



Machine Learning @ Splunk 2019

# The Splunk Machine Learning Toolkit in Action

Iman Makaremi & Harsh Keswani  
Machine Learning PM | Splunk



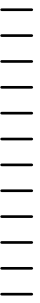
**Iman Makaremi**

Product Manager, ML & AI, Splunk



**Harsh Keswani**

Product Manager, ML & AI, Splunk



# Forward-Looking Statements

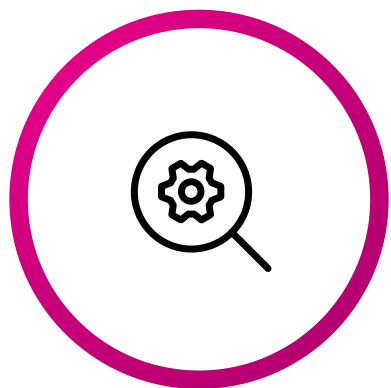


During the course of this presentation, we may make forward-looking statements regarding future events or plans of the company. We caution you that such statements reflect our current expectations and estimates based on factors currently known to us and that actual events or results may differ materially. The forward-looking statements made in the this presentation are being made as of the time and date of its live presentation. If reviewed after its live presentation, it may not contain current or accurate information. We do not assume any obligation to update any forward-looking statements made herein.

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 functionalities described or to include any such feature or functionality in a future release.

Splunk, Splunk>, Turn Data Into Doing, 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. © 2019 Splunk Inc. All rights reserved.

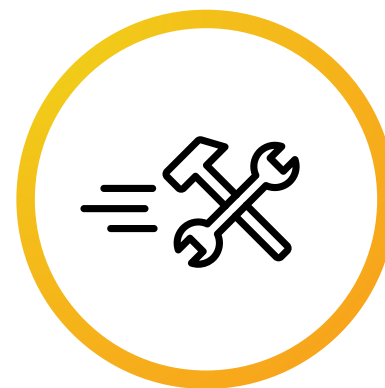
# Agenda



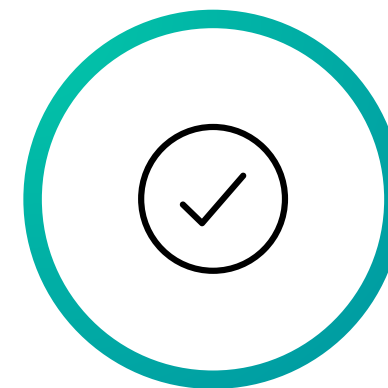
.conf Talk Analysis



ML at Splunk



MLTK in Action



What's New?

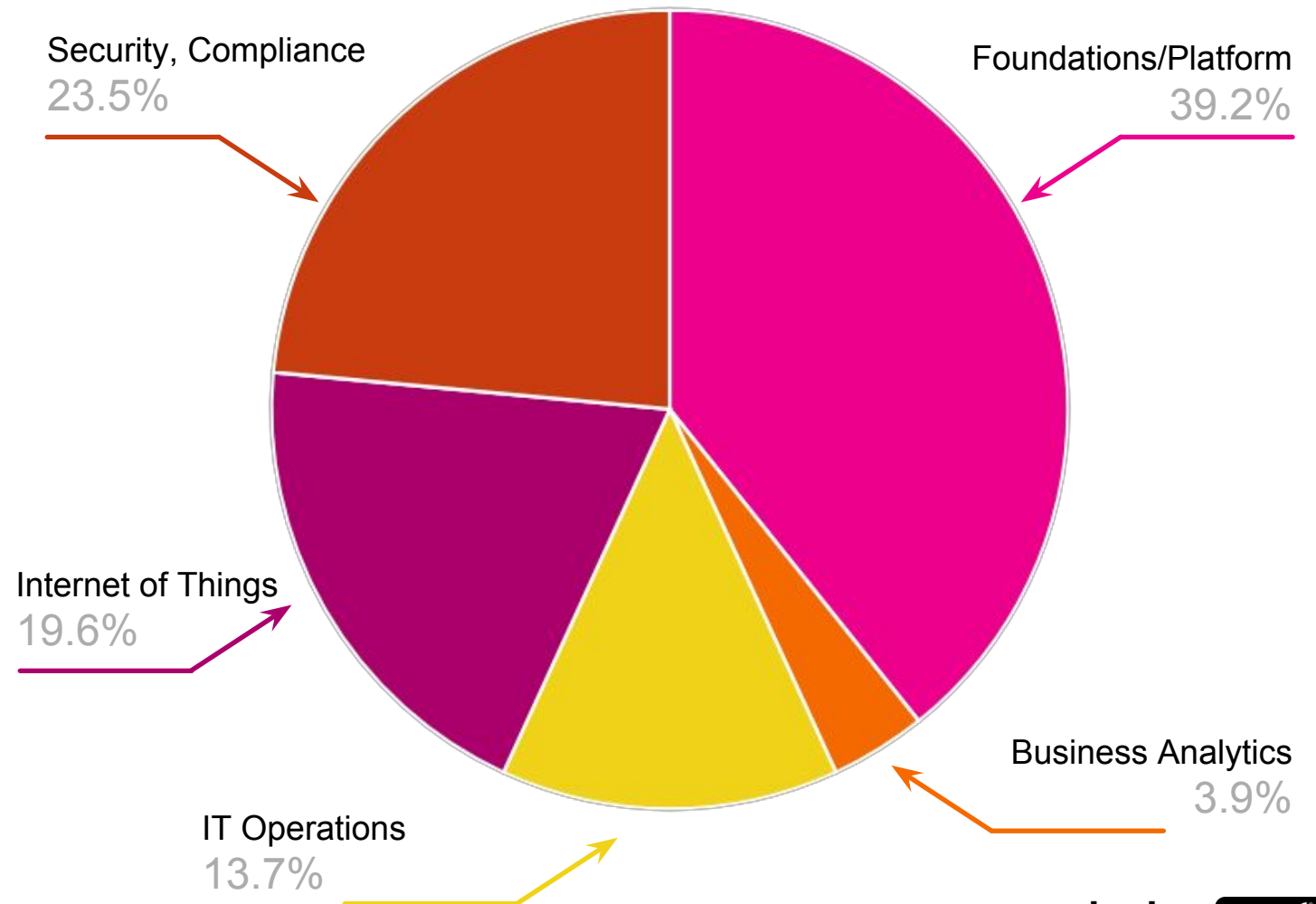


# Talk Data Analysis

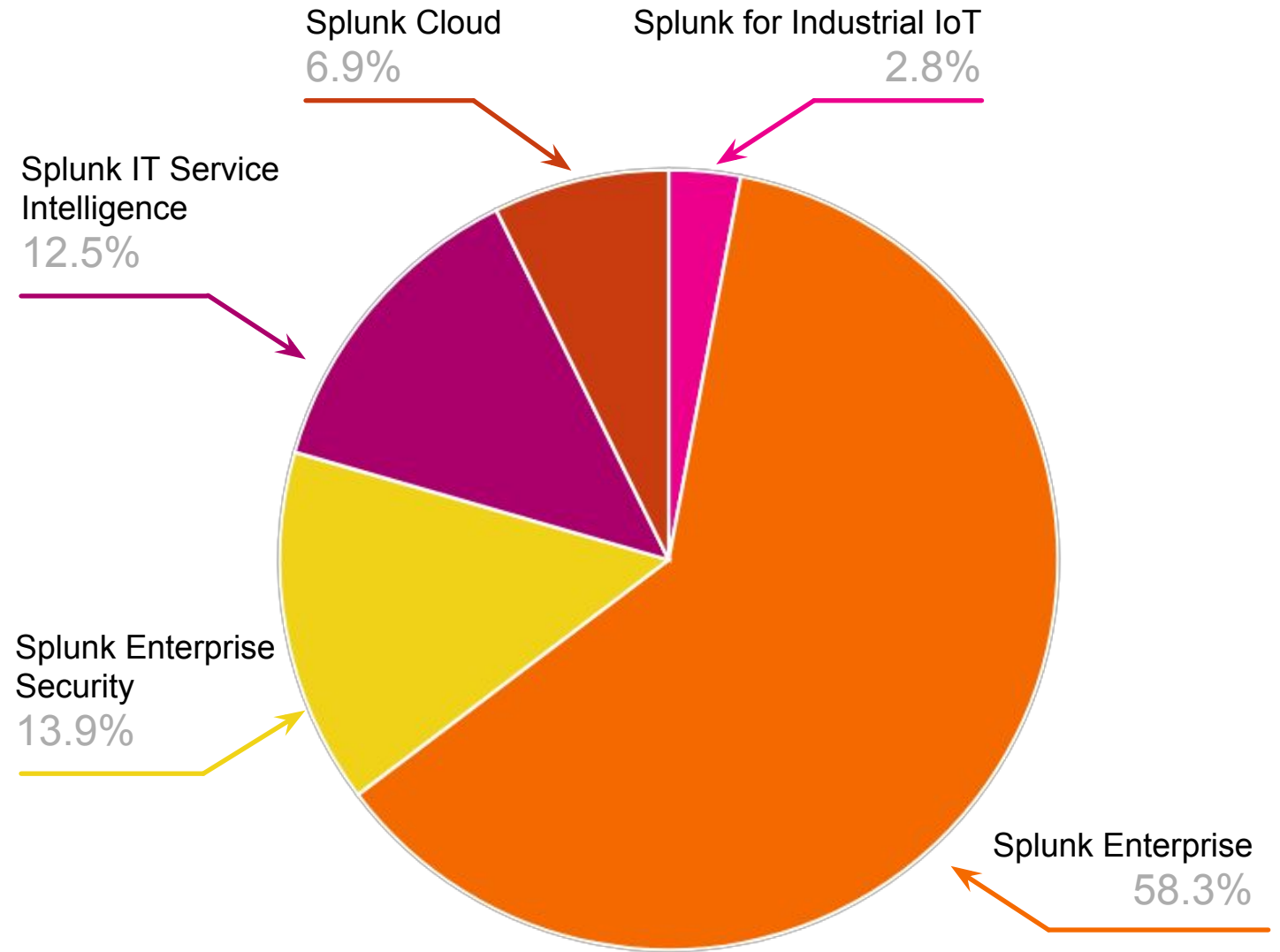
---

MLTK Impact in 2019 .conf Talks

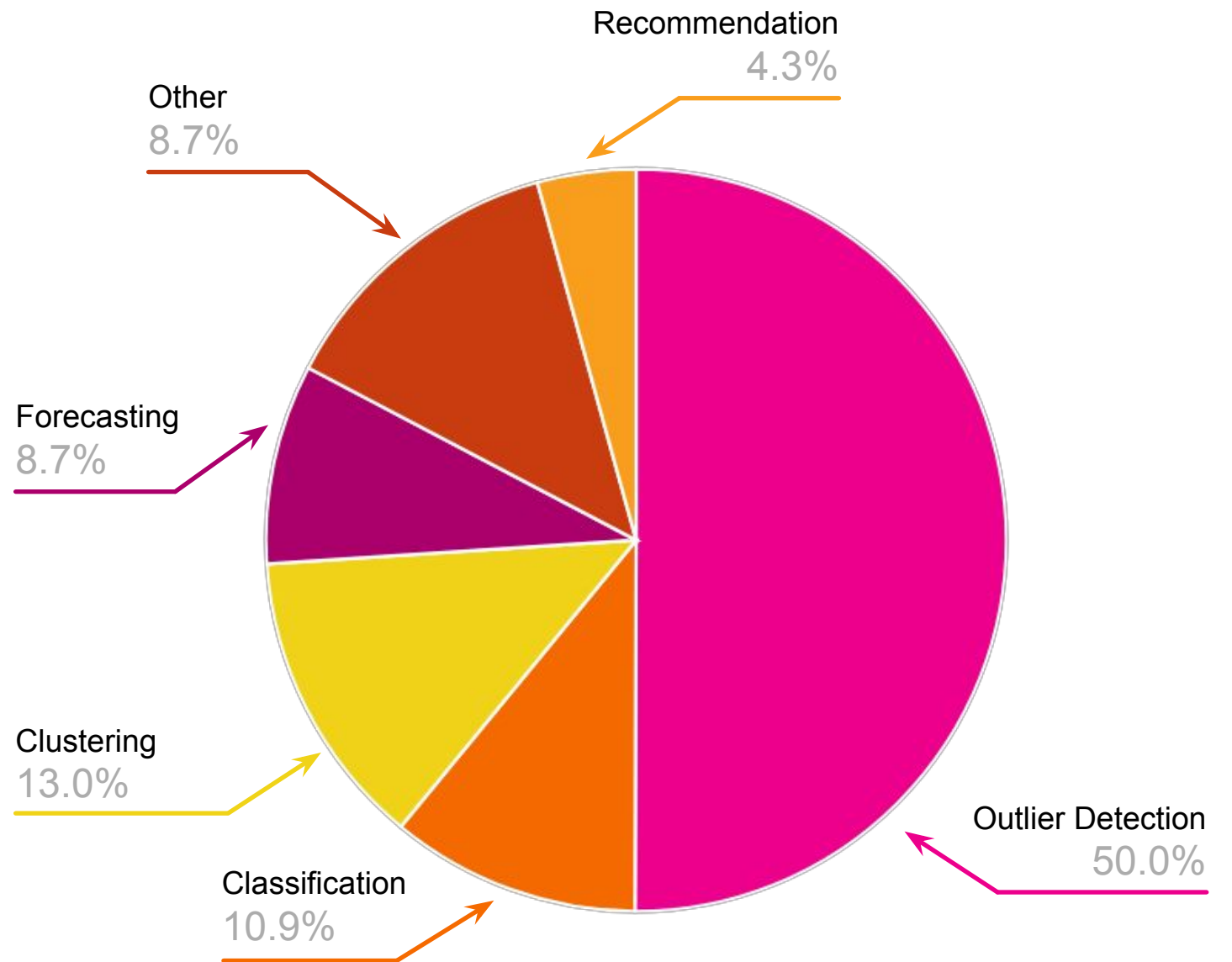
# Talks per Sessions Track



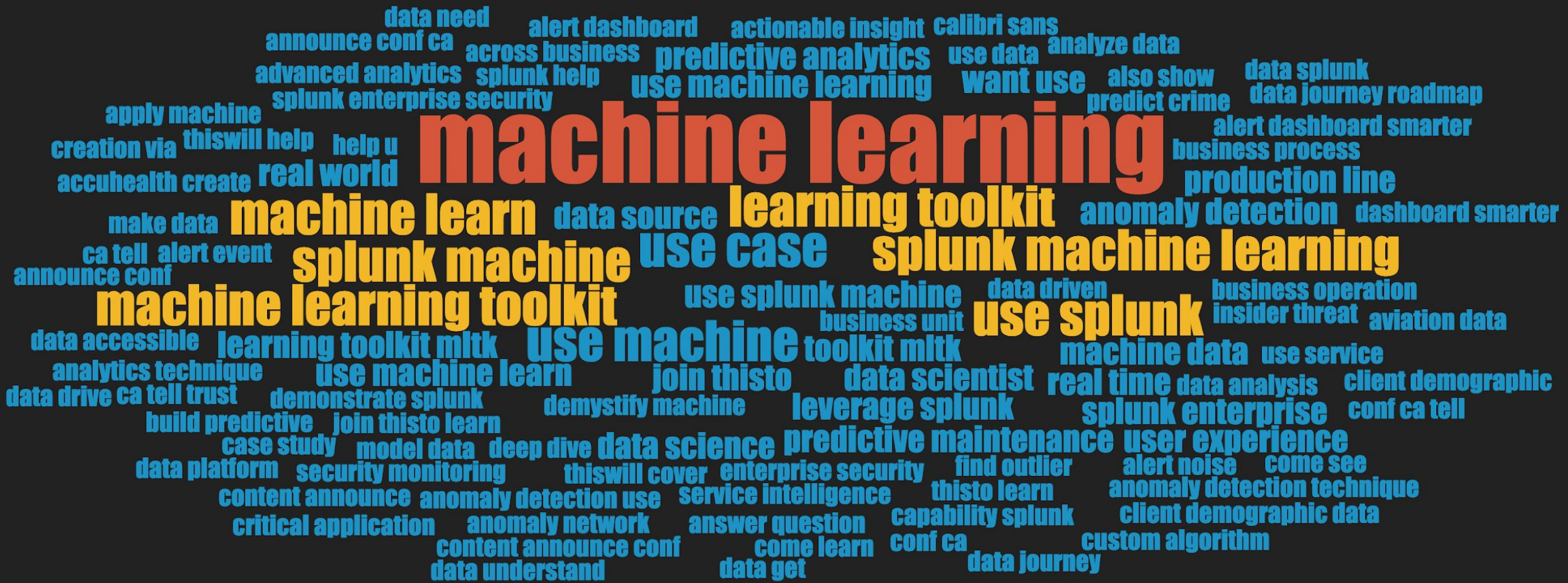
# MLTK Talks Per Product Category

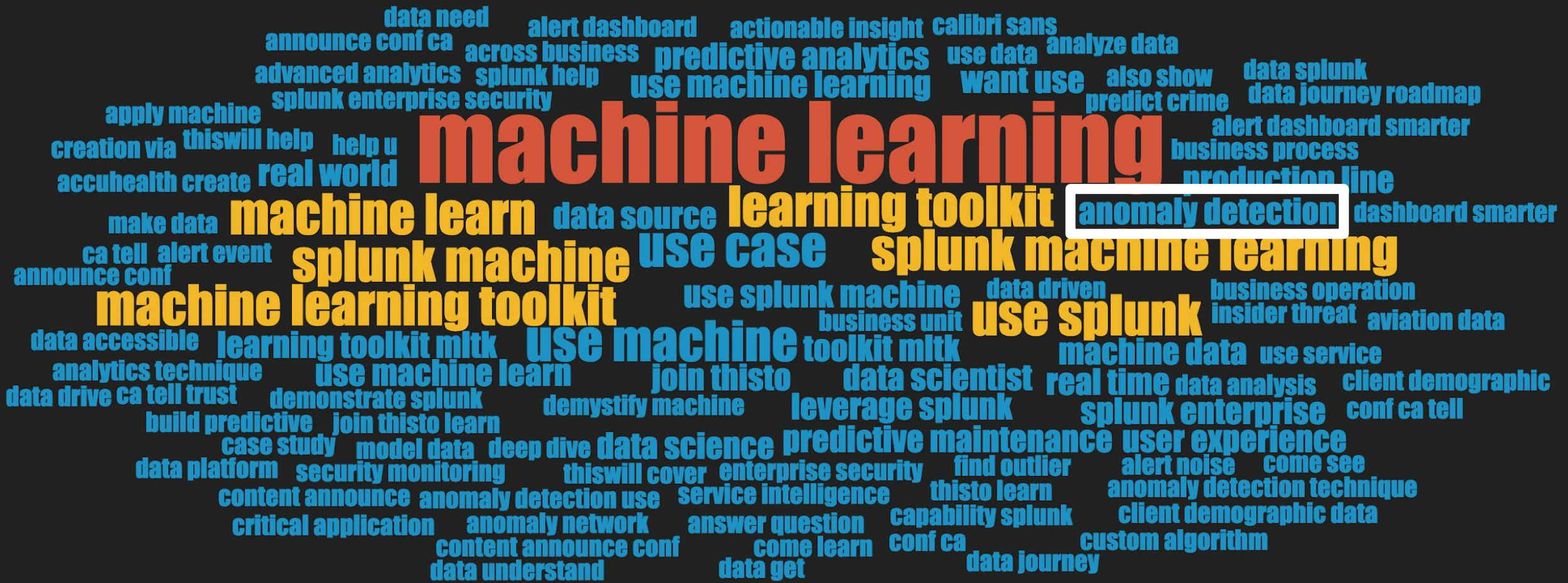


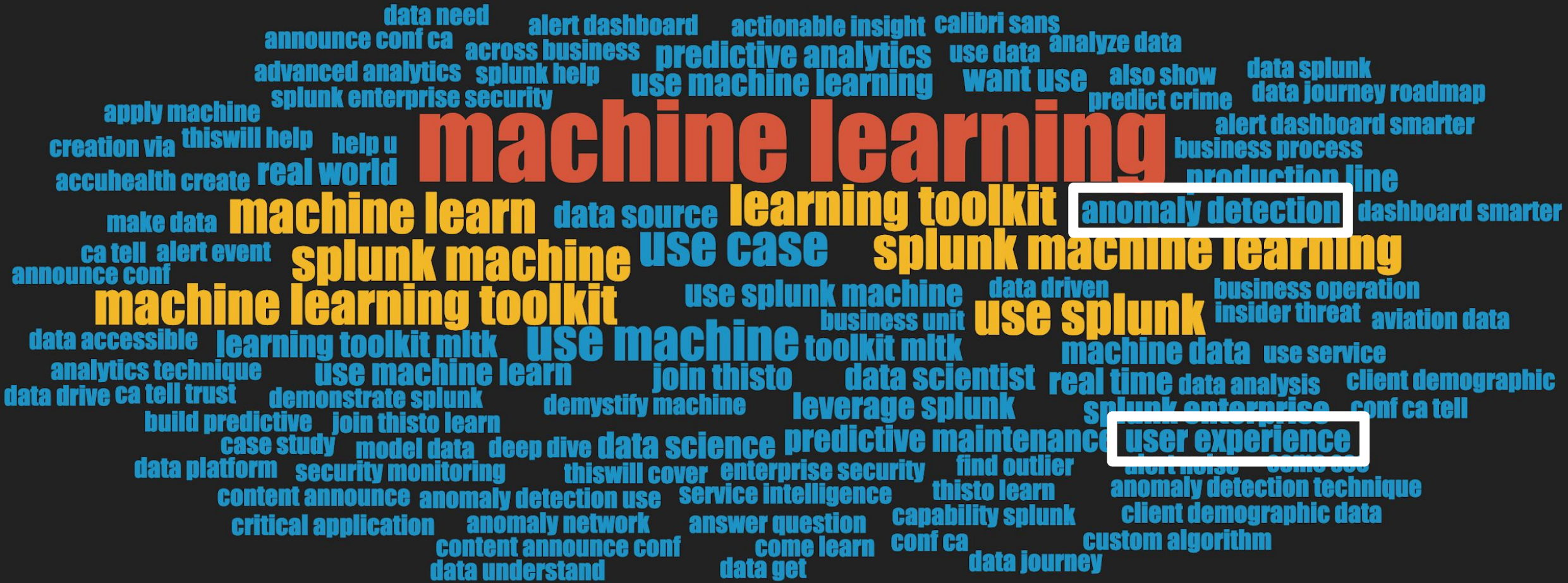
# Talks Per ML Use Case

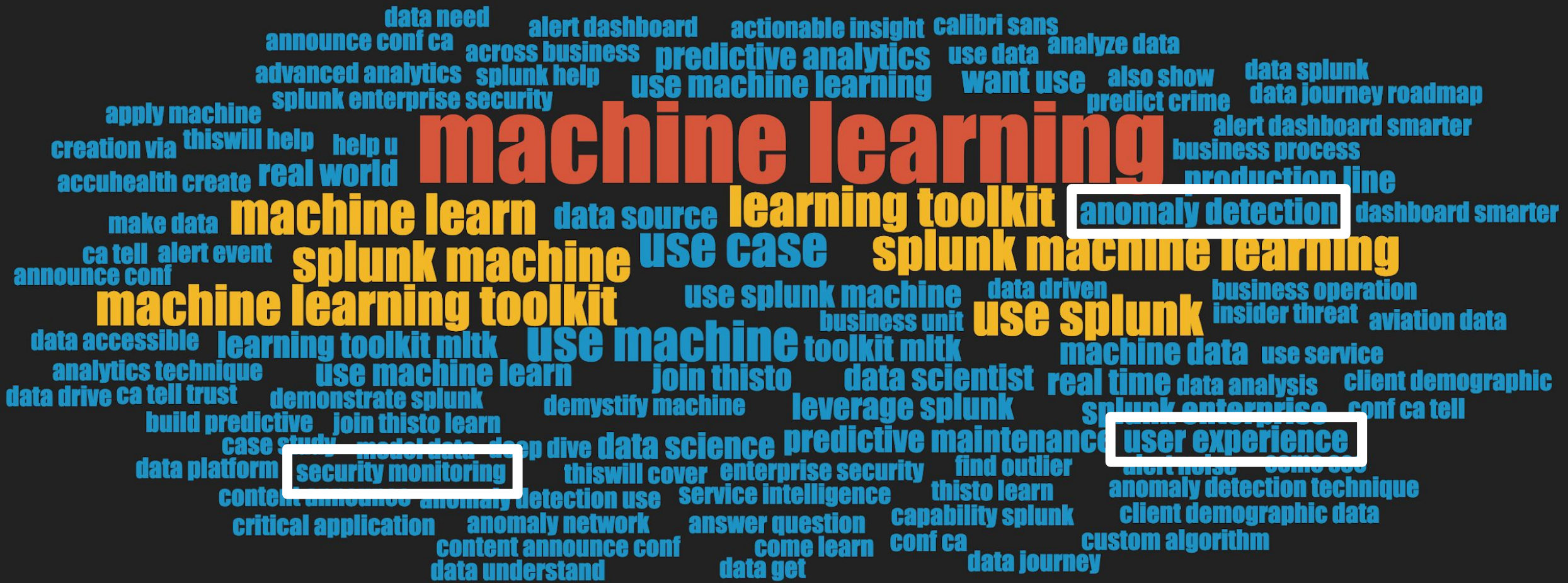


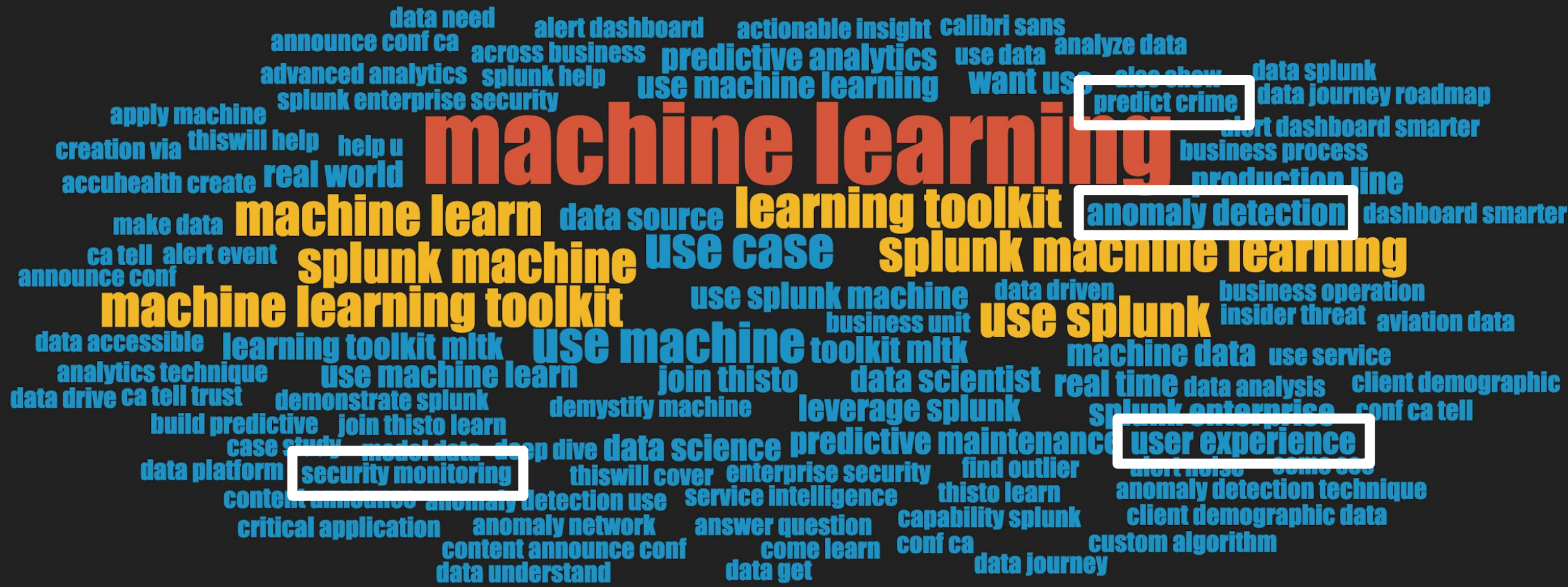












# machine learning

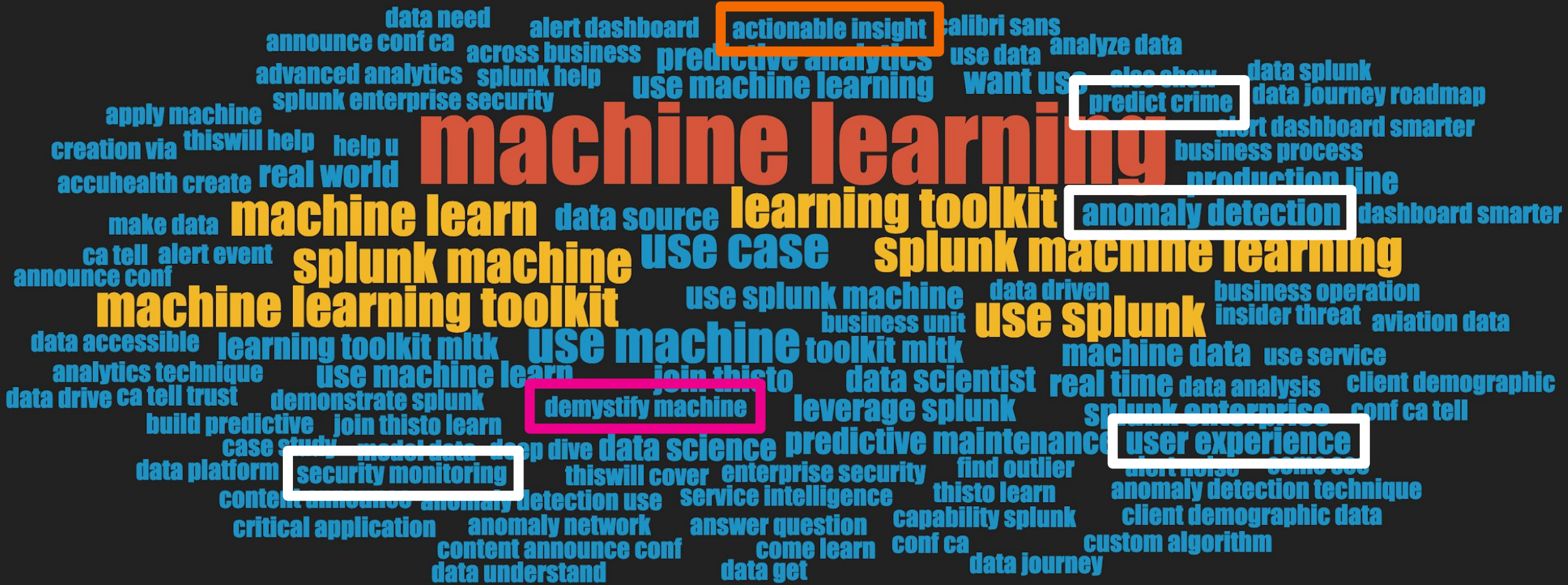
predict crime

anomaly detection

demystify machine

security monitoring

user experience





# MLTK?!

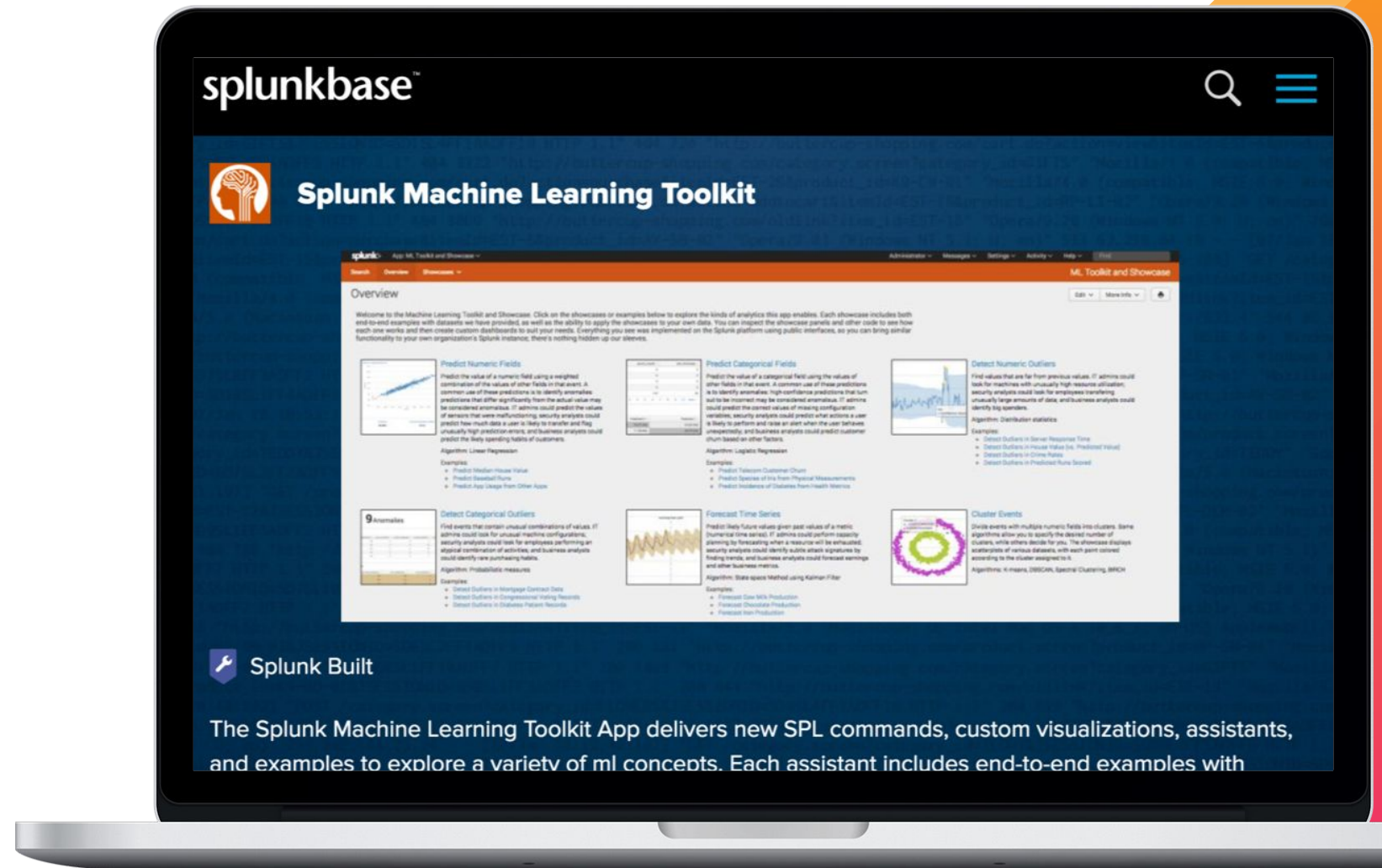
---



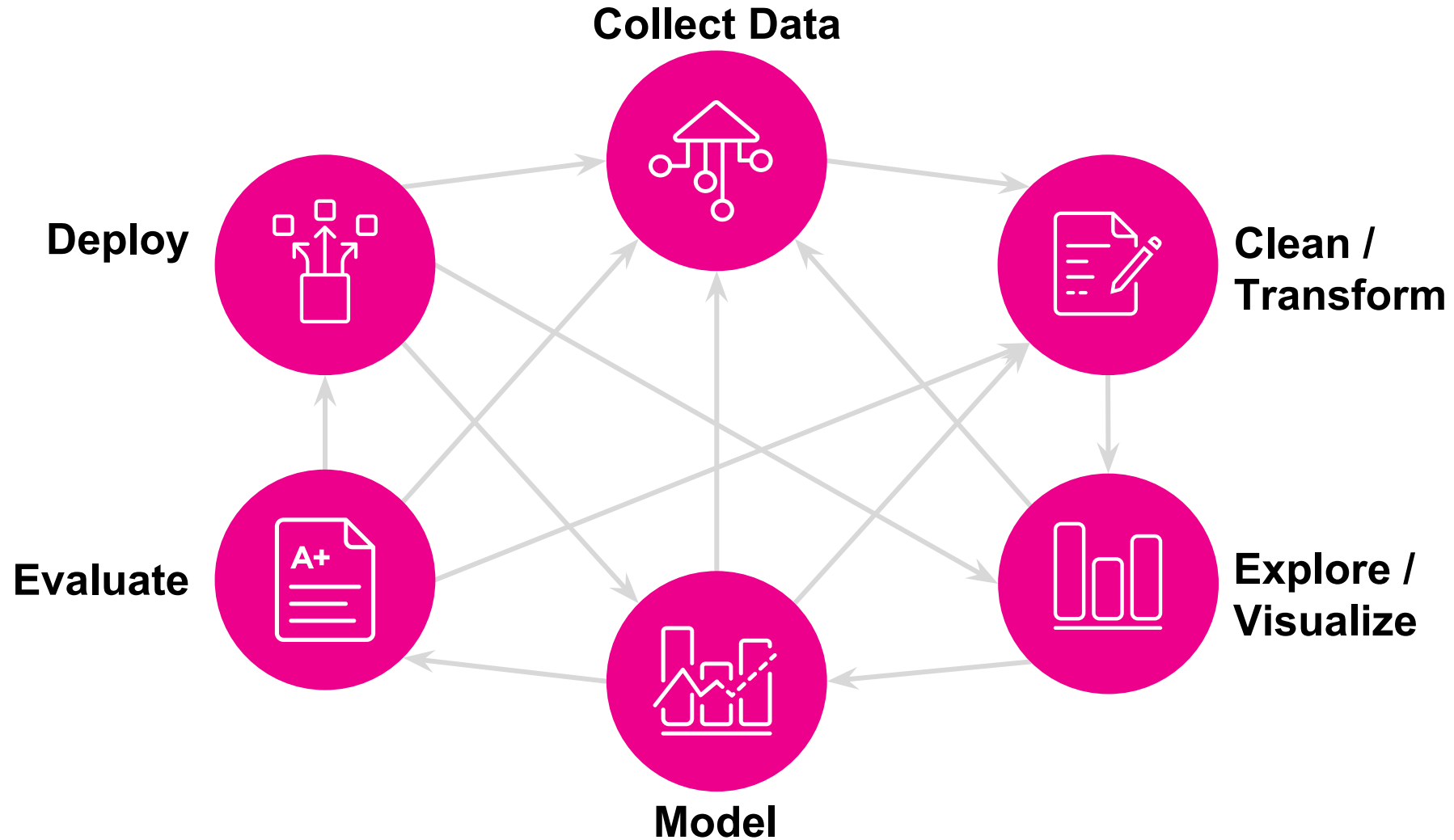
# What?

## Splunk has a Machine Learning Toolkit App!

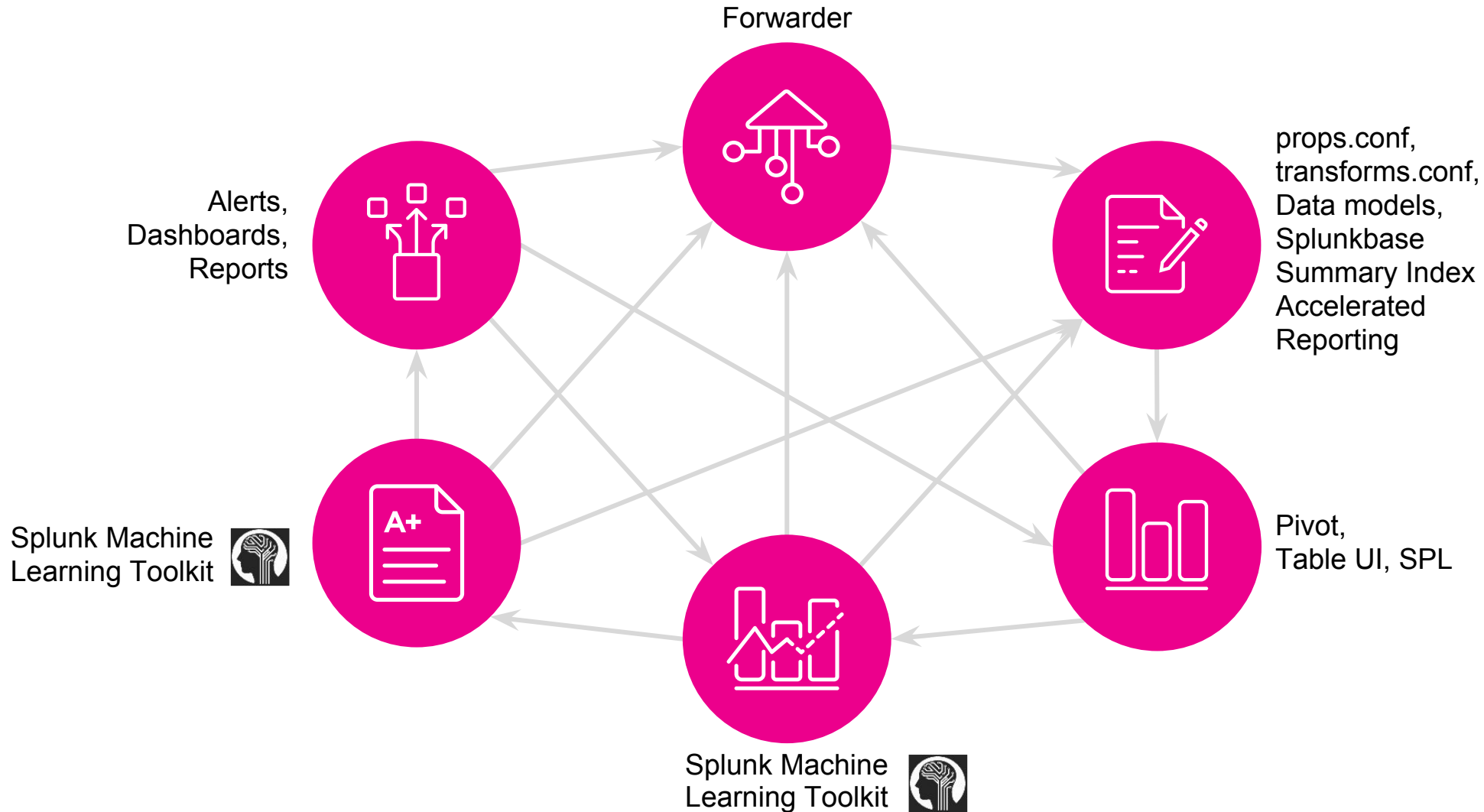
- What is Splunkbase
- What is the App
- Where can I go to learn more



# Machine Learning Process

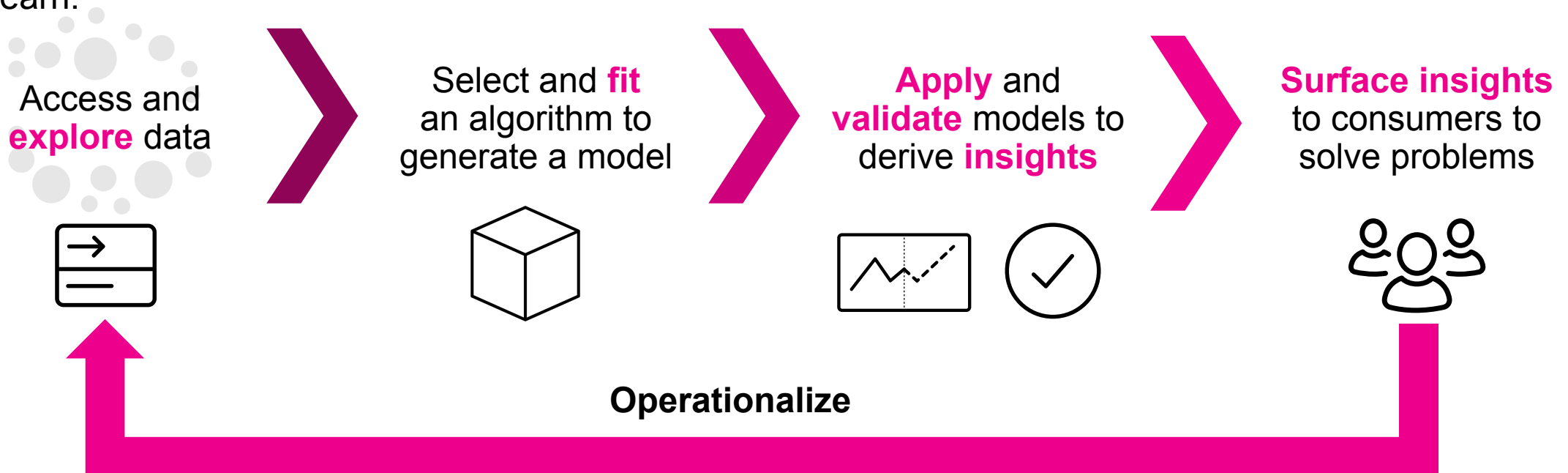


# Machine Learning Process with Splunk



# Splunk Helps Answer Questions with AI and ML

- **Identify a Problem:** <Stuff in the world> requires big time and money investment.
- **Build a Solution:** Build ML model to forecast <possible incidents>, act preemptively and continuously learn.



# MLTK Delivers Easy-to-Use Outcome Assistants

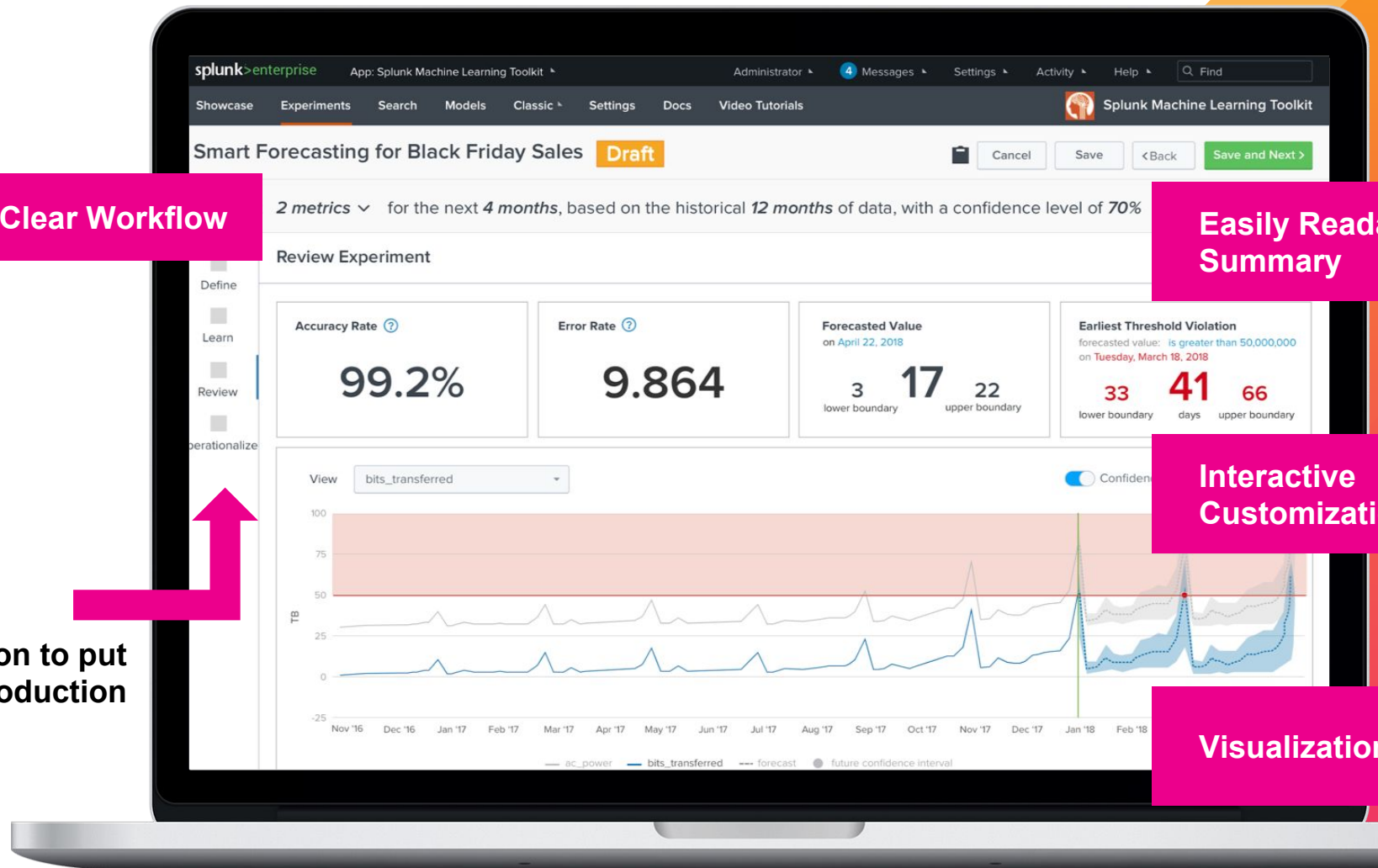
Easy Automation to put  
ML into Production

Clear Workflow

Easily Readable  
Summary

Interactive  
Customizations

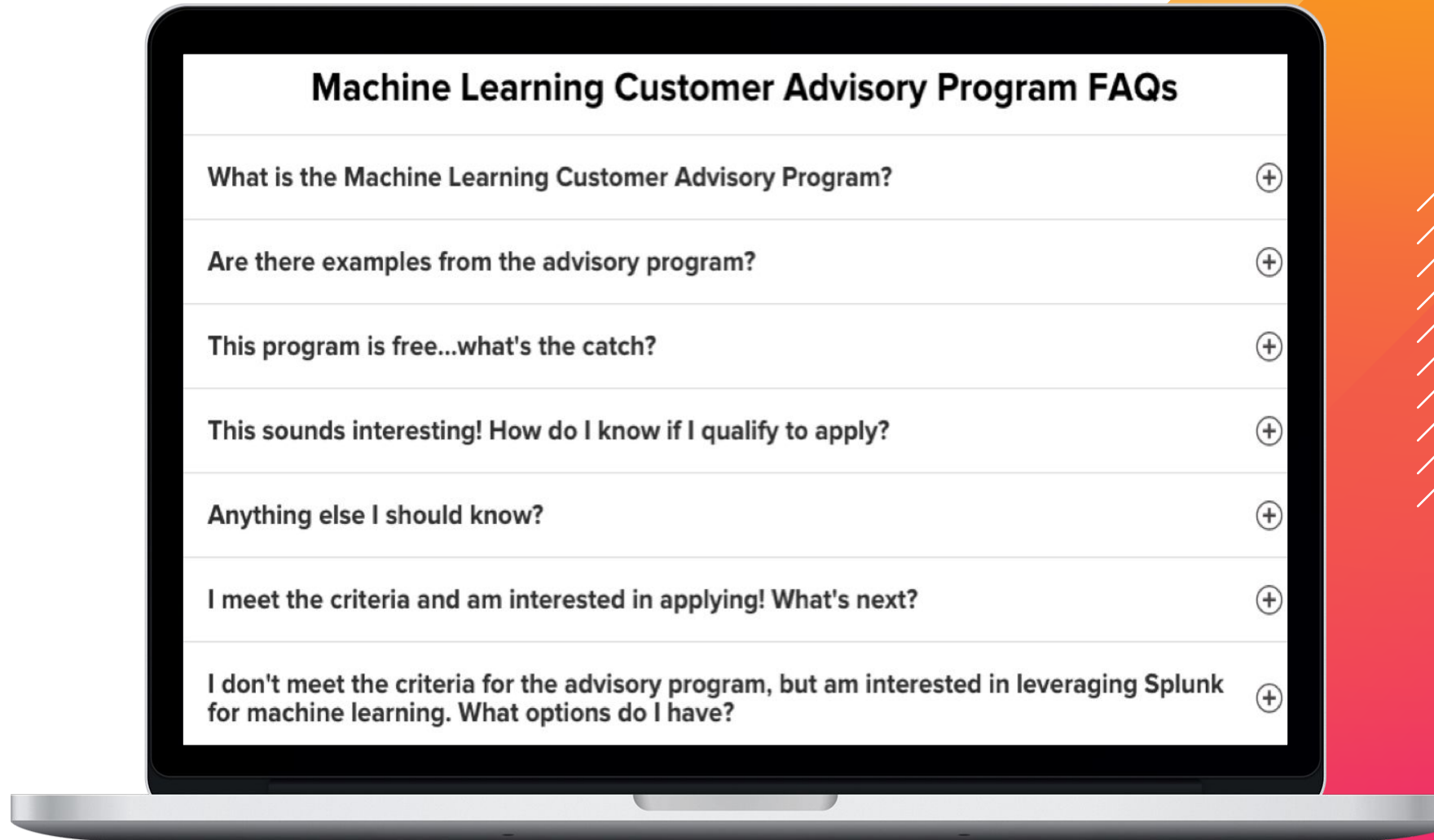
Visualizations



# What is the ML Advisory Program?

Complimentary support of Splunk data science resources to help build a ML use case resulting in a public reference

- Early access to new and enhanced Machine Learning features
- Opportunity to shape the development of the product
- Complimentary assistance in operationalizing a production quality ML model



# ML Advisory Customers

**T-Mobile**

**Siemens**

**Intel**

**StubHub**

**TransUnion**

**Ministry  
of  
Israel**

**Honda**

**BMW**

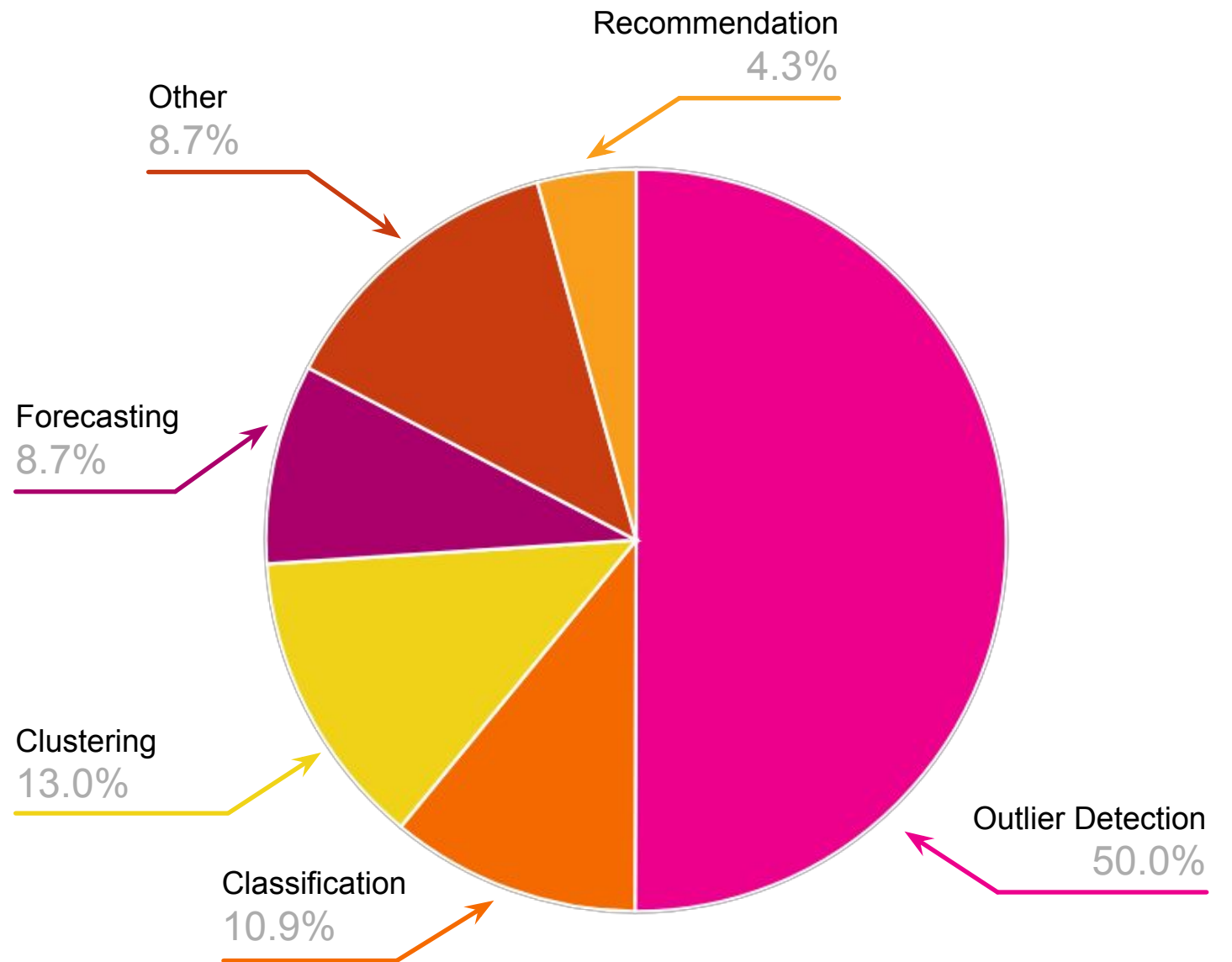


# MLTK in Action

---



# Talks Per ML Use Case





# Outlier Detection

---

**Outlier Detection**  
**Anomaly Detection**  
**Threat Detection**  
**Fraud Detection**  
**Intrusion Detection**  
**Threat Hunting**  
**Outage Prediction**



**T-Mobile**

# Identifying service-impacting events across the deployed enterprise applications.



**FN1366 - Enhanced Anomaly Detection: Join T-Mobile and Splunk as we Deep Dive an Enterprise-IT Operational Use Case**

---

**Wednesday, October 23, 01:45 PM - 02:30 PM**

Iman Makaremi, Principal Product Manager – Machine Learning and AI, Splunk  
Scott Garcia, MTS - Member Technical Staff, T-Mobile



**Stubhub**

# Automatic monitoring of performance KPIs and anomaly discovery within minutes.



**IT1171 - Accelerate your ability to sniff out application exceptions and detect outliers in performance KPIs**

---

**Tuesday, October 22, 04:15 PM - 05:00 PM**

Eurus Kim, Staff ML Architect, Splunk  
PJ Pokhrel, Performance Engineer, Stubhub  
Steve Veio, Performance OPS Manager, StubHub



**Aflac**

# Insurance Fraud Detection



**SEC1904 - The Duck Test: Leverage Machine Learning to Remediate Fraud in Huge Datasets**

---

**Thursday, October 24, 01:00 PM - 01:45 PM**

Matthew Harper, Director, Cyber Crime Prevention, Aflac



**Siemens**

# Datacenter security monitoring for hacker detection.



**SEC1374 - Augment Your Security Monitoring Use Cases with Splunk's Machine Learning Toolkit**

**Thursday, October 24, 11:45 AM - 12:30 PM**

Oliver Kollenberg, Security Consultant, Siemens AG  
Philipp Drieger, Staff Machine Learning Architect , Splunk

# MLTK Offers



Numeric Outlier  
Detection Assistant



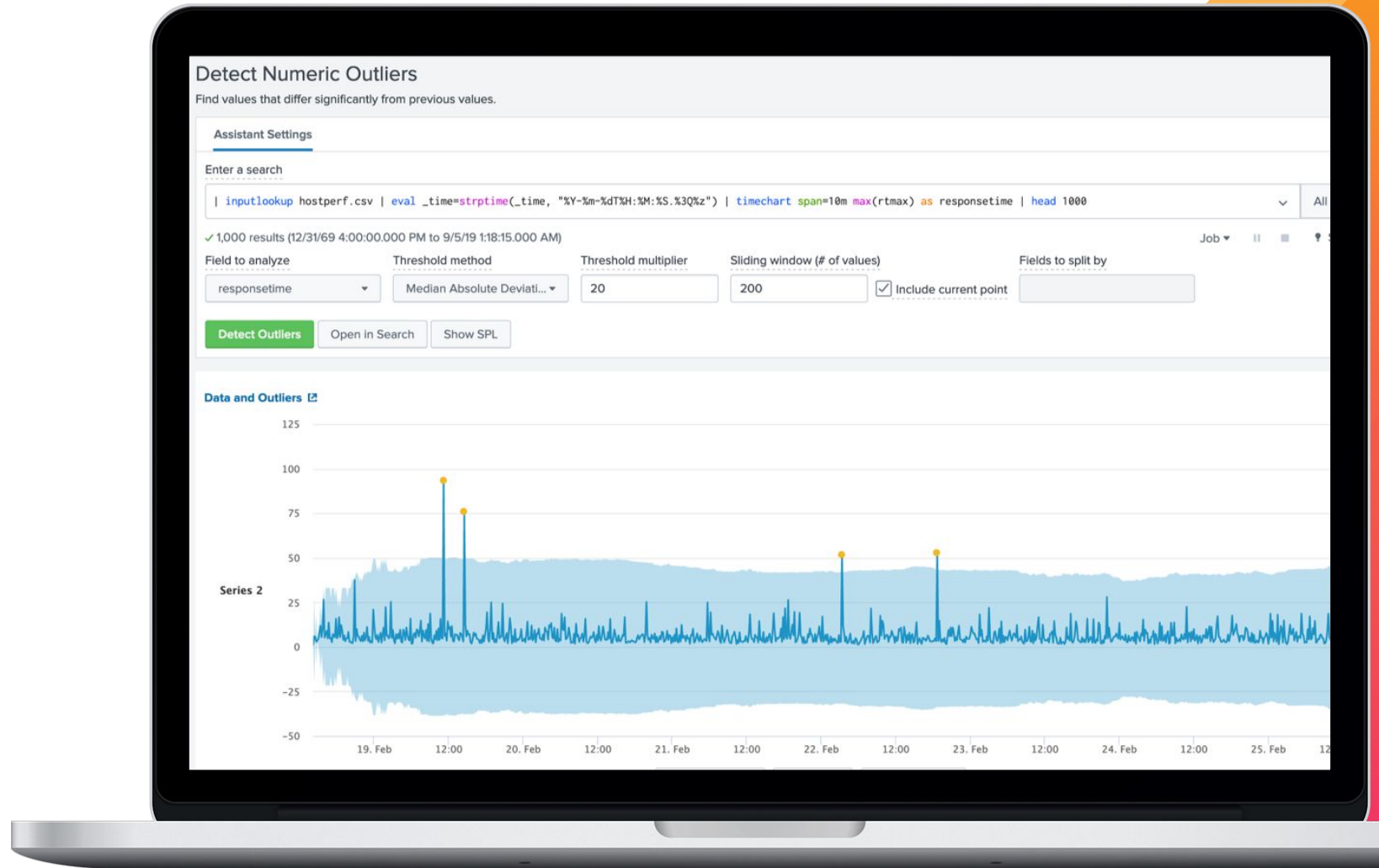
Smart Outlier  
Detection Assistant



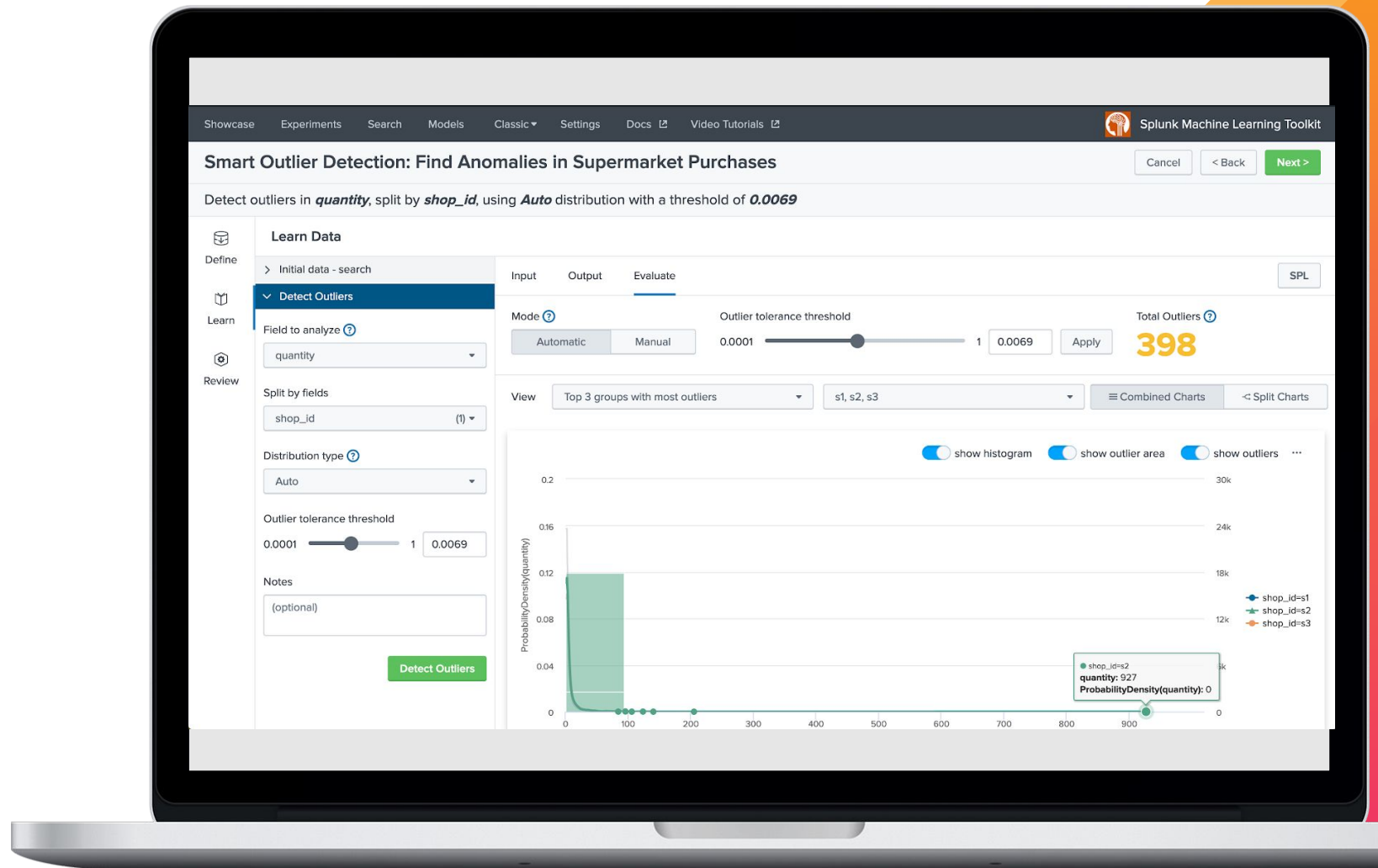
Custom  
ML



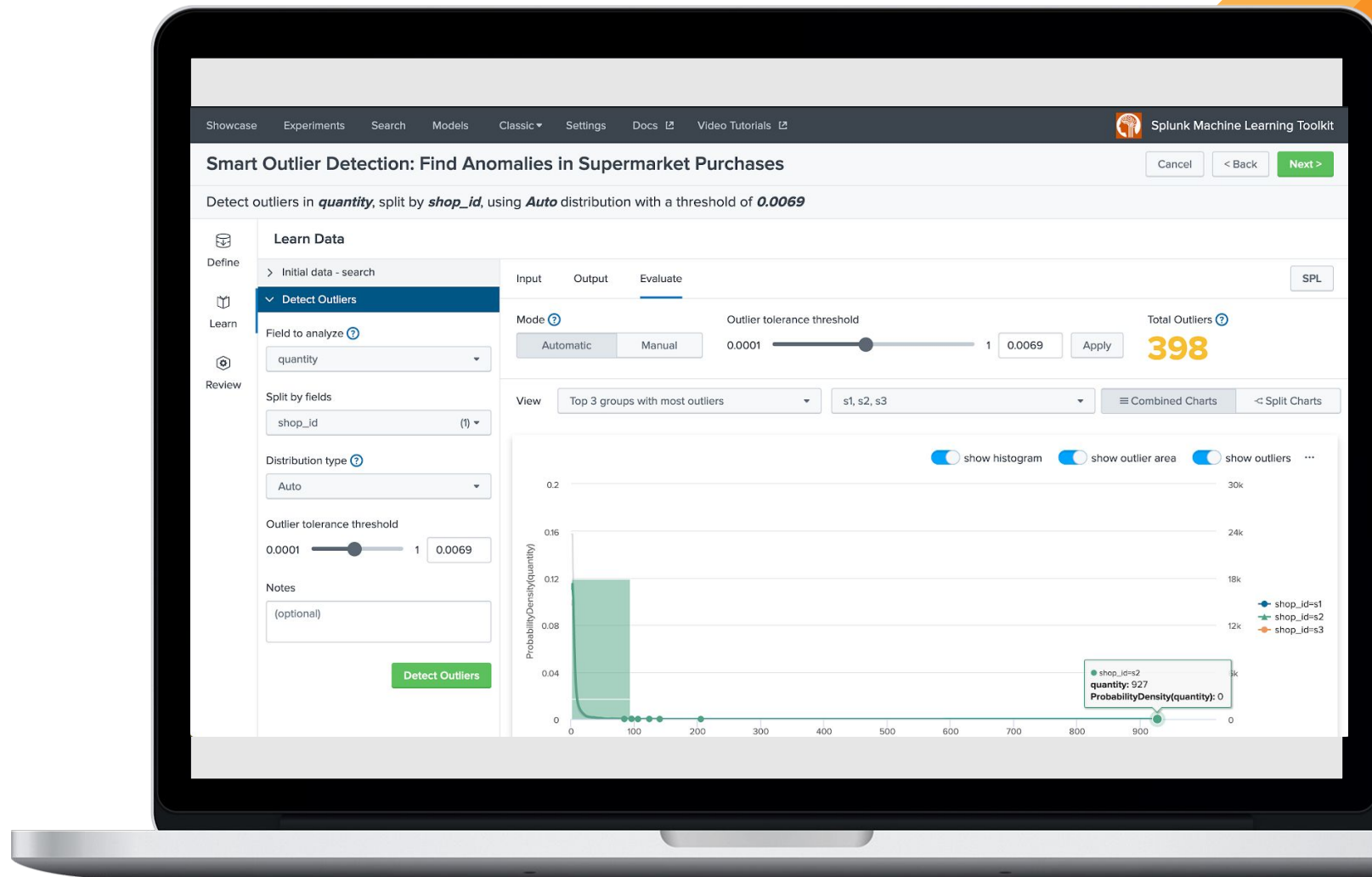
# Numeric Outlier Detection Assistant



# Smart Outlier Detection Assistant



# Smart Outlier Detection Assistant



# Custom ML

- Custom ML Outlier Detection Blog [\(link\)](#)
- Add Custom ML Alg [\(link\)](#)



# Forecasting

---



**T-Mobile**

# Predict future outages by forecasting congestion rates of cell towers.



**IT1722 - Predict Real World Outage using Splunk MLTK**

---

**Wednesday, October 23, 12:30 PM - 01:15 PM**

Vijay Veggalam, Member of Technical Staff, T-Mobile  
Gintaras Gaigalas, Sr. RF Engineer, T-Mobile



TransUnion

# Eliminate out-of-disk outages by forecasting disk usage and creating what-if scenarios.



**FN1137 - Forecasting Disk Usage with Machine Learning – So easy,  
even a cave-person can do it!**

---

**Tuesday, October 22, 01:45 PM - 02:30 PM**

Steve Koelpin, Splunk Advisor, TransUnion  
Alicia Dale, Site Reliability Engineer, TransUnion

# MLTK Offers



Time Series  
Forecasting Assistant



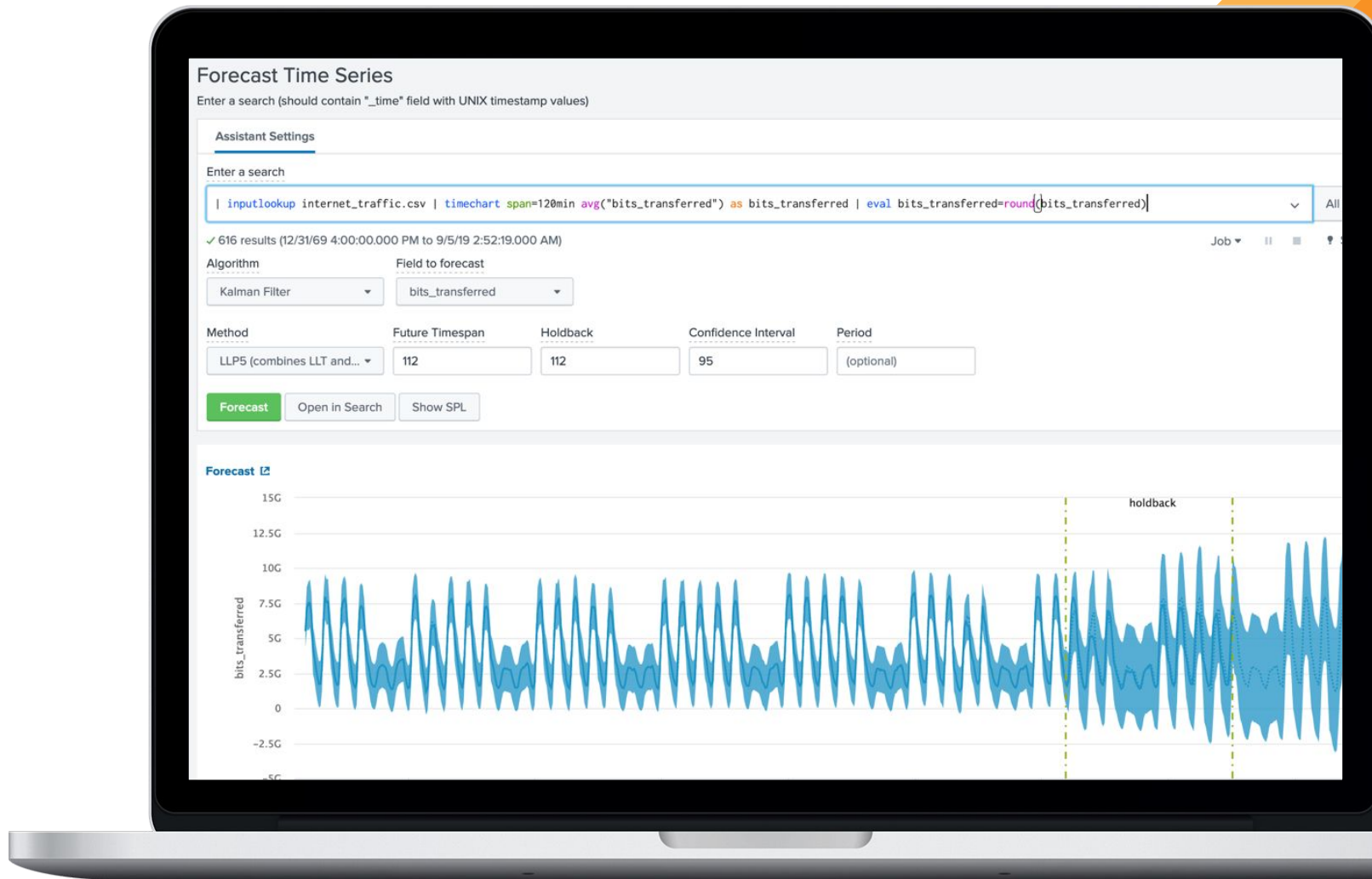
Smart Forecasting  
Assistant



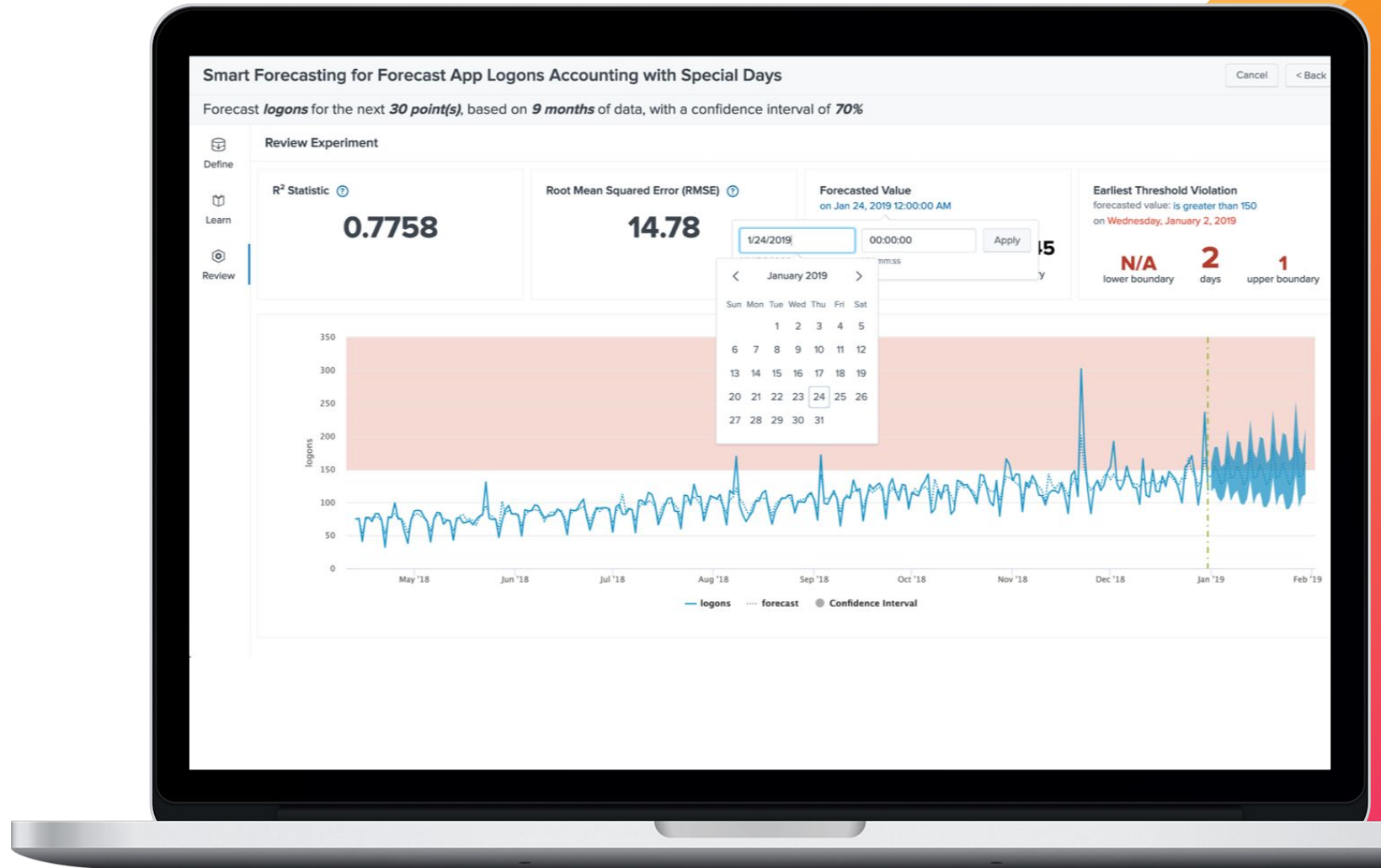
Custom  
ML



# Time Series Forecasting Assistant



# Smart Forecasting Assistant





# Clustering

---



**T-Mobile**

# Predict future outages by forecasting congestion rates of cell towers.



**IT1722 - Predict Real World Outage using Splunk MLTK**

---

**Wednesday, October 23, 12:30 PM - 01:15 PM**

Vijay Veggalam, Member of Technical Staff, T-Mobile  
Gintaras Gaigalas, Sr. RF Engineer, T-Mobile



**Aflac**

# Insurance Fraud Detection



**SEC1904 - The Duck Test: Leverage Machine Learning to Remediate  
Fraud in Huge Datasets**

---

**Thursday, October 24, 01:00 PM - 01:45 PM**

Matthew Harper, Director, Cyber Crime Prevention, Aflac

# MLTK Offers



Cluster Numeric  
Events Assistant



Custom  
ML

# Cluster Numeric Events Assistant



The screenshot displays the 'Cluster Numeric Events' interface. At the top, it says 'Partition events with multiple numeric fields into clusters.' Below this is the 'Assistant Settings' section, which includes a search bar with the query: `| inputlookup disk_failures.csv | search SMART_1_Raw=* | sample 1000`. It shows 1,000 results from a specific time range. The 'Preprocessing Steps' section shows 'StandardScaler' and 'PCA' are selected. The PCA settings include 'Preprocess method' set to 'PCA', 'Select the fields to preprocess' set to 'SS\_SMART\_1\_Raw, ... (5)', and 'K (# of centroids)' set to '3'. A green 'Apply' button is visible. Below this, the 'Algorithm' is set to 'K-means', 'Fields to use for clustering' are 'PC\_1, PC\_2, PC\_3 (3)', and 'K (# of centroids)' is '3'. The 'Save the model as' field contains 'example\_hard\_drives'. At the bottom, there are buttons for 'Cluster', 'Schedule Training', 'Open in Search', and 'Show SPL'. On the right side, a 'Cluster Visualization' window is open, showing a 3x3 grid of scatter plots for PC\_1, PC\_2, and PC\_3. The plots show data points colored by cluster: Cluster 0 (blue), Cluster 2 (yellow), and Cluster 1 (red). A legend on the right of the visualization identifies the clusters.



# What's New?

---



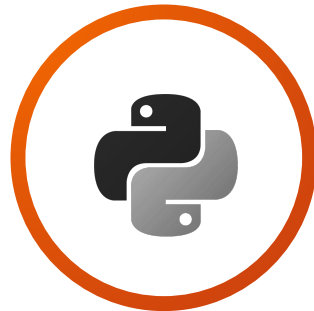
# What's New?

## MLTK 5.0



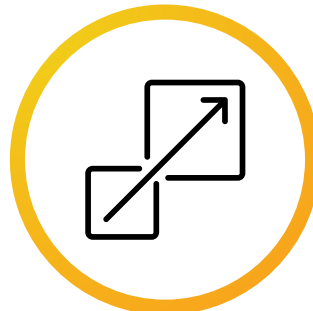
New smart assistants for forecasting and outlier detection, new algorithms, UI refresh, more examples

## PSC 2.0



Upgrade to Python 3.7 for MLTK 5.0 and later versions. Networkx Library Support

## Large Scale Splunk ML PoC



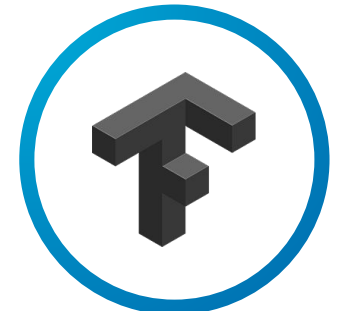
Visit us in the **Innovation Booth** to know more!

## ML on Splunk Cloud Platform



ML-SPL on new Splunk Cloud Platform (SCP), ML-APIs on SCP

## TF 2.0 and PyTorch Support



MLTK Container now supports TensorFlow 2.0 and Pytorch (Only PS Offering).



# How to Reach Us?

---

# Reach out to us for feedback and suggestions



[mlprogram@splunk.com](mailto:mlprogram@splunk.com)



splunk>

# Thank

# You



Go to the .conf19 mobile app to

**RATE THIS SESSION**

