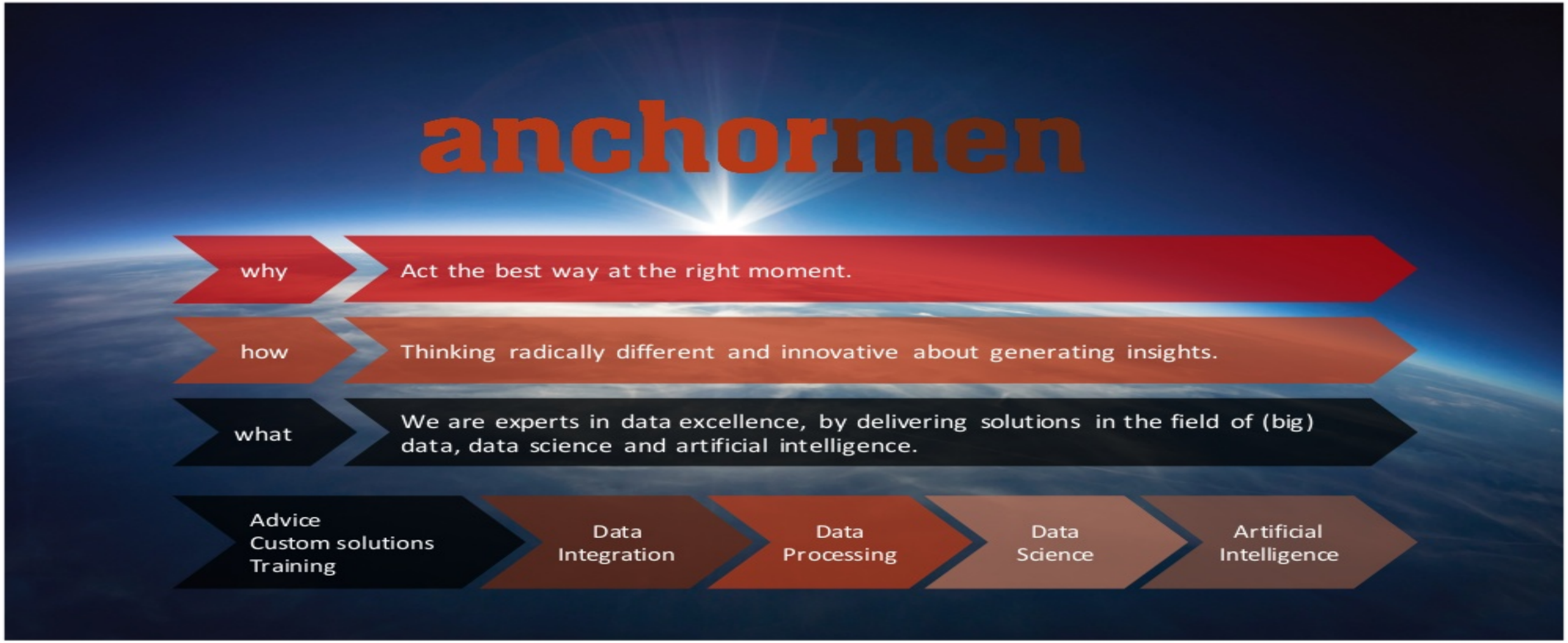

Making the switch
SparkSummit 2016
anchormen
expert in data excellence


anchormen
why
Act the best way at the right moment.
how
Thinking radically different and innovative about generating insights.
what
We are experts in data excellence, by delivering solutions in the field of (big) data, data science and artificial intelligence.
Advice
Custom solutions
Training
Data Integration
Data Processing
Data Science
Artificial Intelligence

Anchormen

anchormen

- We specialize in data excellence:
  - Consumer 360
  - 24/7 Business
  - Search, Match & Find
- **anchormen.nl/careers**

About us

anchormen

Jeroen Vlek

- Lead data engineer
- Struggling with Bloodborne (PS4)

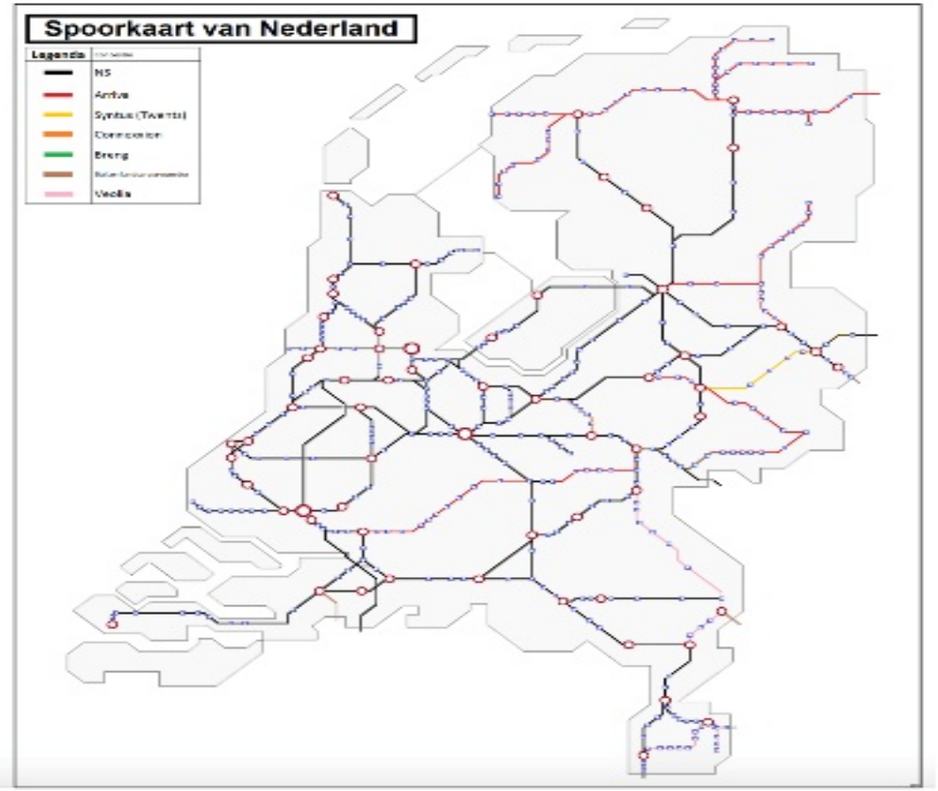
Chris Pool

- Data scientist
- Struggling with diapers

Dutch railways

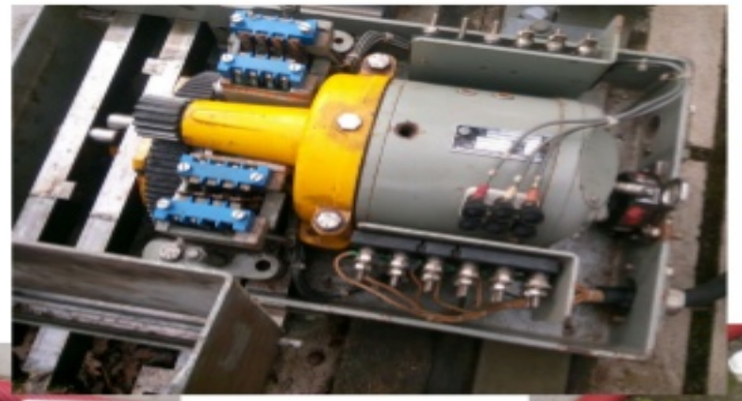
anchormen

- Most used network in Europe
- 3,3 million journeys
- 1.157.260 daily travellers



What does Strukton Rail do?

anchormen



Strukton
Rail

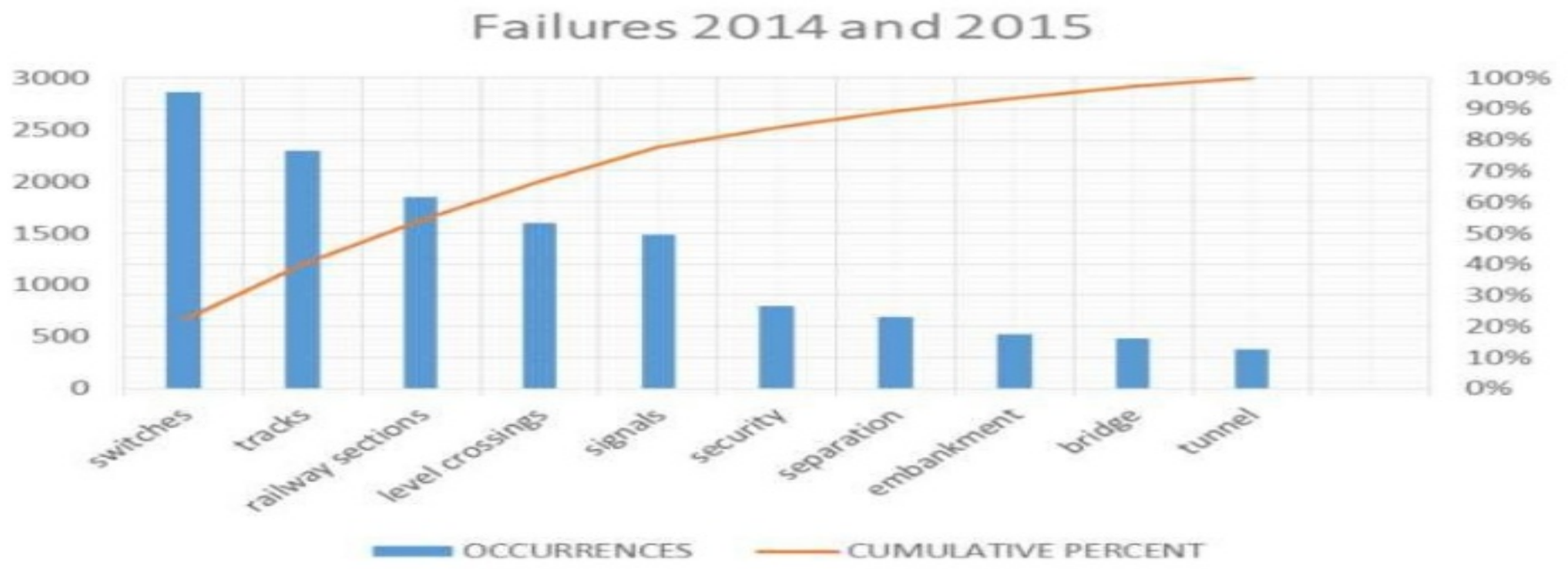
Predictive Maintenance @ Strukton

anchormen

- Less delays and canceling of trains
- Making Strukton the leading company in the field of rail maintenance
- Cost reduction
- Better preparation for repair personnel

Switch Failures

anchormen



Switch Failure Causus

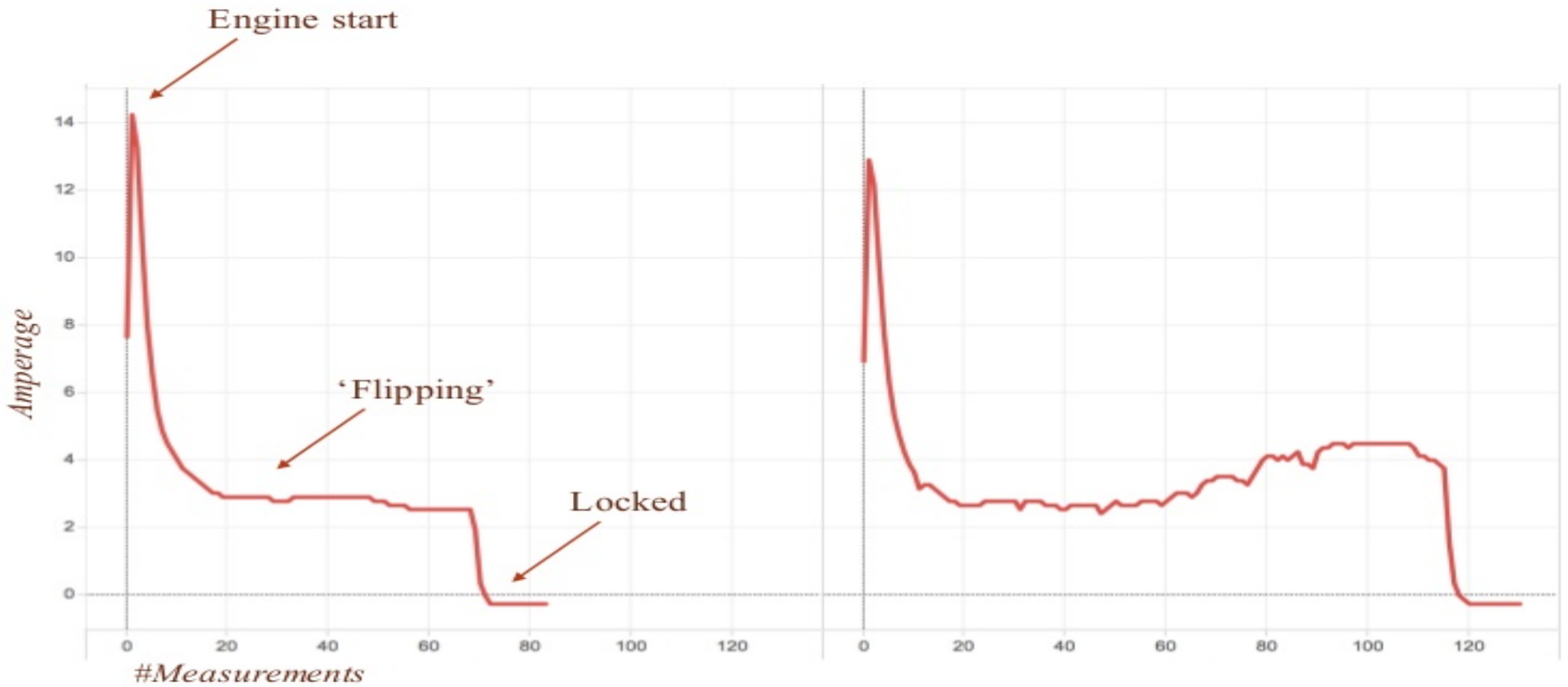
anchormen

Frequently obstructed movements due to:

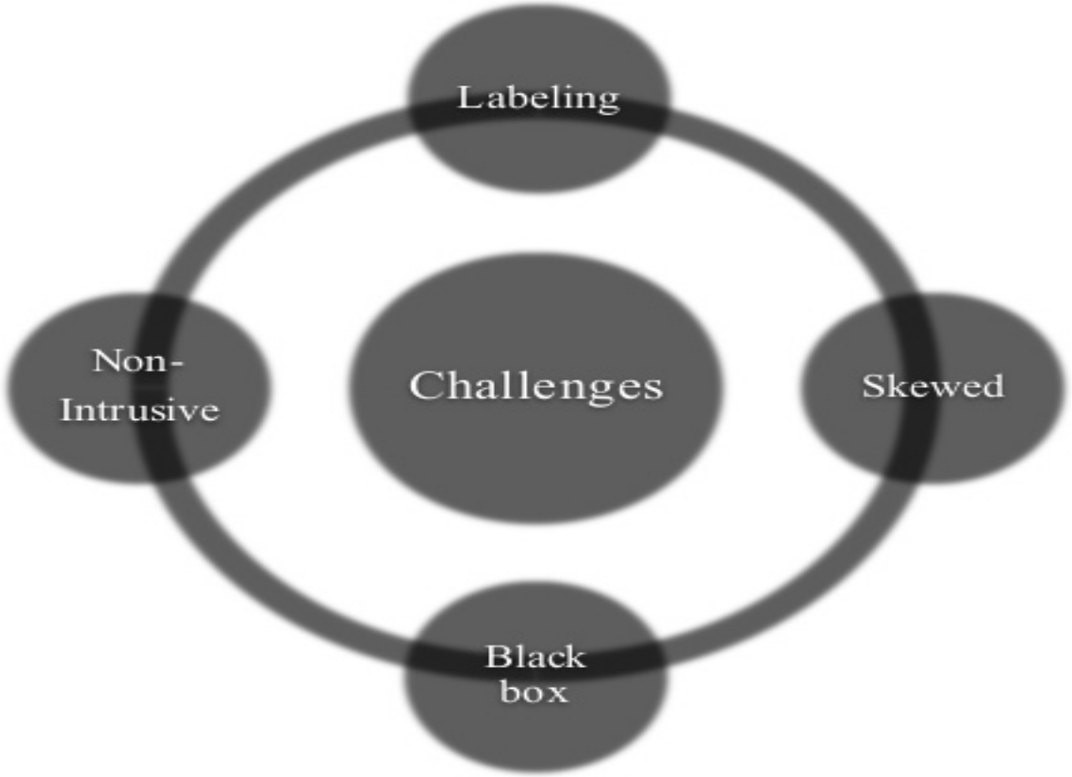
- Poor adjustment of rolling construction
- Lack of grease on slide chairs
- Bent blades
- Electrical problems (worn-out brushes, motor, etc.)

Goal: Predict switch failure

anchormen



anchormen



Problem Definition

anchormen

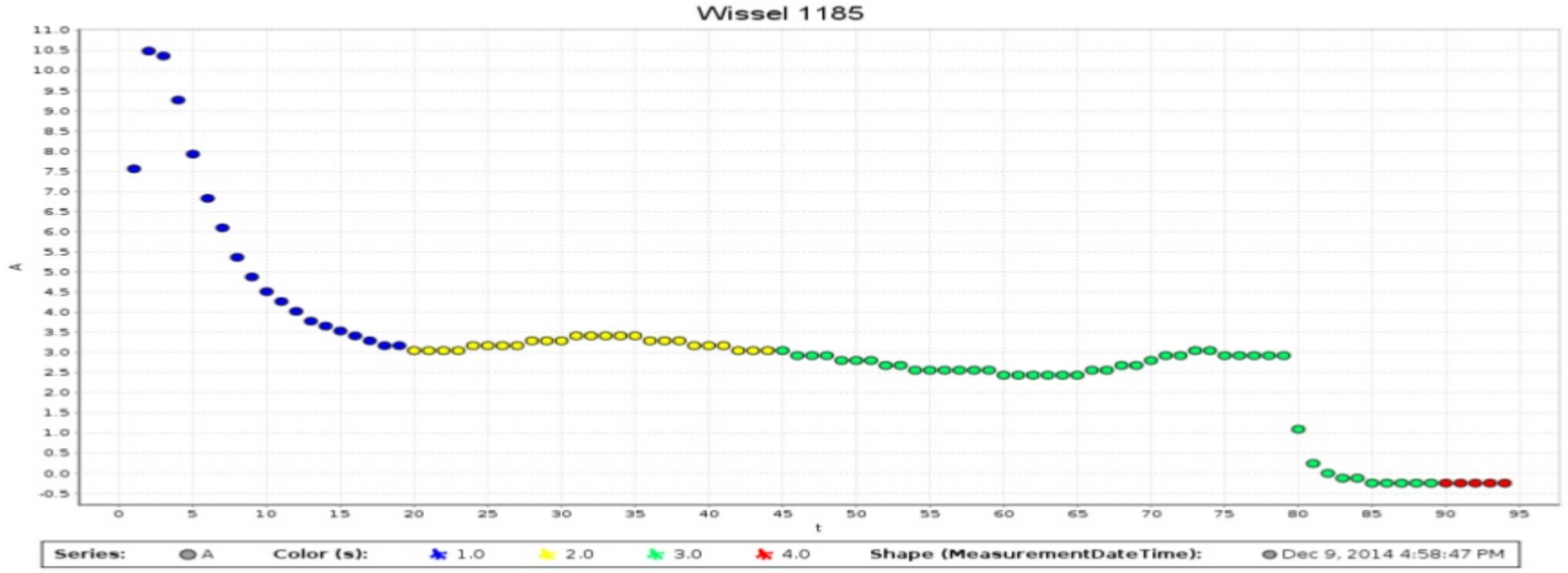
Learn the deviations in the data that indicate an upcoming malfunction

Data

anchormen

Segments

anchormen



Derived features

anchormen

- Features that represent the curve (per segment):
  - Min
  - Max
  - Average
  - Length
  - Difference compared to previous flip
- Features for entire flip
  - Days since last failure
  - Temperature

Normalization and Aggregation

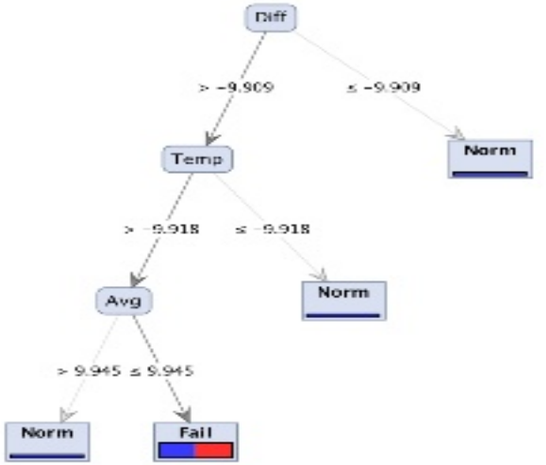
anchormen

- Normalize data using sliding window
- Aggregate per day
  - Min
  - Max
  - First
  - Last
  - Variance
  - Average
  - Count

Model

anchormen

- Decision tree: Will it break within the next 3 weeks or not?
- Strukton: “keep it simple and explainable”
- From days until failure to classes
  - 0-2 days
  - 2-7 days
  - **7-21 days**
  - 21-55 days
  - >55 days



Architecture (current)

anchormen

Why Spark?

anchormen

- Lots of data prep and feature computation
- More switches to be added in the future
- Streaming scenarios:
  - Short term failures
  - Optimize personnel's routes

Results

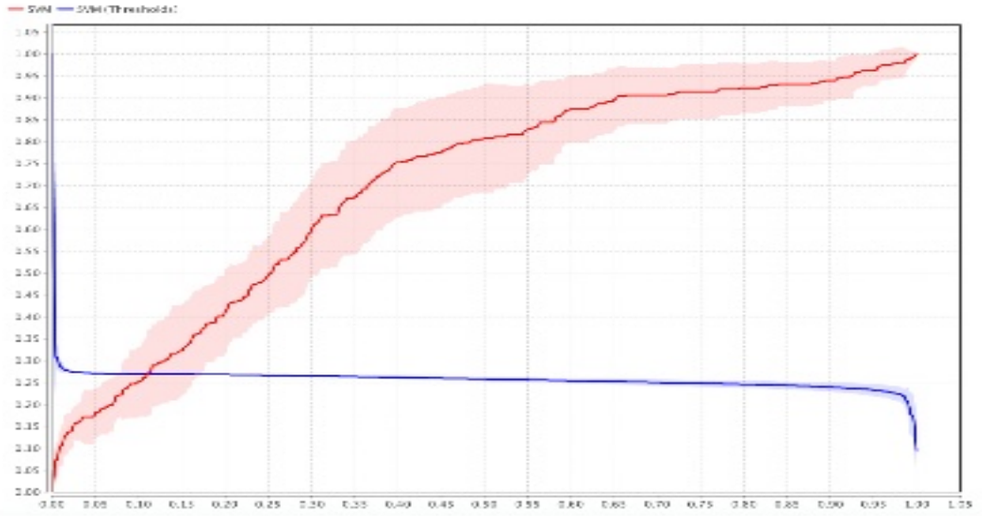
anchormen

	True negative	True positive	class precision
Predicted negative	798	23	97.20%
Predicted positive	1	64	98.46%
class recall	99.87%	73.56%	

Precision vs Recall

anchormen

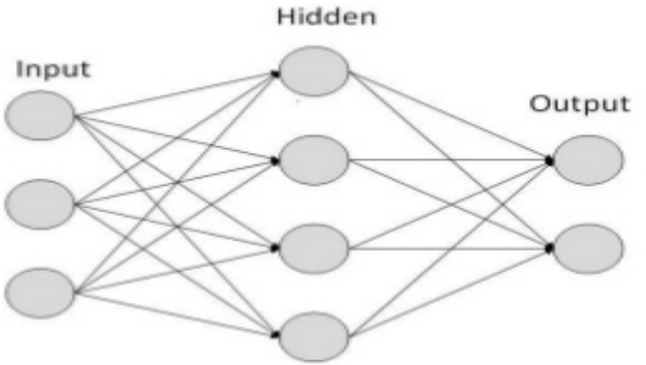
- Precision and recall are easily explained
- Sending a mechanic is cheaper than a fine
- Recall is more important



Future work

anchormen

- Deep learning
- Predict the number of days (regression)
- Predict type of failures
  - Less voltage
  - Too disorderly
  - Not locking: Too frequent
  - Up/down movement



Next steps

anchormen

- Production
- Lambda architecture
- Nation wide roll out

Questions?

anchormen

- info@anchormen.nl
- www.anchormen.nl
- @anchormenBDS


1861
anchormen
Amsterdam
Panamalaan 2K
1019 AZ Amsterdam
Groningen
Rozenburglaan 3
9727 DL Groningen
info@anchormen.nl
www.anchormen.nl