© 2018 SPLUNK INC.

Machine Learning and IoT Use Cases

Nikolas Kourtidis | Senior Partner Sales Engineer
Philipp Drieger | Staff Machine Learning Architect
niko@splunk.com and philipp@splunk.com

October 2018

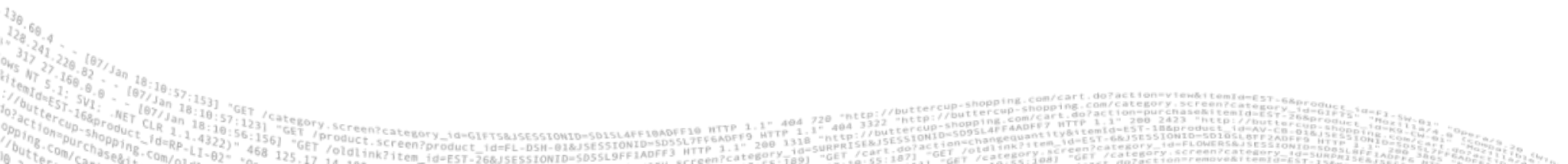
© 2018 SPLUNK INC.

Forward-Looking Statements

During the course of this presentation, we may make forward-looking statements regarding future events or the expected performance of the company. We caution you that such statements reflect our current expectations and estimates based on factors currently known to us and that actual events or results could differ materially. For important factors that may cause actual results to differ from those contained in our forward-looking statements, please review our filings with the SEC.

The forward-looking statements made in this presentation are being made as of the time and date of its live presentation. If reviewed after its live presentation, this presentation may not contain current or accurate information. We do not assume any obligation to update any forward-looking statements we may make. In addition, any information about our roadmap outlines our general product direction and is subject to change at any time without notice. It is for informational purposes only and shall not be incorporated into any contract or other commitment. Splunk undertakes no obligation either to develop the features or functionality described or to include any such feature or functionality in a future release.

Splunk, Splunk>, Listen to Your Data, The Engine for Machine Data, Splunk Cloud, Splunk Light and SPL are trademarks and registered trademarks of Splunk Inc. in the United States and other countries. All other brand names, product names, or trademarks belong to their respective owners. © 2018 Splunk Inc. All rights reserved.



© 2018 SPLUNK INC.

About us

PHILIPP DRIEGER

Staff Machine Learning Architect

philipp@splunk.com

NIKO KOURTIDES

Senior Partner Sales Engineer

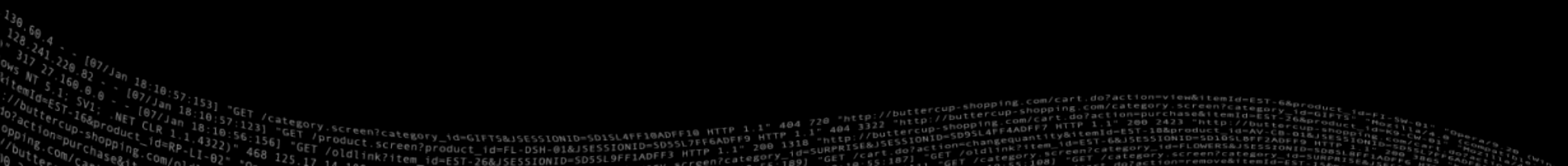
niko@splunk.com

splunk> .conf18

© 2018 SPLUNK INC.

Agenda

- IoT and Machine Learning Use Cases Overview
- Quick Intro in Machine Learning and MLTK Demo
- Customer Success Stories
  - DB Cargo Predictive Maintenance
  - BMW Car Sharing Demand Prediction
  - Continental Industrial Production Optimization
  - Zeppelin Power Systems Predictive Maintenance
    - Deep dive into project (data onboarding and ML)
- Wrap up



© 2018 SPLUNK INC.

IoT and Machine Learning Use Cases

splunk> .conf18

© 2018 SPLUNK INC.

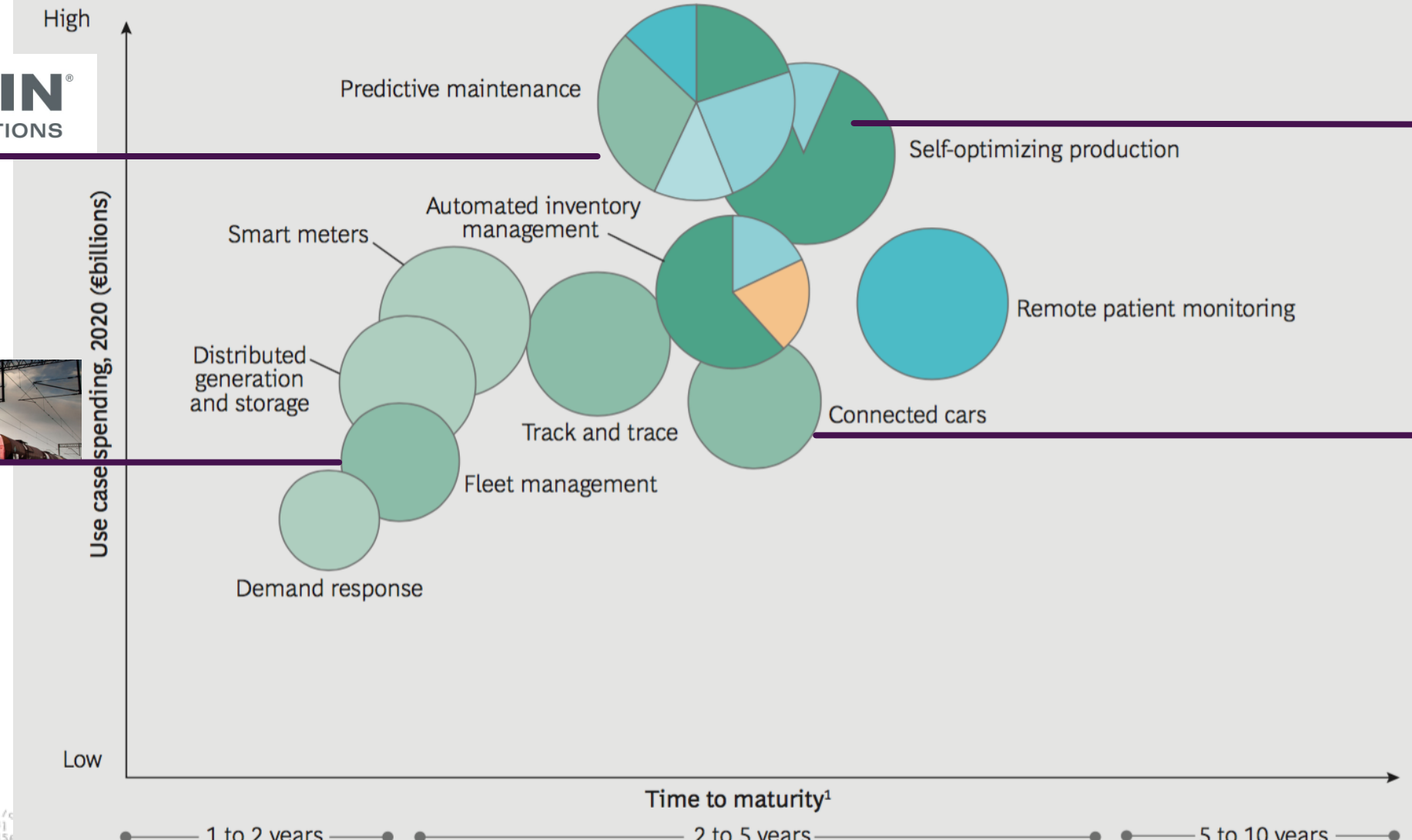
Think IoT from Use Cases

ZEPPELIN
WE CREATE SOLUTIONS

Continental
The Future in Motion

BMW

EXHIBIT 2 | Ten Use Cases Will Drive IoT Growth Through 2020



unk> .conf18

© 2018 SPLUNK INC.

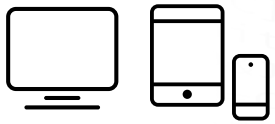
Operationalize Machine Learning

splunk> .conf18

© 2018 SPLUNK INC.

Continuous Data Ingest at Scale

Industrial Assets



Consumer and Mobile Devices

OT

IT

Industrial Data
SCADA, AMI, Meter Reads

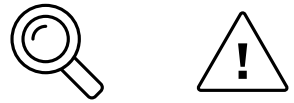
Native Inputs
TCP, UDP, Logs, Scripts, Wire, Mobile

Modular Inputs
MQTT, AMQP, COAP, REST, JMS

HTTP Event Collector
Token Authenticated Events

Technology Partnerships
Kepware, AWS IoT, Cisco, Palo Alto

Engineers | Data Analysts | Security Analysts | Business Users



Search | Alert | Visualize | Predict | Develop

Real Time

splunk>enterprise splunk>cloud

External Lookups/Enrichment

Asset Info | Maintenance Info | Data Stores

splunk> .conf18

© 2018 SPLUNK INC.

Sense and Respond

Every Search Can Use Machine Learning

Industrial Assets

Consumer and Mobile Devices

OT

IT

Real Time

splunk>

Search

Alert

Flash lights

Email

Send an email

Tickets

File a ticket

Third-Party Applications

Trigger process flow

Smartphones and Devices

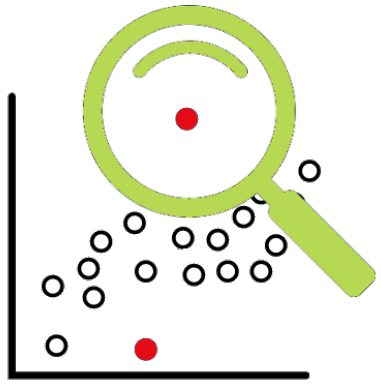
Send a text

splunk> .conf18

© 2018 SPLUNK INC.

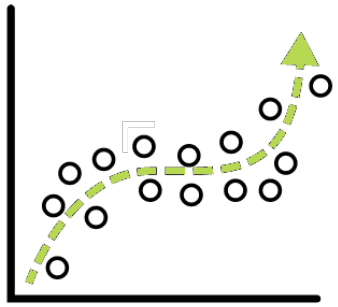
Splunk Customers Want Answers from their Data

Anomaly detection



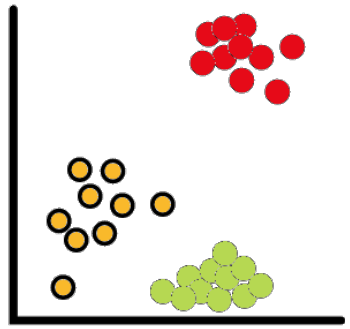
- Deviation from past behavior
- Deviation from peers
- (aka Multivariate AD or Cohesive AD)
- Unusual change in features
- **ITSI MAD Anomaly Detection**

Predictive Analytics



- Predict Service Health Score
- Predicting Events
- Trend Forecasting
- Detecting influencing entities
- Early warning of failure – predictive maintenance

Clustering

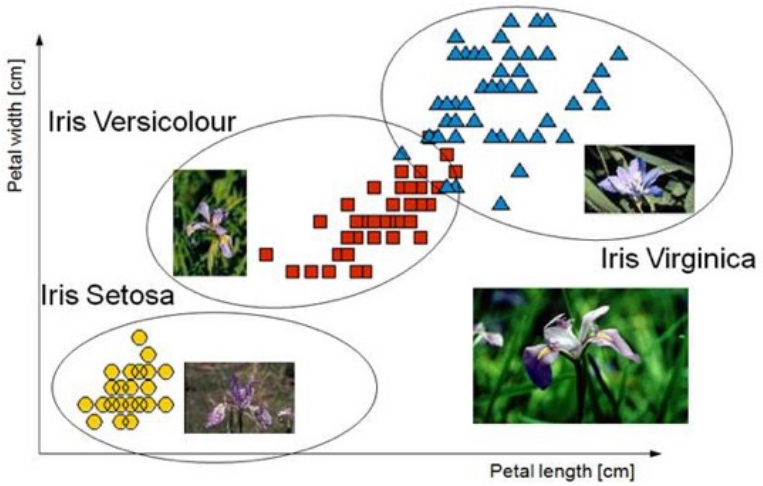


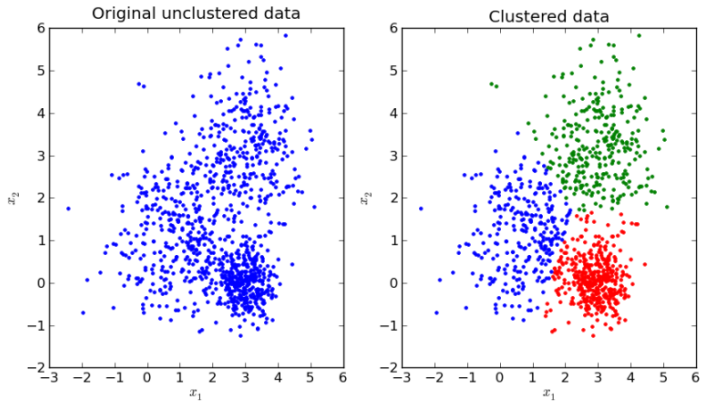
- Identify peer groups
- Event Correlation
- Reduce alert noise
- Behavioral Analytics
- **ITSI Event Analytics**

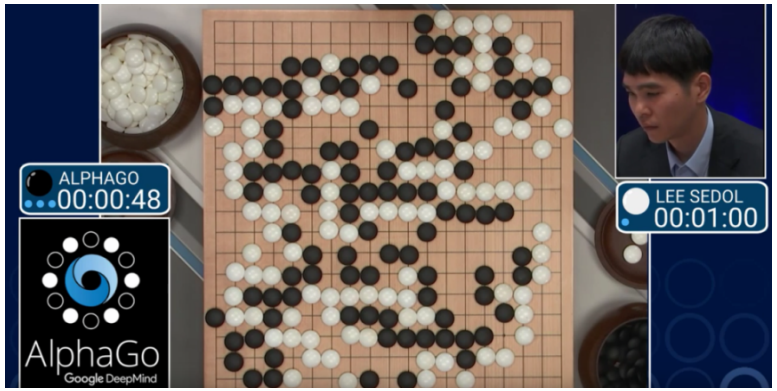
© 2018 SPLUNK INC.

Types of Machine Learning

- Supervised Learning (labeled data)
  - regression
  - classification
- Unsupervised Learning (unlabled data)
  - clustering
  - anomaly detection
- Mixed Models (with Reinforcement or Feedback)
  - human in the loop
  - autonomous systems







© 2018 SPLUNK INC.

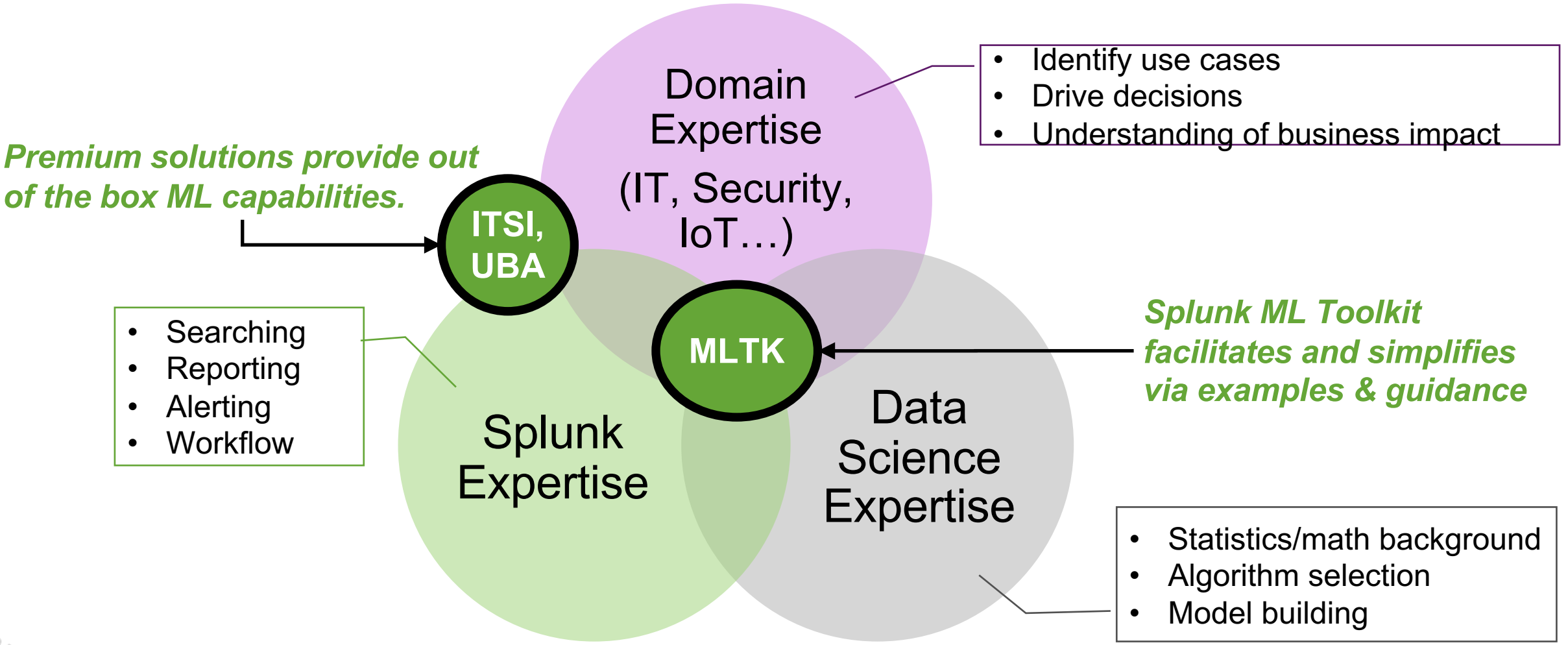
Types of Machine Learning

In context of IoT

- Supervised Learning
  - Build a model on sensor data with labels from known outages
  - Build regression models that reflect “normal” behavior and detect strong residuals as anomalies
- Unsupervised Learning (unlabled data)
  - Build a model to cluster process steps
  - Generate labels with clustering for supervised learning
  - Detect strong signal deviations as anomalies
- Mixed Models (with Reinforcement or Feedback)
  - Present model predictions to domain experts and incorporate feedback into next model trainings
  - Stack multiple models and apply 2nd order ML

© 2018 SPLUNK INC.

Skill Areas for Machine Learning

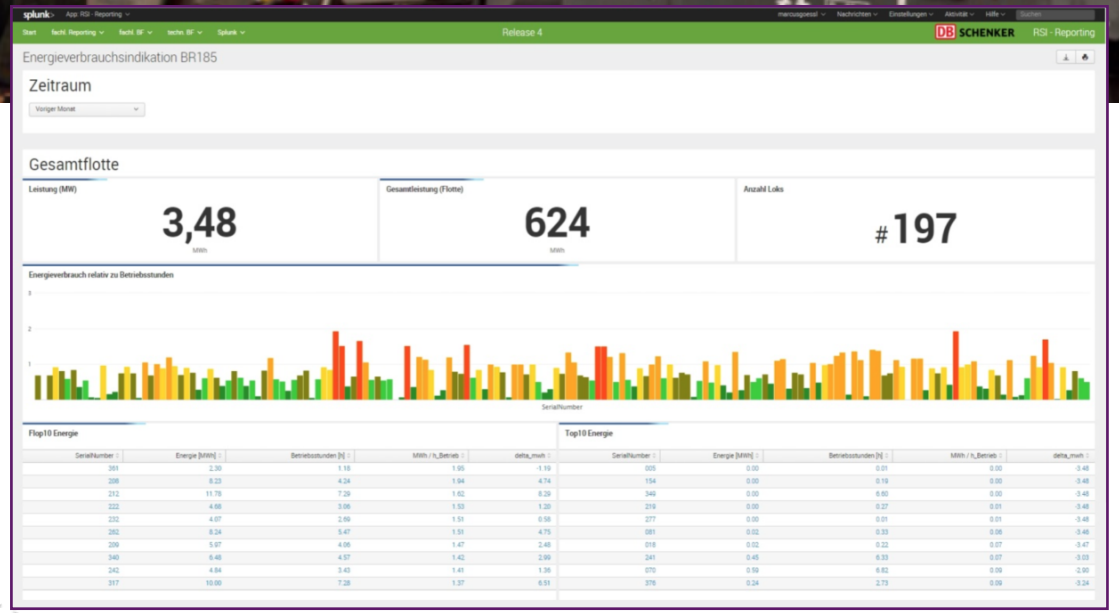


© 2018 SPLUNK INC.

Customer Use Cases

https://www.splunk.com/blog/2018/04/25/operational-intelligence-manufactured-in-germany-splunklive-2018-events-in-germany.html

Rolling Stock Intelligence



- Find and fix errors fast
- 2000 trains connected in 2019 over 16 types
- Root cause analysis, Operational Intelligence
- Fleet Management, Predictive Maintenance

http://conf.splunk.com/files/2016/slides/internet-of-big-rolling-things-at-db-cargos-european-rolling-stock-increased-customer-satisfaction-through-higher-availability-and-reliability.pdf

Saturday

00:00

BMW car sharing

- Use MLTK to predict car sharing demand forecast
- Reflect seasonal and daily trends

MLTK

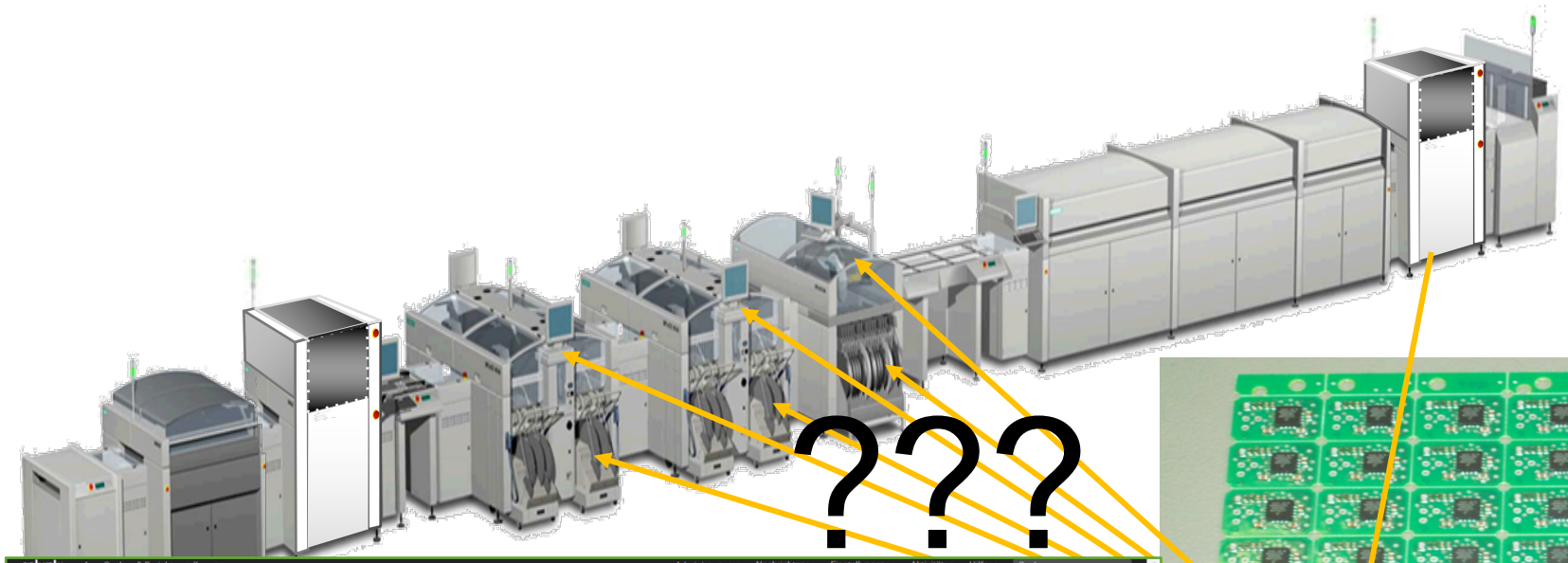
- Non linear algorithm: RandomForest
- Input : lat / lon / time
- Output : car demand

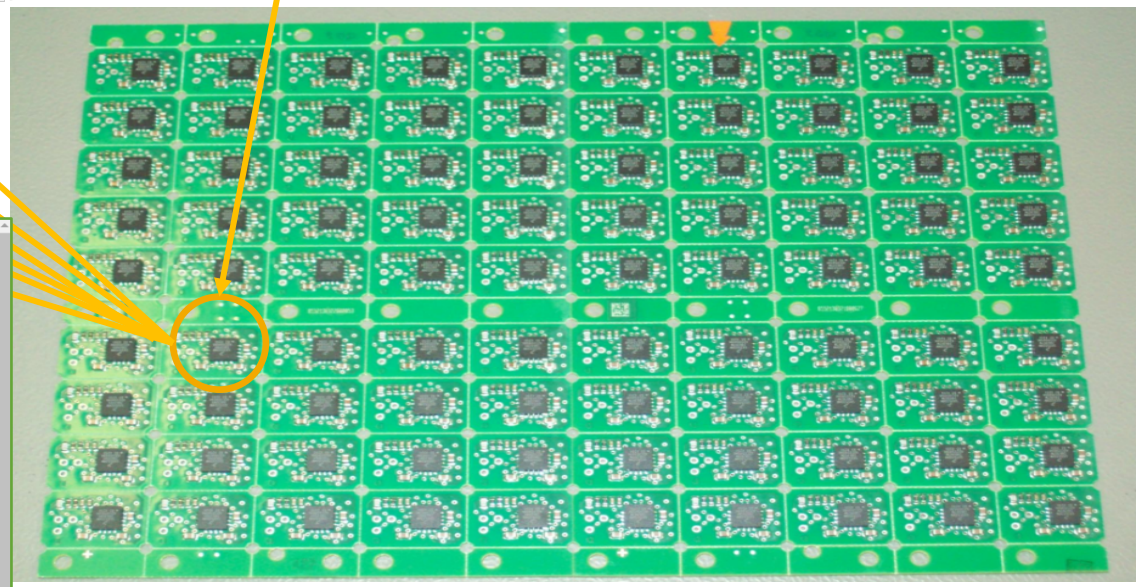
https://www.slideshare.net/Splunk/splunklive-munich-2018-car-sharing-fleet-demand-bmw

© 2018 SPLUNK INC.

Continental
The Future in Motion

SMD Production Line Optimization





splunk> App: Suchen & Berichterstellung — Administrator — Nachrichten — Einstellungen — Aktivität — Hilfe — Suchen

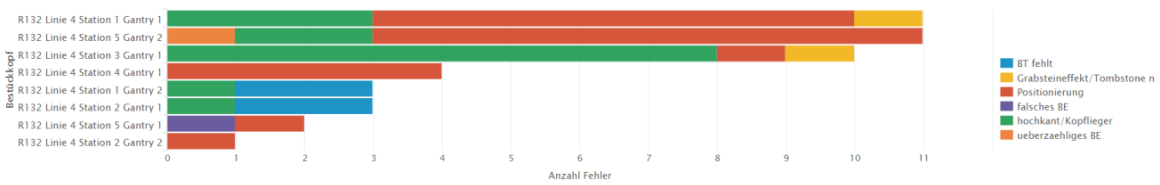
Suche | Datenmengen | Berichte | Benachrichtigungen | Dashboards — Suchen & Berichterstellung

Bestückkopf-Cockpit

Bearbeiten | Exportieren | ...

Zeitauswahl: Letzte 30 Tage — Filter ausblenden

Bestückköpfe mit den höchsten Fehlerbeteiligungen



R132 Linie 4 Station 1 Gantry 1: fehlerrelevante Segmente

Segment	Produkt	Seite	Bauteilposition	Fehler	Anzahl	Zeitpunkte
2	A2C82317700	B	R4017-1	Positionierung	3	06:03:37 04/11/16 05:54:09 04/11/16 05:40:49 04/11/16
5	A2C99042901	A	D9209-2	Positionierung	3	22:42:30 10/11/16 22:42:22 10/11/16 22:41:32 10/11/16
11	A2C99042901	A	C3050-2	hochkant/Kopflieger	1	12:21:18 03/11/16
17	A2C83800700	A	R1005-8	hochkant/Kopflieger	1	16:12:27 05/11/16
17	A2C83800700	A	R402-12	Grabsteineffekt/Tombstone n	1	08:50:19 05/11/16
19	A2C99042901	A	C9223-1	hochkant/Kopflieger	1	12:20:52 03/11/16
8	A2C99042901	B	T4412-2	Positionierung	1	05:58:34 03/11/16

allg. Fehlerhäufigkeit pro Bauteil/Produkt in den letzten 7 Tagen

Produkt	Seite	Bauteilposition	Fehler	Fehleranzahl
A2C82317900	B	D4500-1	hochkant/Kopflieger	5
A2C99042901	A	D9209-2	Positionierung	3
A2C82317900	B	D4500-2	hochkant/Kopflieger	3
A2C82317700	B	R4017-1	Positionierung	3
A2C99043002	B	D4017-1	Positionierung	2
Weitere Fehlerfälle	Weitere Fehlerfälle	Weitere Fehlerfälle	Weitere Fehlerfälle	29

https://www.slideshare.net/Splunk/splunklive-frankfurt-2017-continental

splunk> .conf18

© 2018 SPLUNK INC.

ZEPPELIN®
WE CREATE SOLUTIONS

Predictive Maintenance in Power Plants

- **Goal: Ensure uptime and availability to 100 % for 70 power plants**
- **Reactive to proactive mode with Machine Learning**

https://www.slideshare.net/Splunk/splunklive-munich-2018-use-caseztndkerze-zeppelin

splunk> .conf18

© 2018 SPLUNK INC.

ZEPPELIN®
WE CREATE SOLUTIONS

How raw data looked like

```
[1=WinAC.AVHistory,155]
PL1000-40313;Motorlast;%;0;0,1;0;0;0;1;0;0;0;0;0;0;
0000-0021:03D4 03DE 03DE 03E8 03E8 03D4 03E8 03D4 03FC 03E8 03E8 03E8 03FC 03F2 0406 03D4 03D4 03DE 03F2
03CA 03E8

[2=WinAC.AVHistory,141]
PL1000-40297;Zündspannung Zyl. 2 (linke Bank);%;0;1;0;0;0;0;0;0;0;0;0;0;
0000-0021:0047 0047 0046 0047 004A 0046 0048 0045 0047 0047 0046 0049 0047 0046 0048 0045 0047 0048 0045 0046
0048 0046
…
…
…
[200=WinAC.AVHistory,42]
PL1000-40575;Brennstoffmassenstrom;kg/h;0;0,1;0;0;0;1;0;0;0;0;0;0;
0000-0021:0E41 0E26 0E3C 0E51 0E51 0E50 0E42 0E39 0E3D 0E48 0E51 0E33 0E2F 0E4A 0E48 0E3A 0E31 0E48 0E4D
0E2F 0E26
```

https://www.slideshare.net/Splunk/splunklive-munich-2018-use-casezndkerze-zeppelin

splunk> .conf18

© 2018 SPLUNK INC.

ZEPPELIN®
WE CREATE SOLUTIONS

How raw data looked like

```
[1=WinAC.AVHistory,155]
PL1000-40313;Motorlast;%;0;0,1;0;0;0;1;0;0;0;0;0;0;
0000-0021:03D4 03DE 03DE 03E8 03E8 03D4 03E8 03D4 03FC 03E8 03E8 03E8 03FC 03F2 0406 03D4 03D4 03DE 03F2
03CA 03E8 03F2 03D2 03E8
[2=WinAC.AVHistory,141]
PL1000-40297;Zündspannung Zyl. 2 (linke Bank);%;0;1;0;0;0;0;0;0;0;0;0;0;
0000-0021:0047 0047 0046 0047 004A 0046 0048 0045 0047 0047 0046 0049 0047 0046 0048 0045 0047 0048 0045 0046
0048 0046
…                0045 0048 0046
…
…
[200=WinAC.AVHistory,42]
PL1000-40575;Brennstoffmassenstrom;kg/h;0;0,1;0;0;0;1;0;0;0;0;0;0;
0000-0021:0E41 0E26 0E3C 0E51 0E51 0E50 0E42 0E39 0E3D 0E48 0E51 0E33 0E2F 0E4A 0E48 0E3A 0E31 0E48 0E4D
0E2F 0E26


                 0E33 0E50 0E39
```

Metric — Wert für Minute 22

Unit — Wert für Minute 23

Value — Wert für Minute 24

https://www.slideshare.net/Splunk/splunklive-munich-2018-use-casezndkerze-zeppelin

splunk> .conf18

© 2018 SPLUNK INC.

ZEPPELIN®
WE CREATE SOLUTIONS

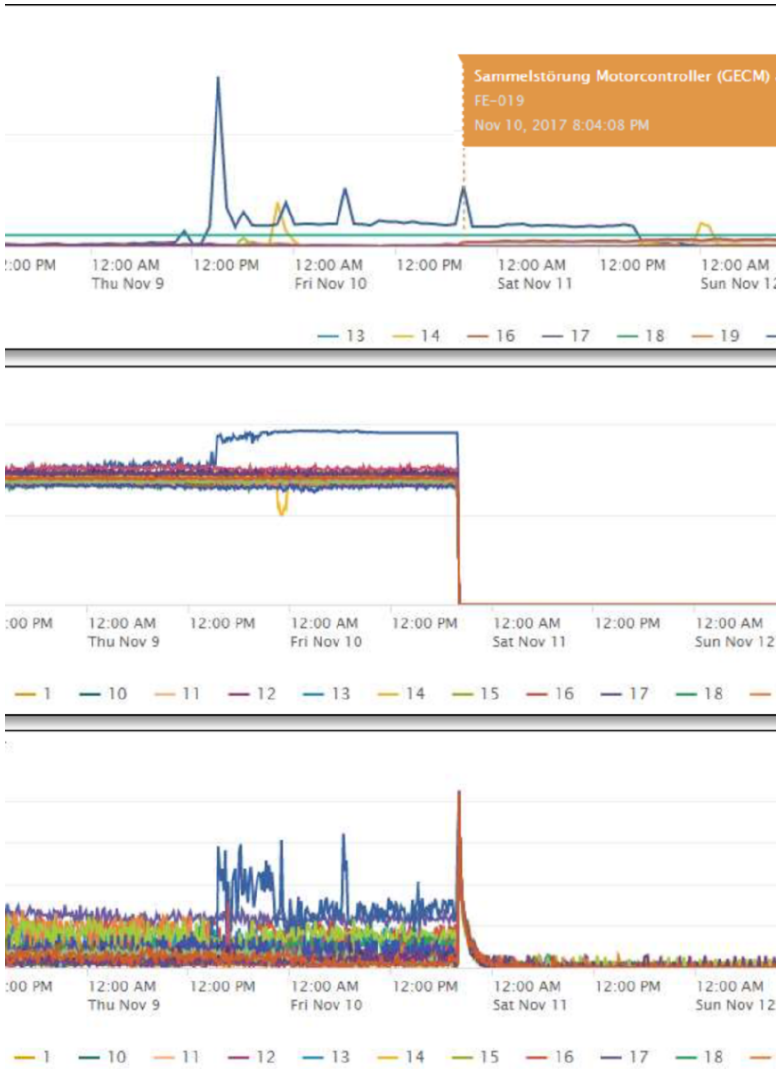
Deep dive
A bit pseudo SPL

Unsupervised Approach 1:

```
index=motor
| timechart max(voltage)
by cylinder
| fit KMeans k=1 *
```

Unsupervised Approach 2:

```
| gentimes start=-30
increment=3h
| map search="
search index=motor
| fields voltage ...
| anomalydetection
| eval time=$starthuman$,
start = $starttime$,
end = $endtime$"
| collect index=anomalies
```



Supervised Approach 1:

- Use outages as labels
- Multiply labels as declining risk score before outage
- Train either categorical prediction model (target risk score as step function) or train numerical prediction (target risk score as numeric value)

Ensemble for the end game :

Combination of multiple approaches give the final signal

splunk> .conf18

© 2018 SPLUNK INC.

Predictive Maintenance Project: lessons learnt

Workflow and achieved project milestones

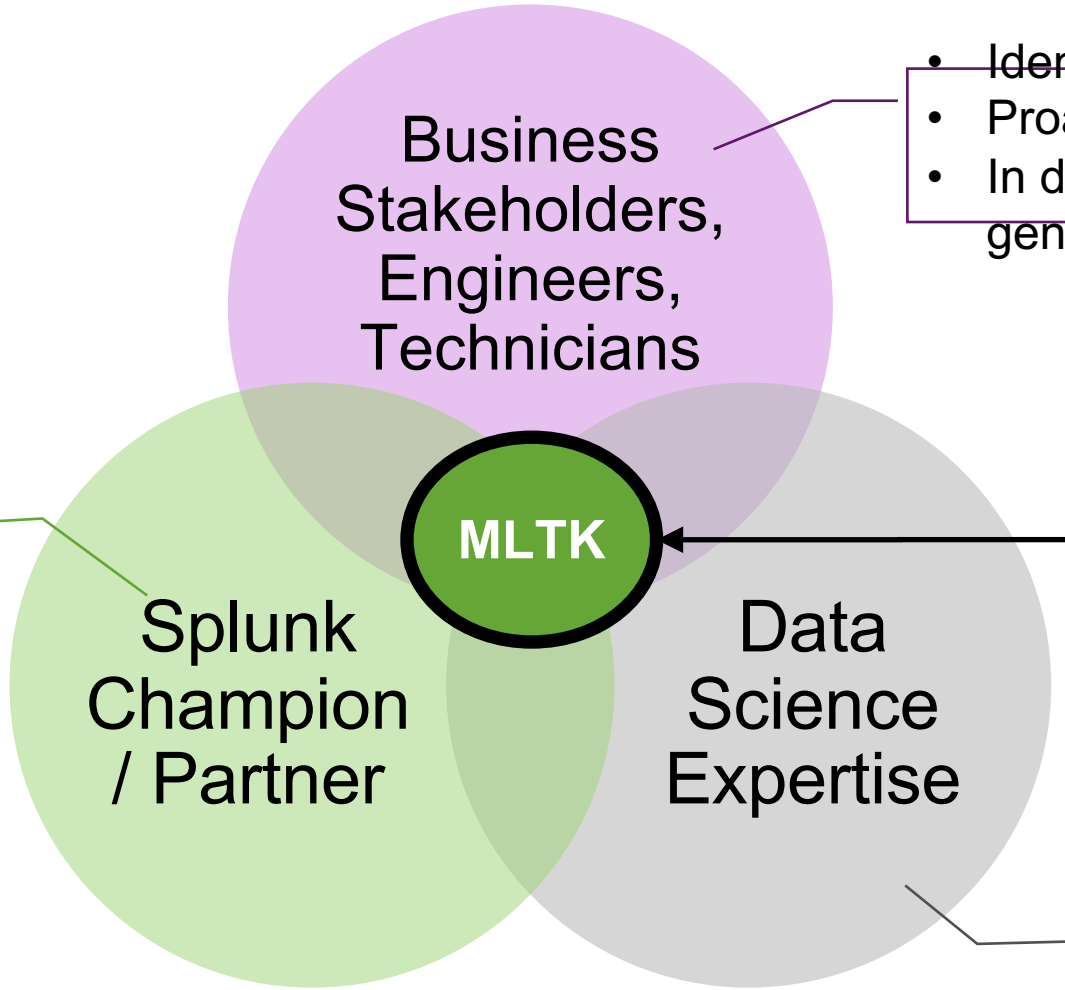
Decision for :	Data prep	Domain expertise	Modelling	Piloting
Stakeholders decisions aligned	Data collection and preparation	Engineers and domain expertise	Create and evaluate 4 models	Tailor custom dashboards
			Results verified by business	Continuous feedback
Project scoped	mid 2017		**fall 2017**	Work, verify and collect feedback
				2018

- EDA + Domain Knowledge is key
  - 70% data prep, cleaning, transformation etc.
  - Consultation of engineers and service technicians
- Effectively 3 days to evaluate 4 modelling approaches
- Quick agile iterations with the customer

splunk> .conf18

© 2018 SPLUNK INC.

Success Formula for the Zündkerze!



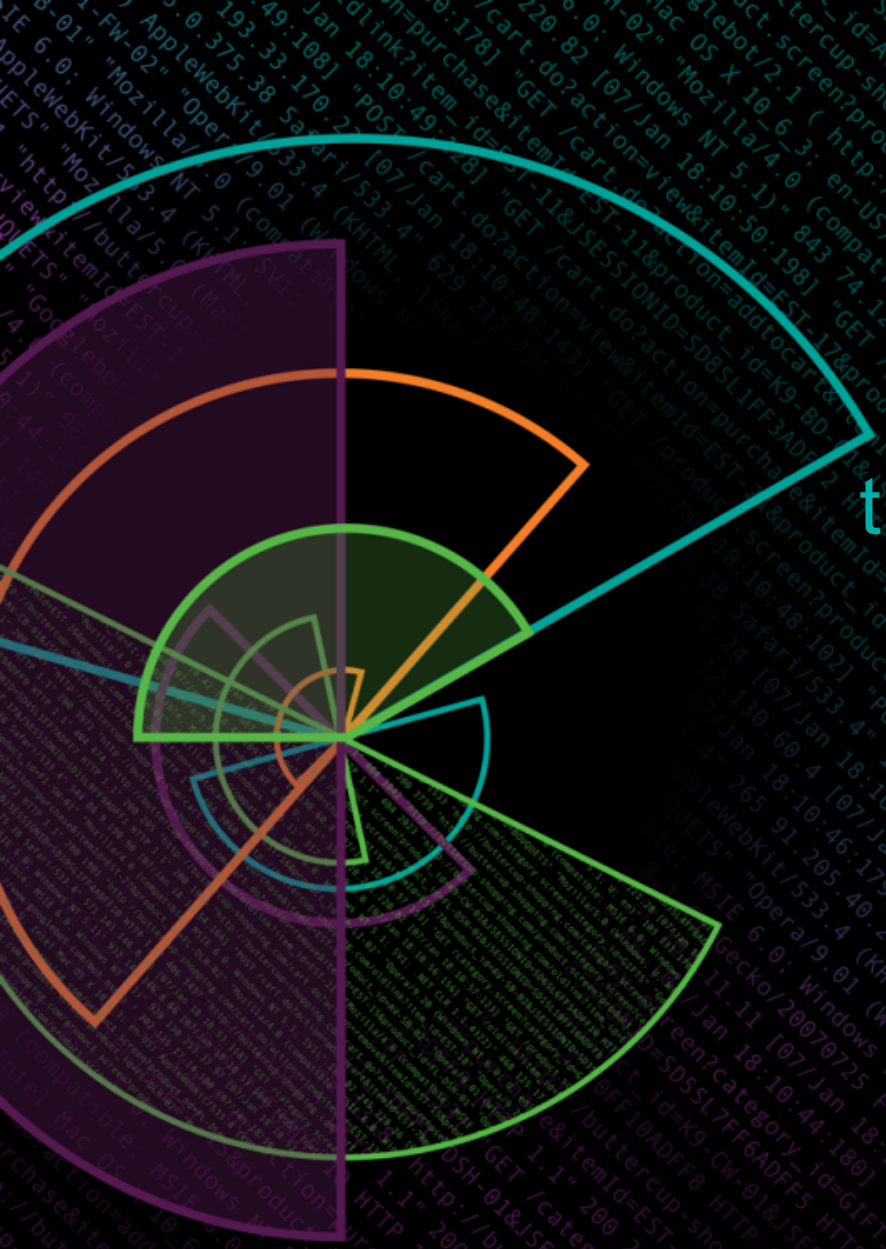
- ~~Identify saving potentials~~
- Proactive handling of outages
- In depth understanding of the power generator physics

- Data onboarding
- SPL for data transformation
- Dashboarding and Alerting

Splunk ML Toolkit helped to quickly glue collaborative modelling efforts together!

- Statistics/math background
- Algorithm selection
- Model building

© 2018 SPLUNK INC.



"Employing the model developed with the MLTK we have been able to identify the early detection of 5 - 7 failures per system leading to approximate savings of €150,000 per year."

Rene Ahlgrim, Data Science Manager
Zeppelin Powersystems

© 2018 SPLUNK INC.

Wrap up

© 2018 SPLUNK INC.

Wrap up

- **Where to start?**
  Take a clear defined business challenge and get started with a simple, limited scope and expand from there
- **Data journey:**
  Data, EDA, Modeling, Operationationalize, Complexity.
- **Domain Expertise is crucial**

© 2018 SPLUNK INC.

Wrap up

- **Get support:**
  Leverage internal data science departments or Splunk or Splunk Partner (ML Advisory)
- **Use the metrics store:**
  IoT data mostly comes in perfect shape for metrics store. Accelerate your analytics
- **No fear:**
  many IoT data sets are not far from IT ones: Metrics of CPU, Mem etc. / Events of Outages
- **Get started!**

© 2018 SPLUNK INC.

Q&A

© 2018 SPLUNK INC.

Thank You

Don't forget to rate this session in the .conf18 mobile app