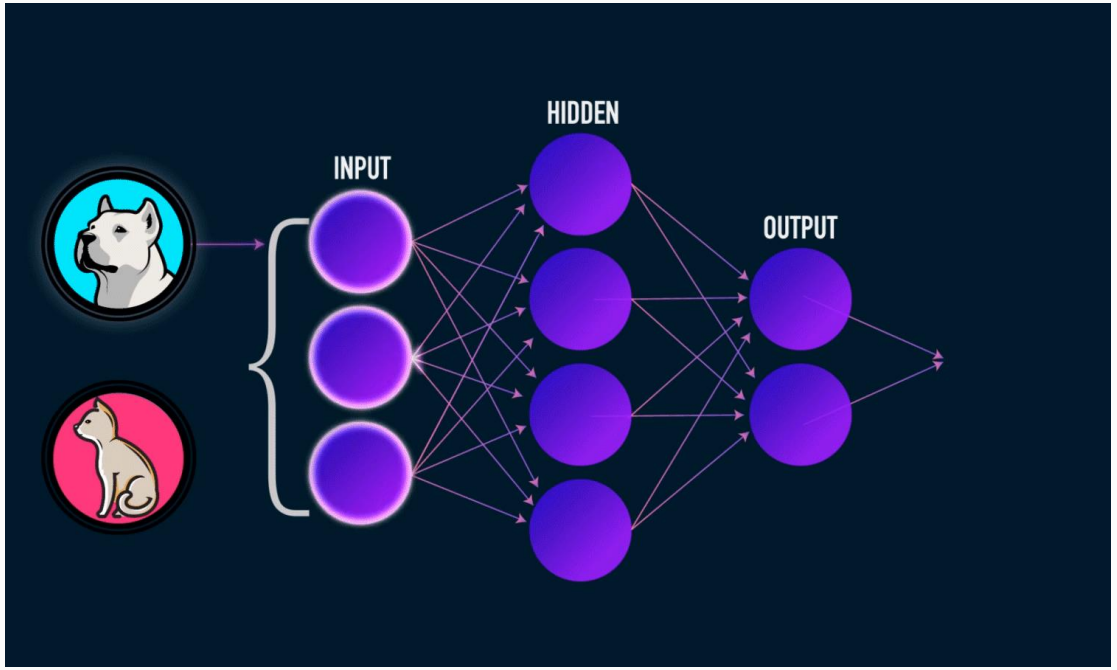
victor.backman@advectas.se

DEEP LEARNING

Complex model.

Can solve difficult problems.

Needs a lot of resources (data and computational power).



victor.backman@advectas.se

zalando

- Zalando has a research team consisting of 15 doctors in machine learning and are at the forefront of AI in clothing retail, releasing a lot of open-source tools.
- They represent every garment mathematically based on how they look (fashion-DNA), which makes up their analytics core.
- By using these representations of unstructured data, they can enhance their applications.

victor.backman@advectas.se

REPRESENTATION OF GARMENTS

zalando research

victor.backman@advectas.se

APPLICATION: STREET-TO-SHOP

zalando research

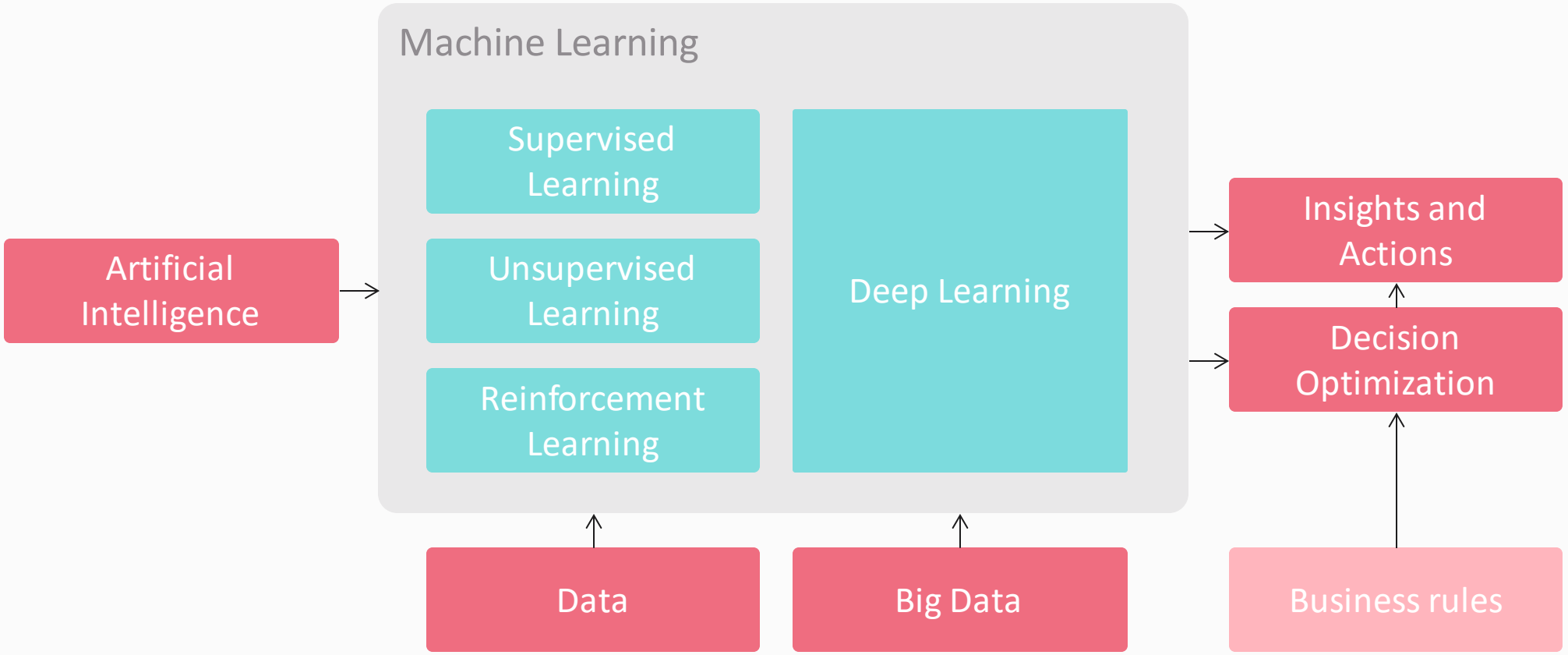
victor.backman@advectas.se

APPLICATION: STREET-TO-SHOP

zalando research

victor.backman@advectas.se

SUPERPERSONAL



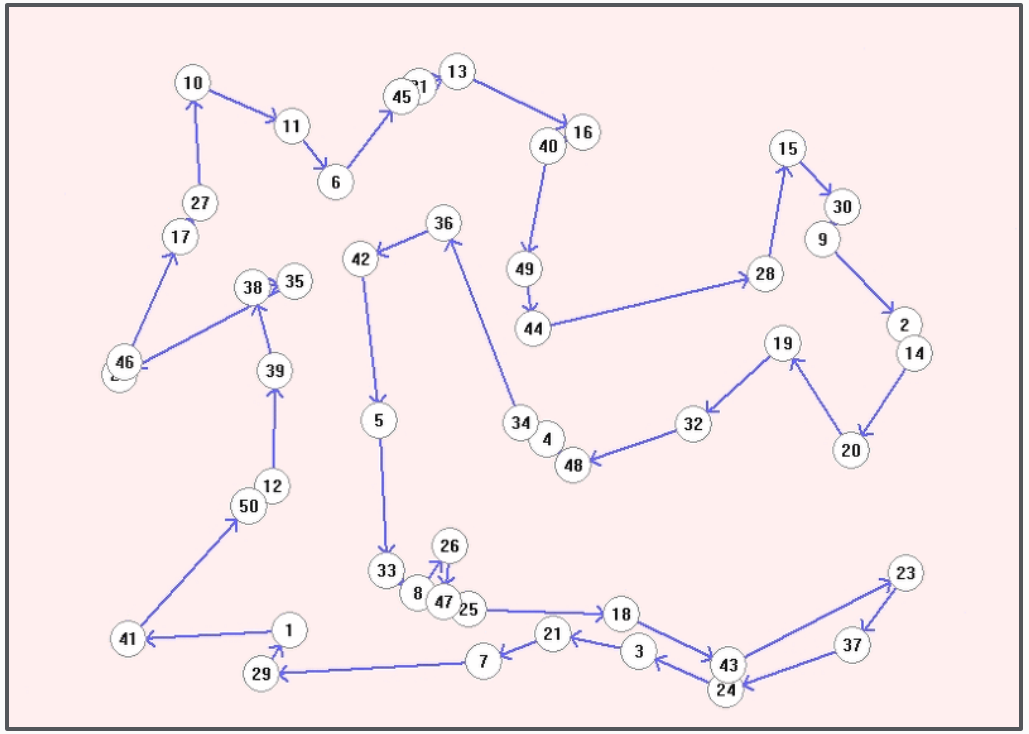
victor.backman@advectas.se

DECISION OPTIMIZATION

Calculate optimal decisions

Does not learn from training data. Optimization is done based on **constraints**.

Machine learning and predictions can help in defining those constraints (e.g the predicted demand).



OPTIMIZE SHIFT SCHEDULING

Challenge

24/7 operations and other staff-intensive operations have a great need to put in place reliable and resource efficient shift schedules. With tools like Excel, the process is time consuming and the risk is great for mistakes and sub-optimality.

Solution

With decision optimization, the process can be streamlined. Together with machine learning, demand can first be forecasted and the schedule optimized automatically thereafter.

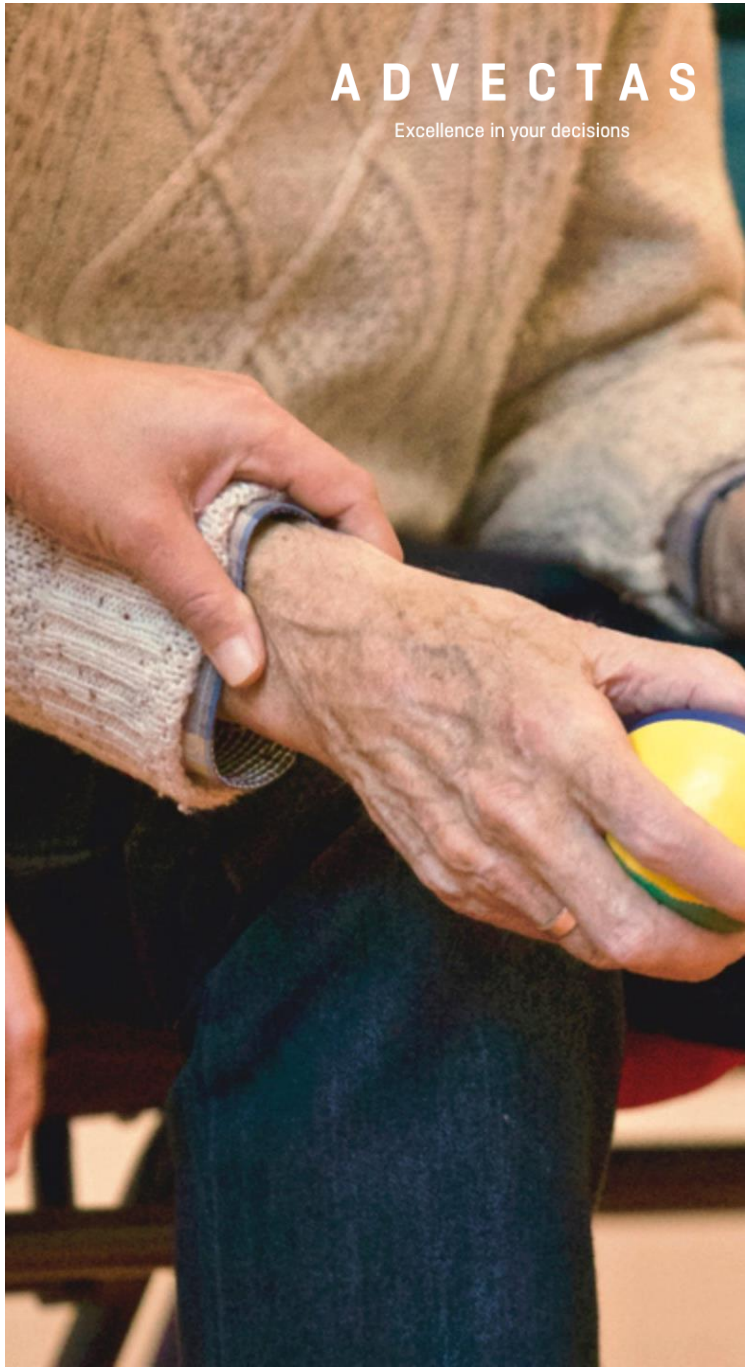
Tech

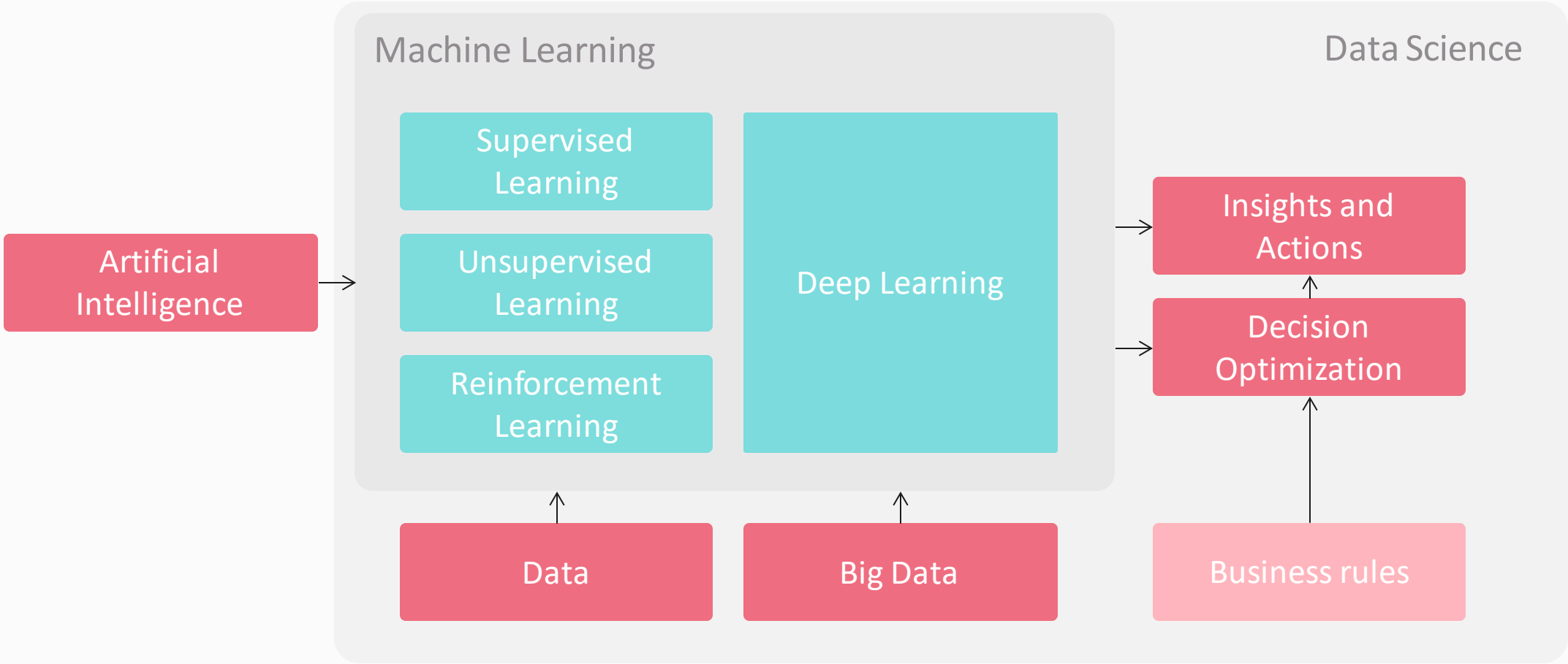
Developed in Python with optimization engines like CPLEX.

Business value

An automated scheduling process can save countless hours of administrative time. An optimal schedule also ensures an efficient use of resources.

< Swedish municipality >





victor.backman@advectas.se

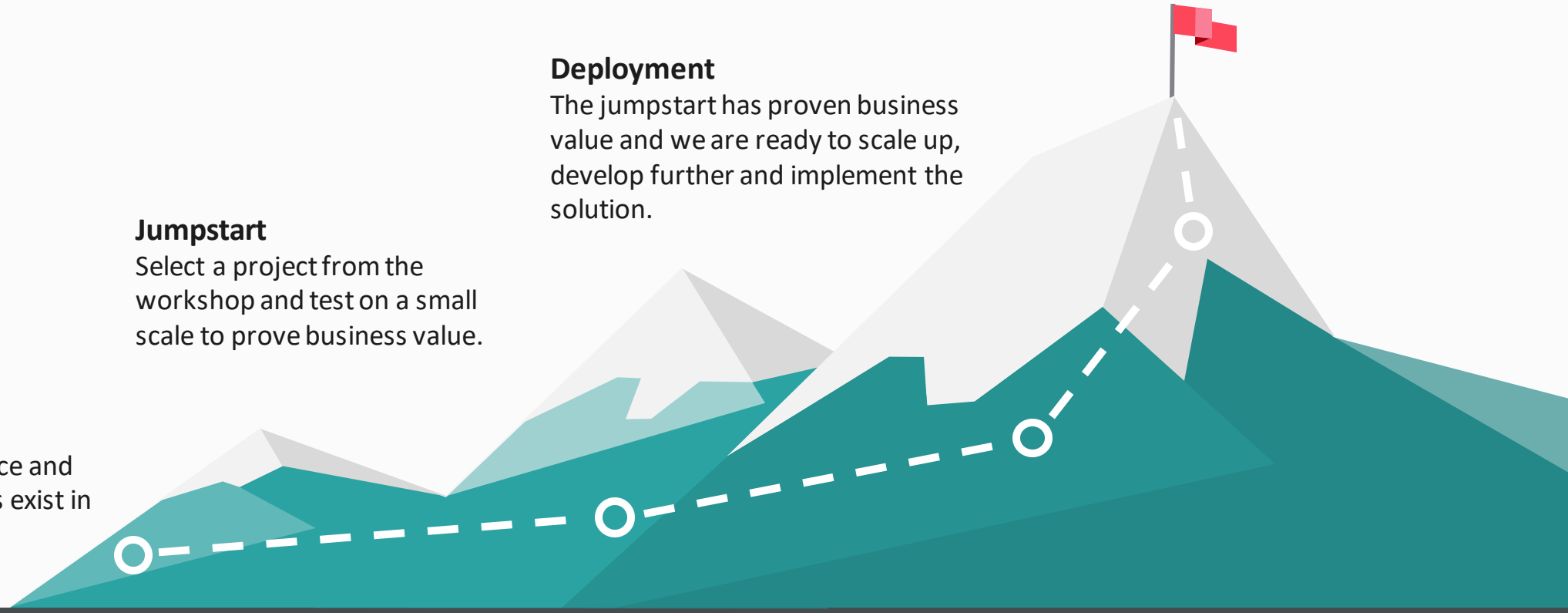
THE DATA SCIENCE JOURNEY

Enrich business process
Integrate the result into the existing process and educate the business on how it should be consumed.

Deployment
The jumpstart has proven business value and we are ready to scale up, develop further and implement the solution.

Jumpstart
Select a project from the workshop and test on a small scale to prove business value.

Data Science Workshop
What is Data Science and what opportunities exist in your business?



KEY TAKEAWAYS

- The possibilities with AI in marketing/sales are prosperous
- AI is a combination of machine learning models
- Machine learning only solves a specific problem
- You need to apply some business logic to your predictions
- Get started – and fail fast!

victor.backman@advectas.se

THANK YOU!

Victor Bäckman
0732-61 63 50
victor.backman@advectas.se

ADVECTAS

Excellence in your decisions