

Machine Learning @ Splunk 2019

The Splunk Machine Learning Toolkit in Action

Iman Makaremi & Harsh





#### 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.

splunk> .conf

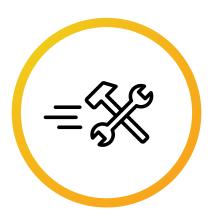
## Agenda



.conf Talk Analysis



ML at Splunk





MLTK in Action

What's New?





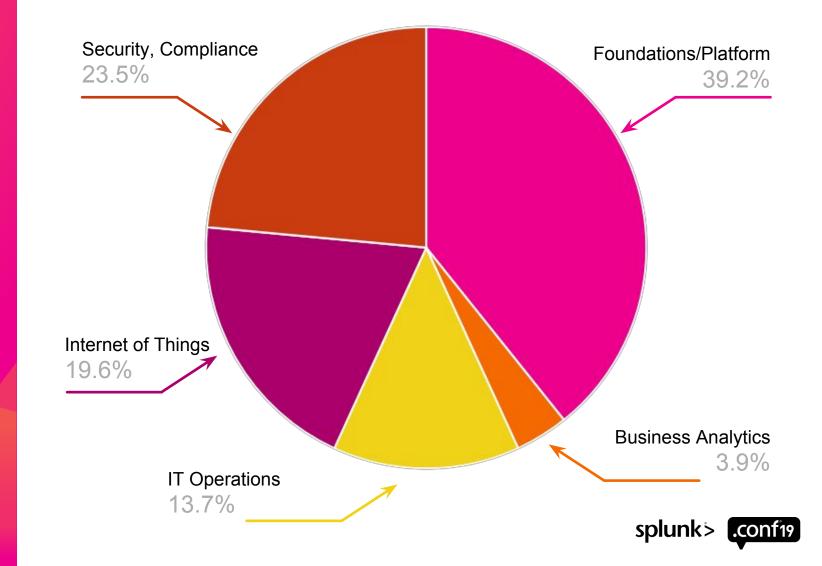


## **Talk Data Analysis**

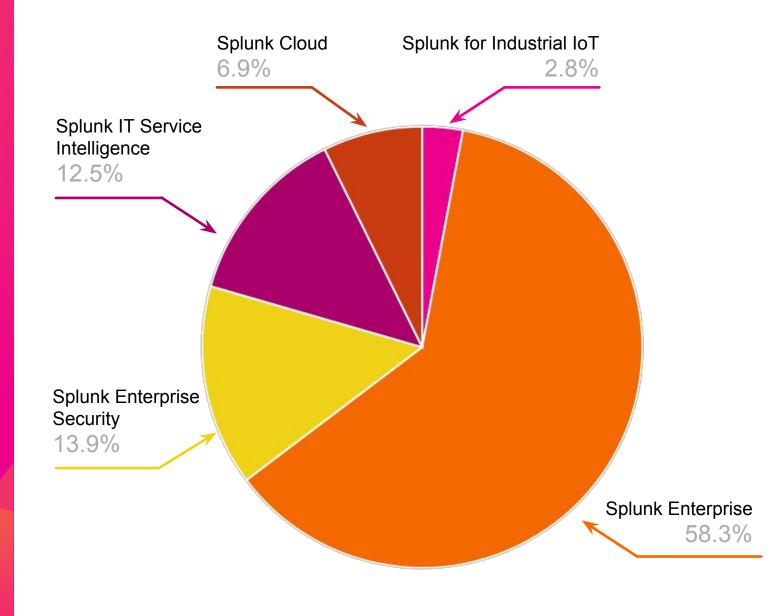
MLTK Impact in 2019 .conf Talks



#### Talks per Sessions Track



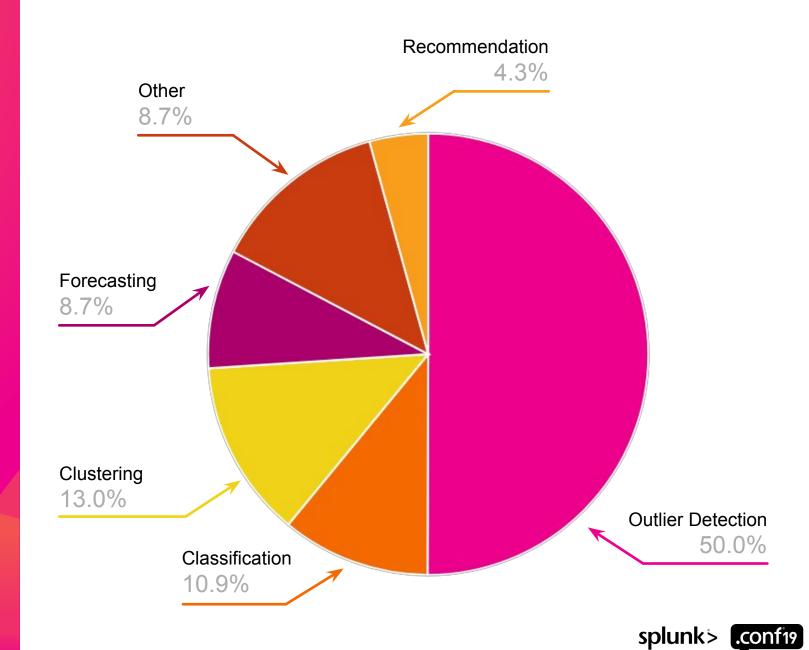
#### MLTK Talks Per Product Category







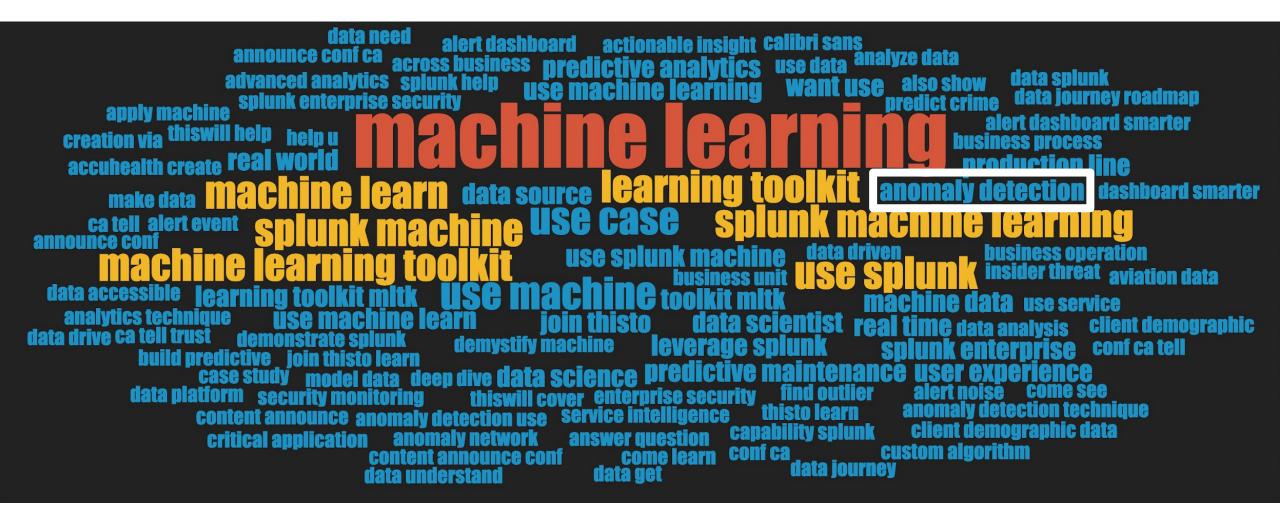
#### Talks Per ML Use Case



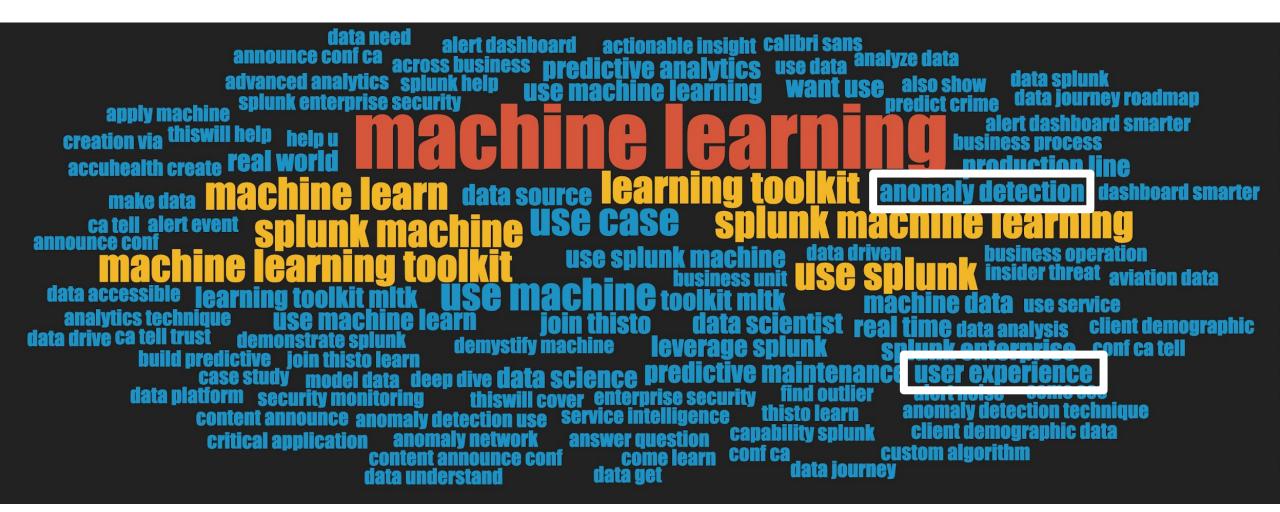


data need actionable insight calibri sans alert dashboard use data <sup>analyze</sup> data announce conf ca across business G data splunk advanced analytics splunk help also show want USC IS **F** data journey roadmap splunk enterprise security nredict crime apply machine alert dashboard smarter creation via thiswill help hein u process accuhealth create MA ashboard smarter Г make data S ca tell alert event announce husiness operation use so nsider threat aviation data data access use service analytics techi TRA scientist client demographic  $\square$ data drive Ca dem demystify machine conf ca tell IAVATA SA build predictive case stu deen dive come see data platform terprise security detection technique and serv Ce ligence content SA client demographic data capability splunk question critical application answer custom algorithm conf ca come learn unce conf data journey data get

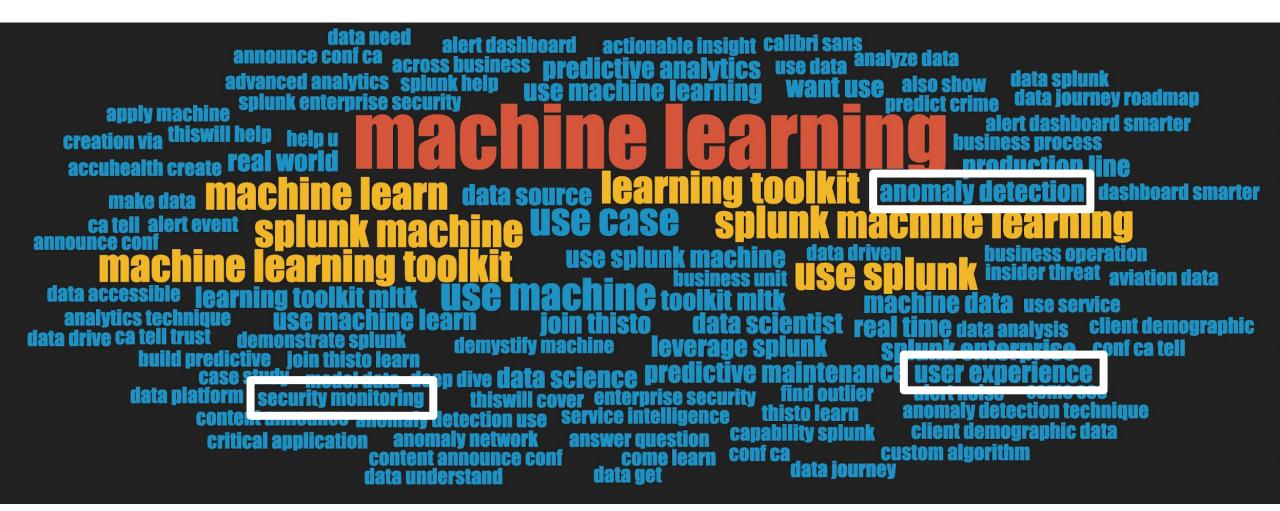




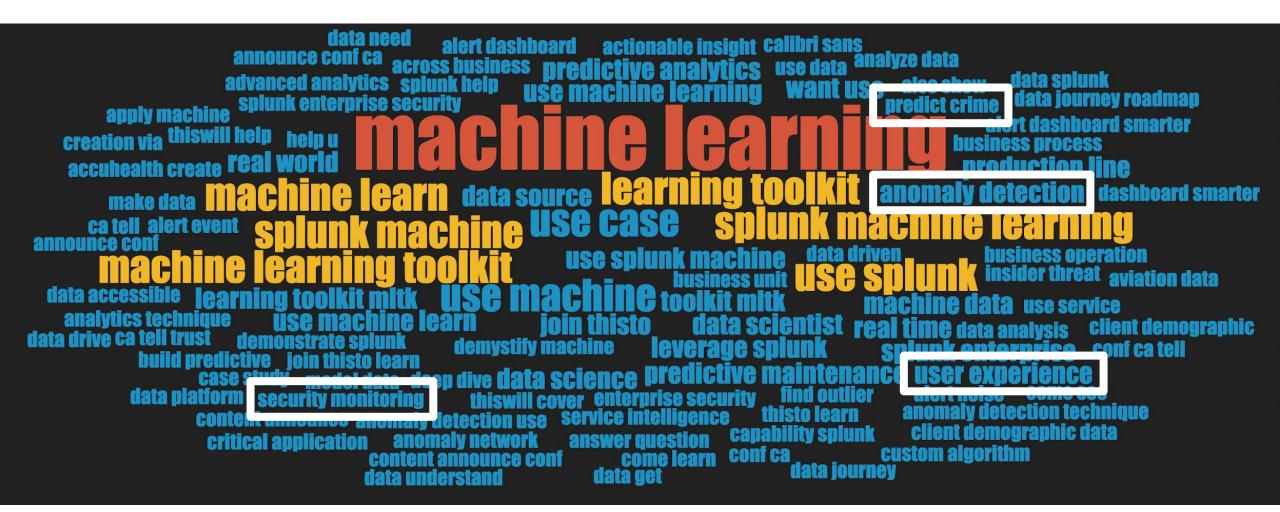




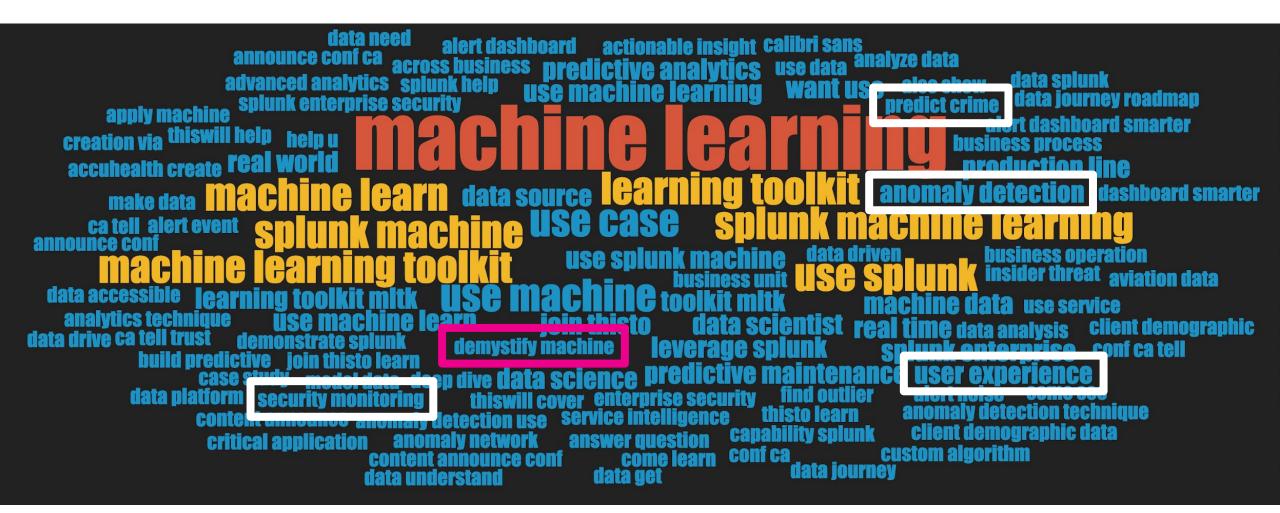




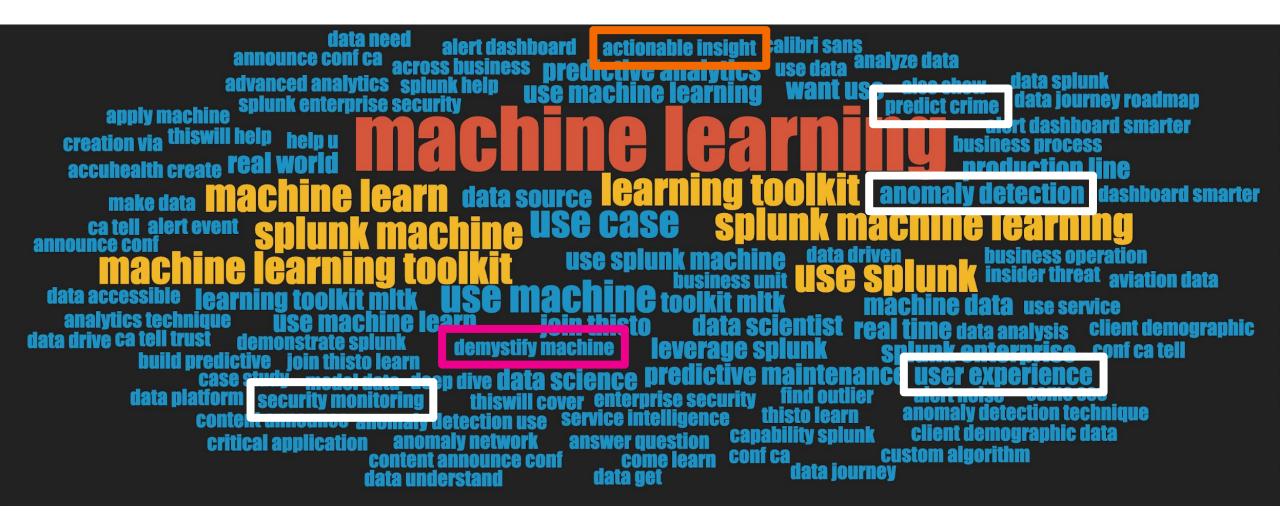
















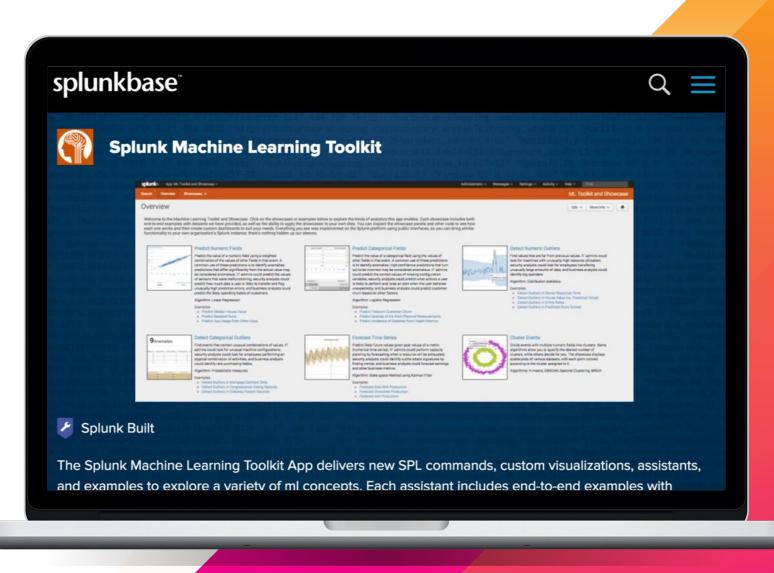
## **MLTK**?!



## What?

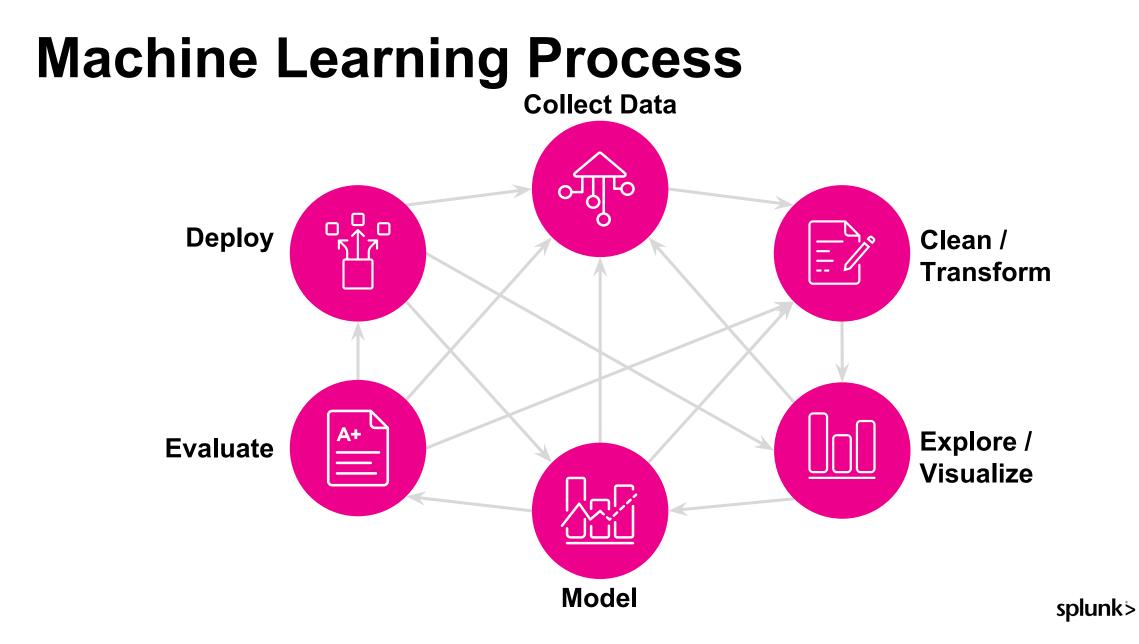
#### Splunk has a Machine Learning Toolkit App!

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



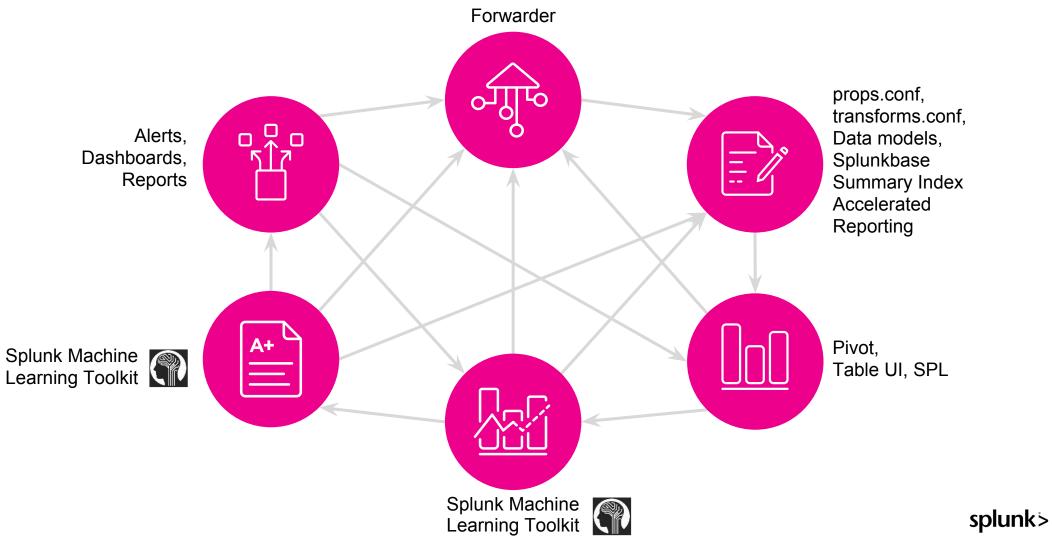


.conf19



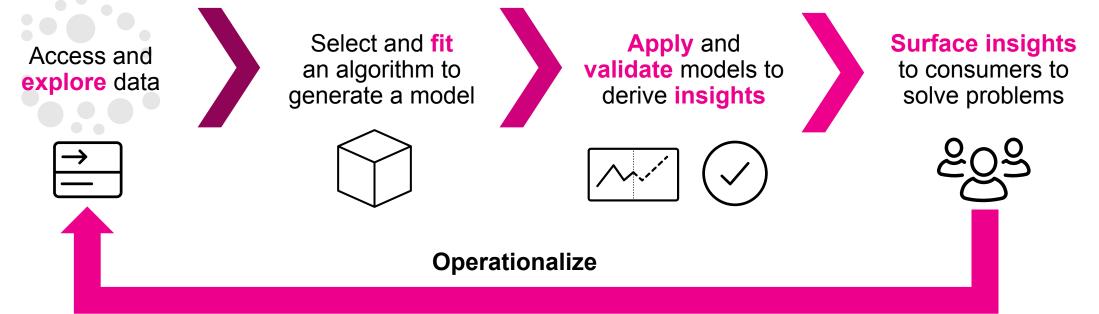
.conf19

#### **Machine Learning Process with Splunk**



## Splunk Helps Answer Questions with AI and ML

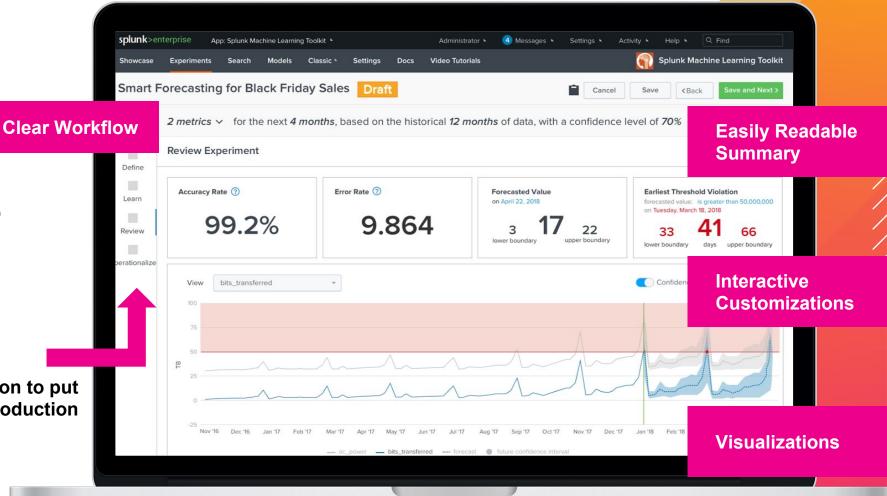
- Identify a Problem: < Stuff in the world> requires big time and money investment.
- Build a Solution: Build ML model to forecast <possible incidents>, act preemptively and continuously learn.





#### MLTK Delivers Easy-to-Use Outcome Assistants

Easy Automation to put ML into Production





## 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

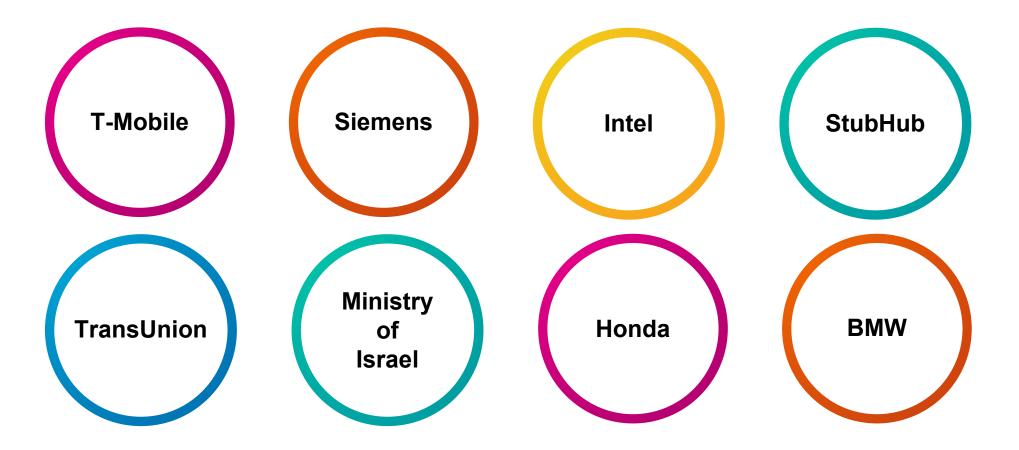
- Early access to new and enhanced Machine Learning features
- Opportunity to shape the development of the product
- Complimentary assistance in operationalizing a production quality ML model

#### Machine Learning Customer Advisory Program FAQs

What is the Machine Learning Customer Advisory Program?	( + )
Are there examples from the advisory program?	+
This program is freewhat's the catch?	+
This sounds interesting! How do I know if I qualify to apply?	+
Anything else I should know?	+
I meet the criteria and am interested in applying! What's next?	( + )
I don't meet the criteria for the advisory program, but am interested in leveraging Splunk for machine learning. What options do I have?	+



## **ML Advisory Customers**





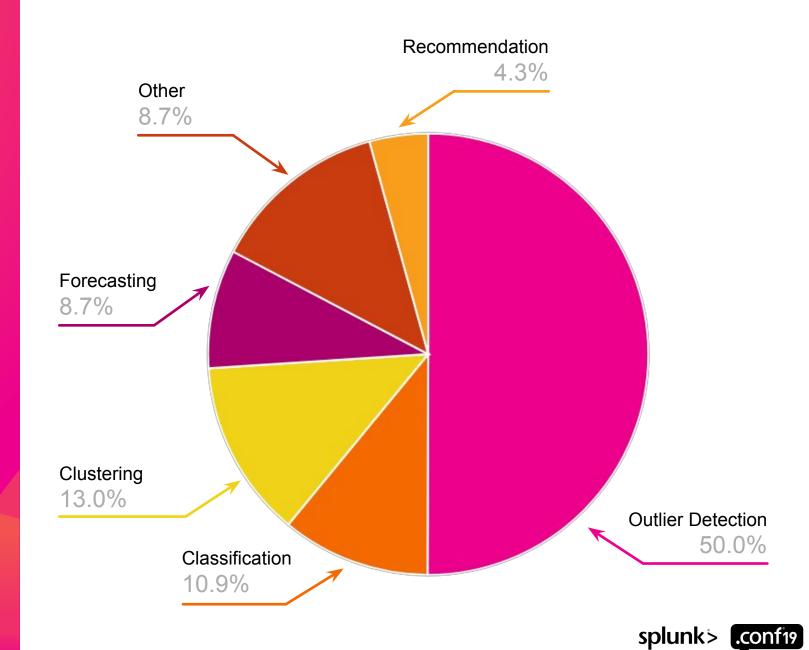


## **MLTK in Action**

Sensor Sensei



#### Talks Per ML Use Case





## **Outlier Detection**

Hore brain, Surgery

311



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



# **T** · · Mobile · **T-Mobile**

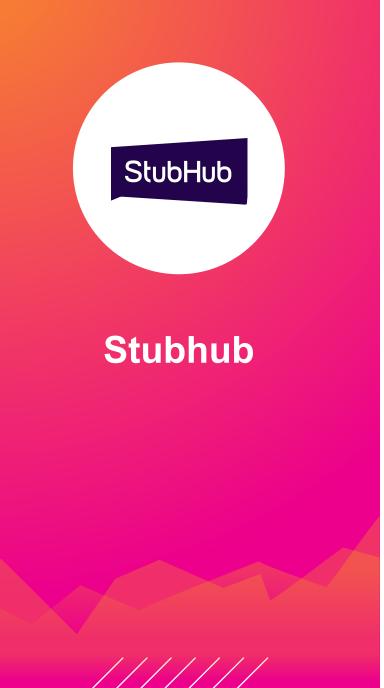
## Identifying service-impacting events across the deployed enterprise applications.

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





# 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



#### 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



Affac.

Aflac

#### SIEMENS

#### 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



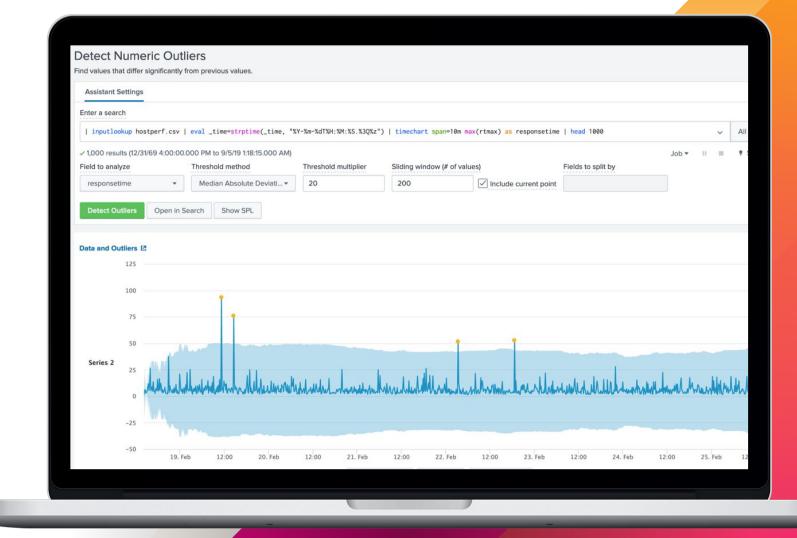
Smart Outlier Detection Assistant



Custom ML



#### Numeric Outlier Detection Assistant





#### Smart Outlier Detection Assistant

	e Experiments Search Models	Classic - Settings Docs & Video Tutoriais &	Splunk Machine Learning Tool
Smar	t Outlier Detection: Find And	malies in Supermarket Purchases	Cancel < Back Next >
Detect	outliers in <i>quantity</i> , split by <i>shop_id</i> , u	sing <i>Auto</i> distribution with a threshold of <i>0.0069</i>	
Ð	Learn Data		
Define	> Initial data - search	Input Output Evaluate	SPL
Ŭ	✓ Detect Outliers		
Learn	Field to analyze 🕜	Mode  Outlier tolerance threshold Automatic Manual O.0001  1 0.0069 Apply	Total Outliers 🕜
٢	quantity -	Automatic Manual 0.0001 1 0.0069 Apply	398
Review	Split by fields	View Top 3 groups with most outliers 💌 s1, s2, s3 💌	Combined Charts
	shop_id (1) 👻		
	Distribution type 🕜	show histogram 💽 show o	outlier area 🛛 🚺 show outliers 📟
	Auto 👻	02	30k
	Outlier tolerance threshold	0.16	24k
	0.0001 1 0.0069		
	Notes	0.12 Visio	18k
	(optional)	(Approx 0.12 (Approx 0.12 (Appr	→ shop_id=s' → shop_id=s' 12k
			12k 🔶 shop_id=s
		0.04	<b>y:</b> 927
	Detect Outliers		ilityDensity(quantity): 0
	Detect Outliers	Probabi	



#### Smart Outlier Detection Assistant

	e Experiments Search Models	Classic▼ Settings Docs L2 Video Tutorials L2	Splunk Machine Learning Too
Smar	Outlier Detection: Find An	iomalies in Supermarket Purchases	Cancel < Back Next >
Detect	outliers in <i>quantity</i> , split by <i>shop_id</i> ,	using <i>Auto</i> distribution with a threshold of <i>0.0069</i>	
E	Learn Data		
Define	> Initial data - search	Input Output Evaluate	SPL
$\square$	✓ Detect Outliers		
Learn	Field to analyze 🕥	Mode (?)         Outlier tolerance threshold           Automatic         Manual         0.0001         1         0.0069         Appl	Total Outliers ⑦
٢	quantity -		_ 390
Review	Split by fields	View Top 3 groups with most outliers	≡ Combined Charts < Split Charts
	shop_id (1) -		
	Distribution type 🕜	C show histogram C sho	ow outlier area 🚺 show outliers 🚥
	Auto 👻	0.2	30k
	Outlier tolerance threshold	0.16	24k
	0.0001 1 0.0069	(Athread and a second and as	
	Notes	012 Other title	18k ≁◆ shop_id=si
	(optional)		tick
			hop_id=s2
	Detect Outliers	0.04	antity: 927 babilityDensity(quantity): 0
	Detect Outliers	0.04	antity: 927



## **Custom ML**

Custom ML Outlier Detection Blog (link)

Add Custom ML Alg (link)



## Forecasting

Looking for trouble.



# **T**··Mobile• **T-Mobile**

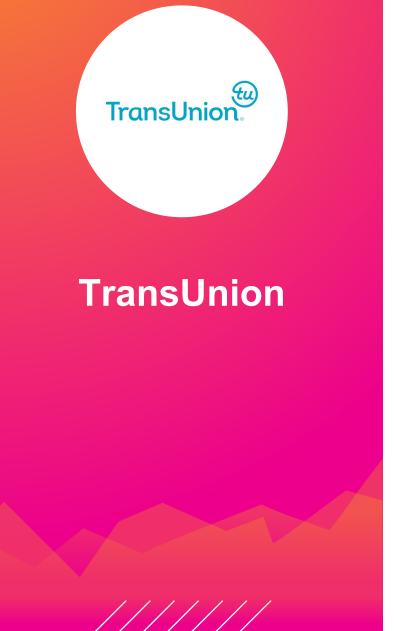
## Predict future outages by forecasting congestion rates of cell towers.

#### 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





## 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

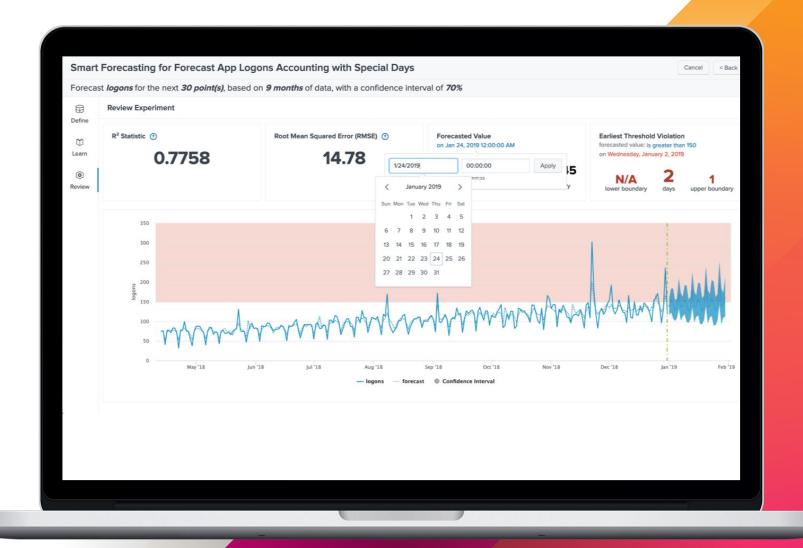


## Time Series Forecasting Assistant

Assistant Settings								
Enter a search								_
inputlookup internet_traf	fic.csv   timechart s	<pre>span=120min avg("bits</pre>	s_transferred")	ferred   eval bits_tran	sferred=round(bits_transf	erred)	~	
✓ 616 results (12/31/69 4:00:00.0 Algorithm	000 PM to 9/5/19 2:52:19 Field to forecast	0.000 AM)				► doL	11 11	
Kalman Filter 🔹	bits_transferred	•						
Method	Future Timespan	Holdback	Confidence Interval	Period				
LLP5 (combines LLT and	112	112	95	(optional)				
Forecast 12					1	holdback		
15G 12.5G 10G	1.10	m	1001-11	111 111		holdback		
15G 12.5G 10G	AAAAA					holdback		A DESCRIPTION OF THE OWNER OWNER OF THE OWNER OWNER OWNER OF THE OWNER OWNE OWNER
15G 12.5G 10G						holdback		
15G 12.5G 10G 7.5G 5G 5G 5G						holdback		Character and a second second
12.5G						holdback		A Distance of the second



## Smart Forecasting Assistant





## Clustering

Can you SPL?



# **T**··Mobile• **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



### 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

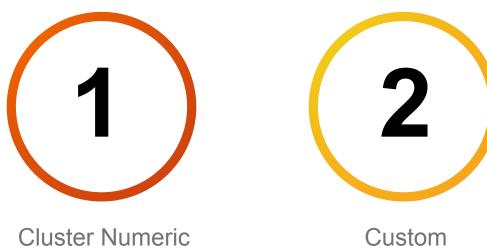
Matthew Harper, Director, Cyber Crime Prevention, Aflac



Affac.

Aflac

## **MLTK Offers**



Events Assistant

Custom ML



### Cluster Numeric Events Assistant

Assistant Settings		Cluster Visualization 12			
Enter a search   inputlookup disk_failures.csv   search SMART_1_Raw=*   sample 1000 1,000 results (12/31/69 4:00:00.000 PM to 9/5/19 3:22:48.000 AM)</th <th>10 2 0</th> <th>9 5 PC.2 30</th> <th>15 <u>9</u> 10 <sup>PC_3</sup> 20</th> <th>•</th>		10 2 0	9 5 PC.2 30	15 <u>9</u> 10 <sup>PC_3</sup> 20	•
Preprocessing Steps          StandardScaler         PCA         Preprocess method         PCA         Apply	Select the fields to preprocess. K (# c SS_SMART_1_Raw, (5) ¥ 3	-10 -20 20 15 5 0 5 -5 -5 20 20 20 20 20 20 20 20 20 20 20 20 20	· · · ·		• Clus • Clus • Clus
+ Add a step Preview Resul		-10			
Algorithm K-means	Fields to use for clustering     K (# of centroids)     Save the model as       PC_1, PC_2, PC_3     (3) •     3     example_hard_drives				
Cluster Schedule Training	Open in Search Show SPL				







## What's New?

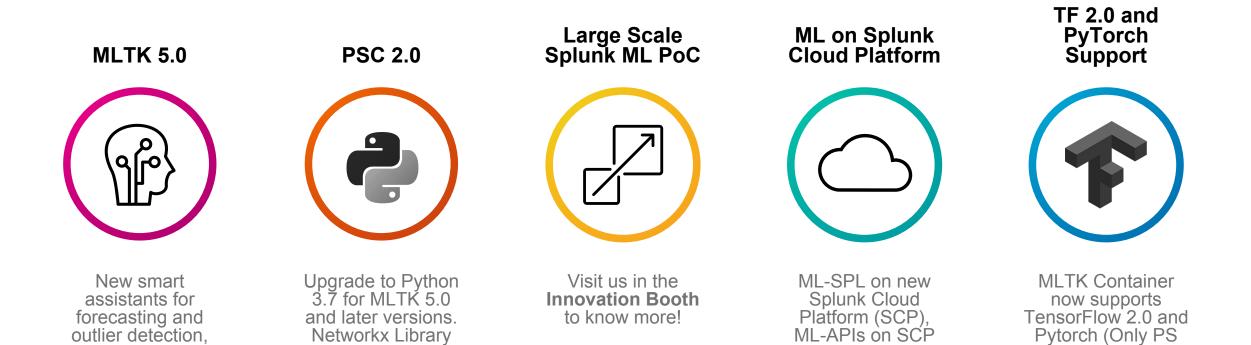


## What's New?

new algorithms,UI

refresh, more examples

Support



splunk > .conf19

Offering).



## How to Reach Us?



## Reach out to us for feedback and suggestions

mlprogram@splunk.com



© 2019 SPLUNK INC

 $\bigcirc$ 



## Thank



Go to the .conf19 mobile app to

**RATE THIS SESSION**