

splunk® > live!

APRIL 13, 2017 | SAN FRANCISCO

splunk>live!

APRIL 13, 2017 | SAN FRANCISCO

Advanced Analytics and Machine Learning in Splunk

Pierre Brunel | Sr. Sales Engineer

Safe Harbor Statement

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. © 2017 Splunk Inc. All rights reserved.

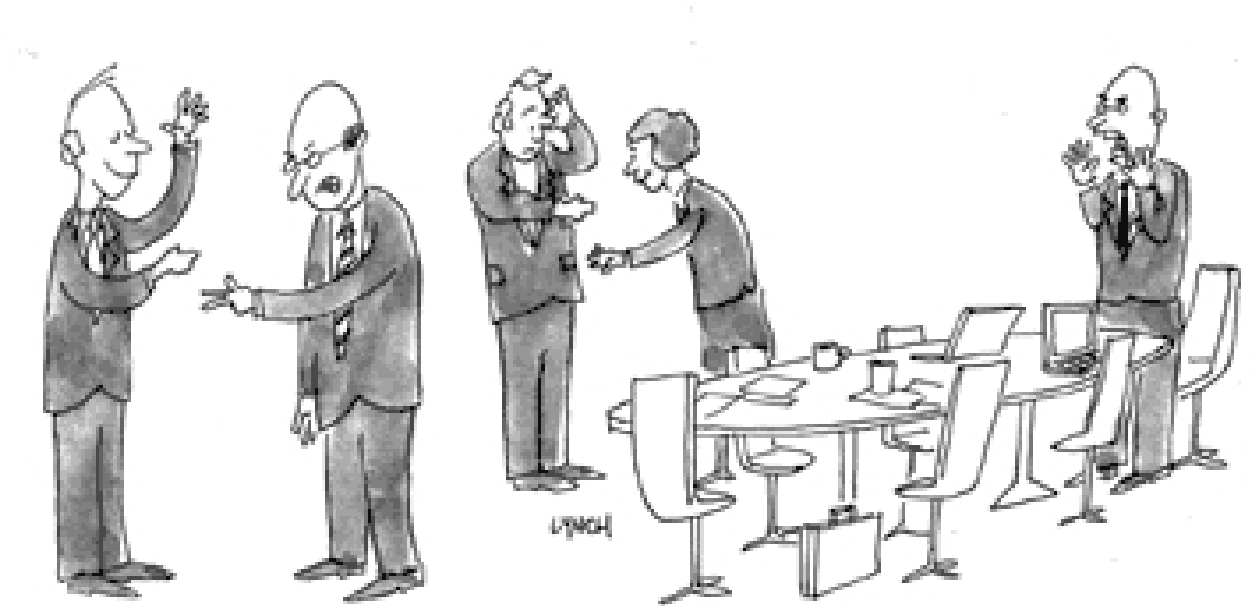
Why Machine Learning?



Humans are good at
learning, but we get lost
in volume and details...

- ▶ Improve decision making
- ▶ Uncover hidden trends or relationships
- ▶ Alert on deviations
- ▶ Forecast or anticipate incidents

All of this requires diverse data from across many silos. Lots of unstructured, real-time data.



"Now that we've made our first billion, we gotta stop making decisions by playing 'paper, scissors, rock.'"

Run the Business in Real Time

Security Operations Center

IT Operations Center

Business Operations Center

Descriptive
(BI Tools, Data Lakes)

Grey space

Predictive
(Models)



Data From the Past

$T - \text{a few days}$



Real-Time Data

$T + \text{a few days}$

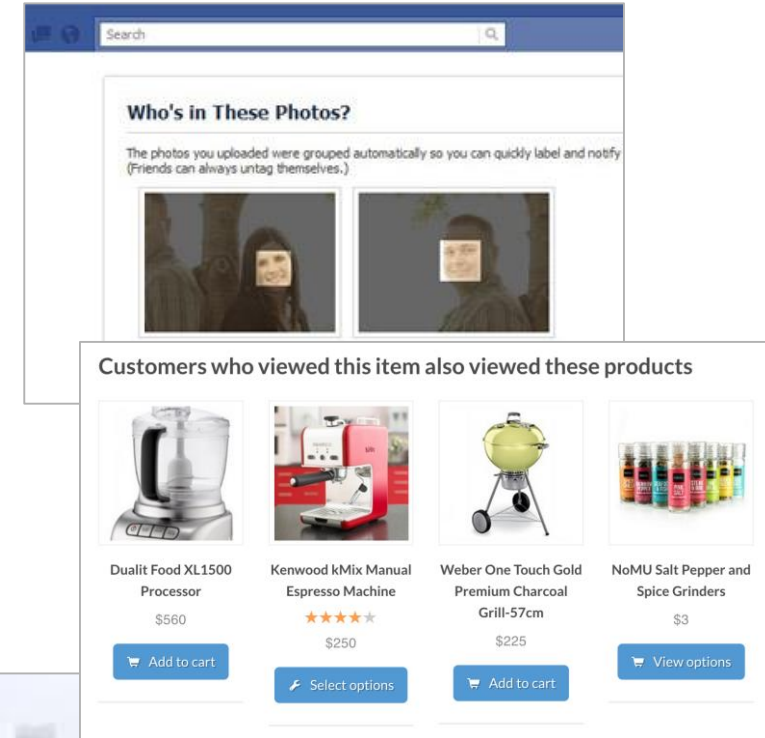


Statistical Forecast

What is Machine Learning?

ML is All Around You!

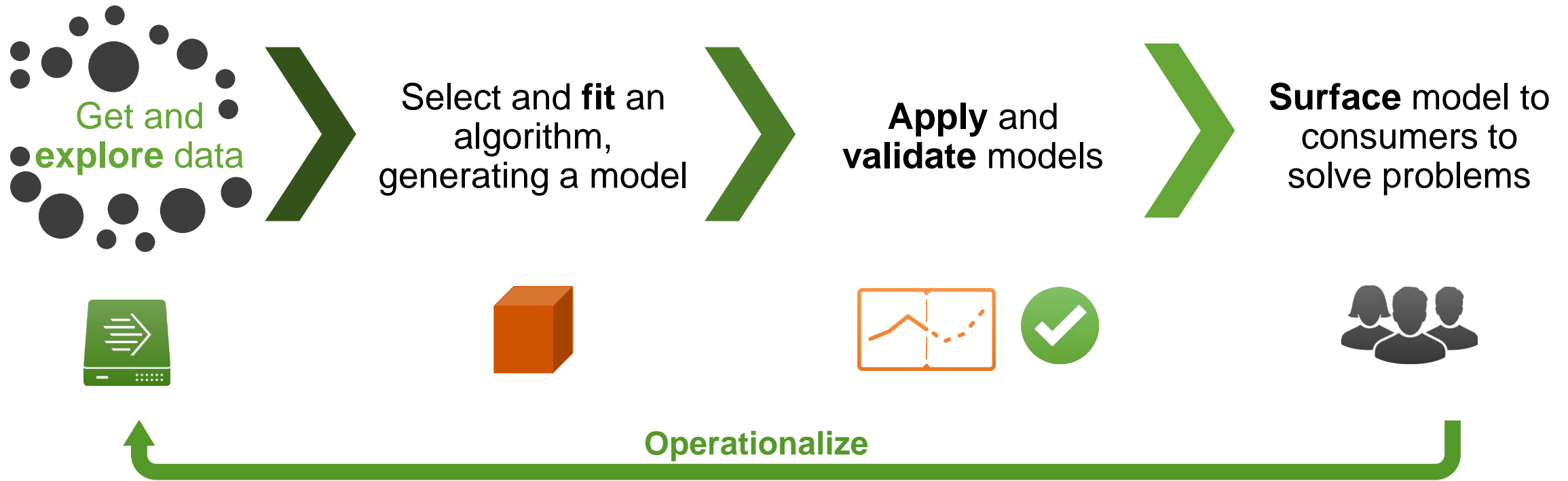
- ▶ **Face detection:** find faces in images
- ▶ **Spam filtering:** identify SPAM messages
- ▶ **Shopping recommendations:** predict what customers would like to buy
- ▶ **Fraud detection:** identify credit card transactions which may be fraudulent in nature
- ▶ **Weather forecast:** predict whether or not it will rain tomorrow; estimate daily max/min



The ML Process

Problem: <Stuff in the world> causes big time & money expense. Value Hypothesis

Solution: Build ML model to forecast <possible incidents>, act pre-emptively & learn



Splunk's Machine Learning Tour

Overview of ML at Splunk



CORE PLATFORM
SEARCH



PACKAGED PREMIUM
SOLUTIONS



MACHINE LEARNING
TOOLKIT

splunk > Platform for Operational Intelligence



Search Includes Machine Learning

Core Platform Search is a powerful and highly flexible interface built with ML

splunk> App: Search & Reporting ▾ Messages ▾ Settings ▾ Activity ▾ Help ▾ Find

Search Datasets Reports Alerts Dashboards Search & Reporting

🔍 New Search Save As ▾ Close

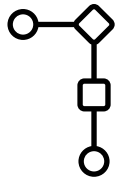
index="oidemo" | stats count(bytes) as TRAFFIC by host | anomalydetection TRAFFIC Last 60 minutes ▾ 🔍

✓ 51,139 events (3/29/17 4:02:00.000 PM to 3/29/17 5:02:44.000 PM) No Event Sampling ▾ Job ▾ || ■ ➔ 🖨 ⬇ ⚡ Smart Mode ▾

Events Patterns Statistics (2) Visualization

20 Per Page ▾ ⚡ Format ▾ Preview ▾

host ▾	TRAFFIC ▾	log_event_prob ▾	max_freq ▾	probable_cause ▾	probable_cause_freq ▾
webserver-01	4217	-2.3168	0.75587	TRAFFIC	0.09859
webserver-03	4284	-2.3168	0.75587	TRAFFIC	0.09859



Splunk IT Service Intelligence

One of several Premium Solutions with Packaged ML



Define services,
entities and KPIs



Monitor and
troubleshoot



Analyze
and detect

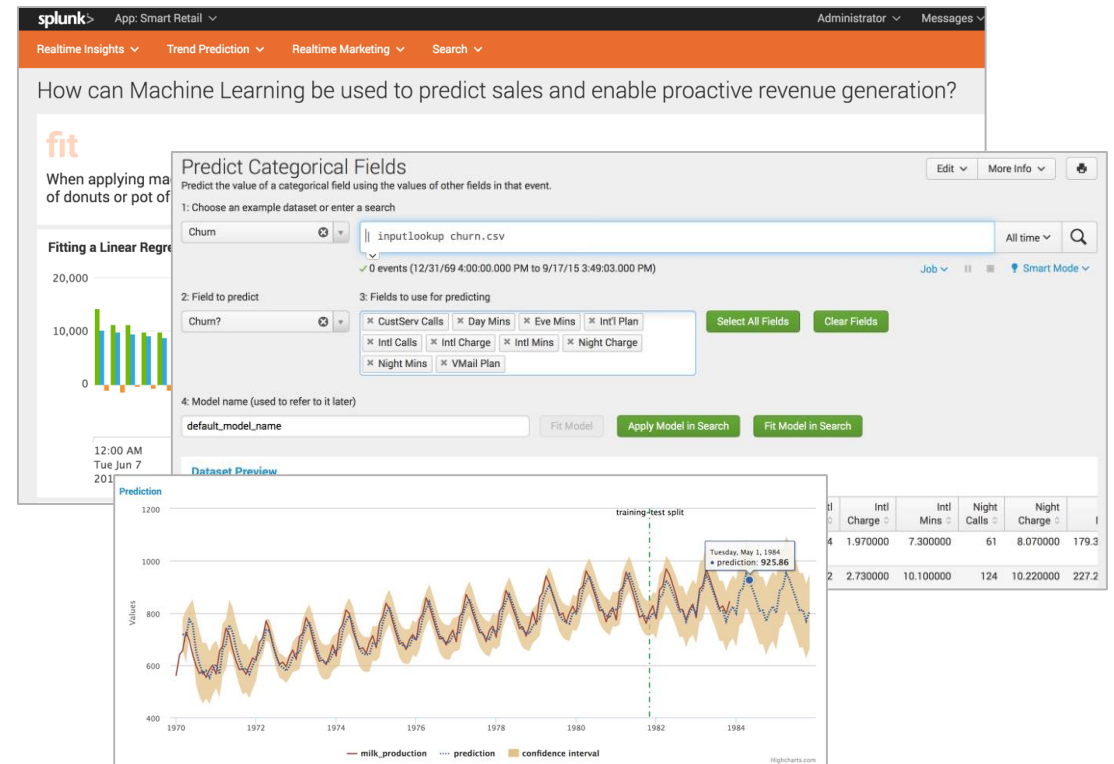
Data-Defined, Data-Driven Service Insights
Adaptive Thresholds and Anomaly Detection



Splunk Machine Learning Toolkit

Extends Splunk platform functions and provides a guided modeling environment

- ▶ **Assistants:** Guided model building, testing and deployment for common objectives
- ▶ **Showcases:** Interactive examples for typical IT, security, business and IoT use cases
- ▶ **Algorithms:** 25+ standard algorithms included with the toolkit
- ▶ **ML Commands:** New SPL commands to fit, test and operationalize models
- ▶ **Python for Scientific Computing Library:** Access to 300+ open source algorithms

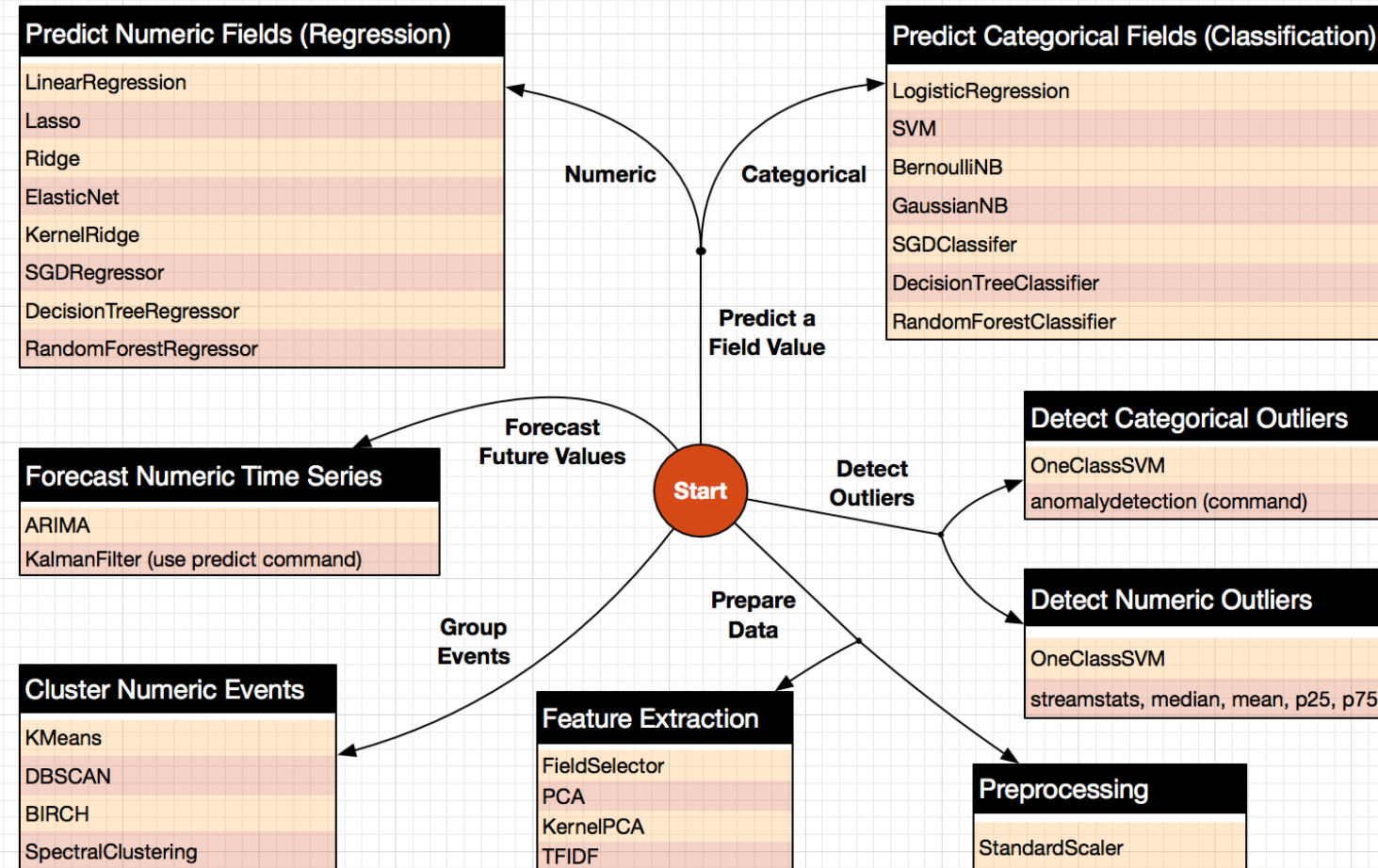


Build custom analytics for any use case

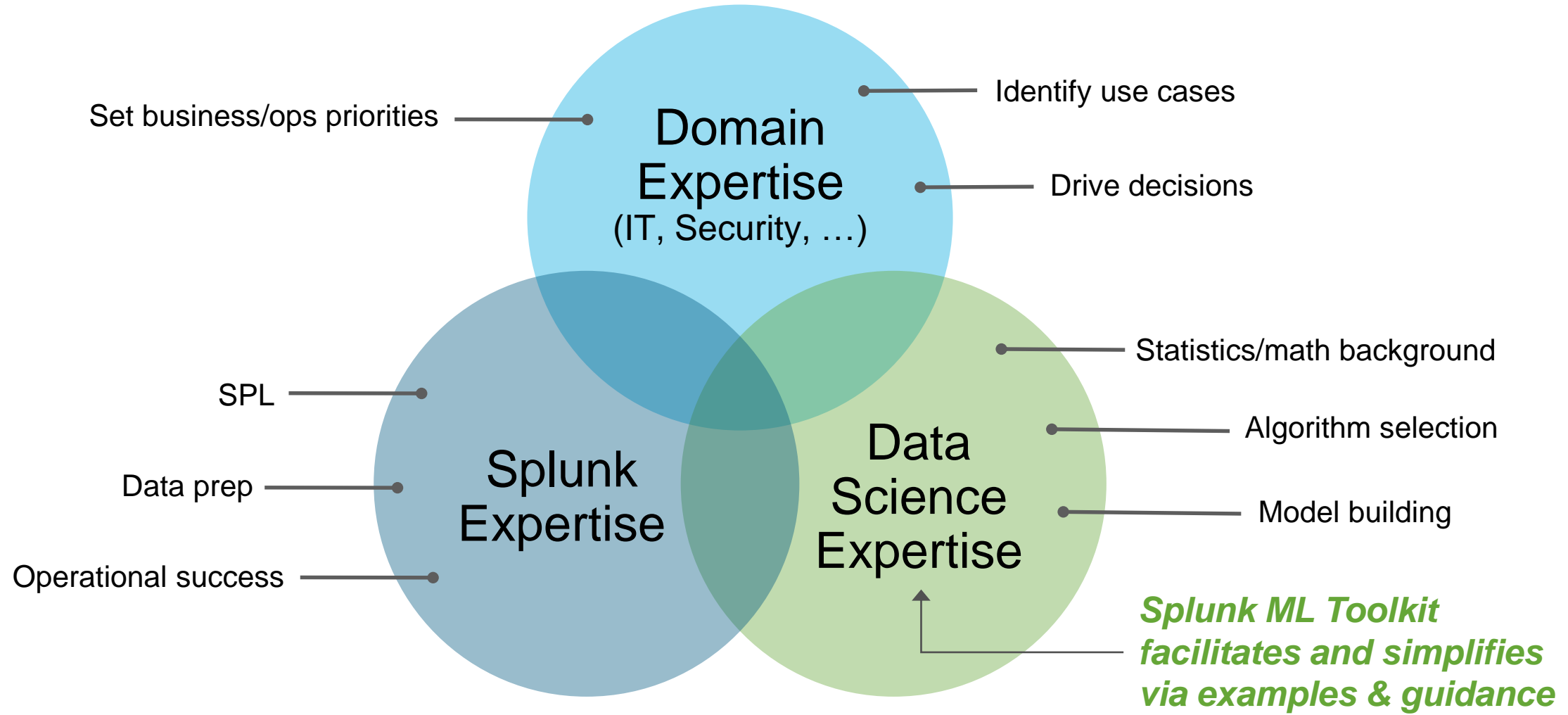
Algorithms Supported (v2.0)



Machine Learning Toolkit



Custom Machine Learning – Success Formula

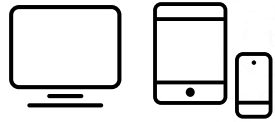


Splunk Architecture and Machine Learning

Continuous Data Ingest at Scale



Industrial Data
SCADA, AMI, Meter Reads



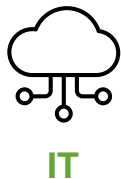
Native Inputs
TCP, UDP, Logs, Scripts, Wire, Mobile

Consumer and Mobile Devices

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>live!

Sense and Respond

Every Search Can Use
Machine Learning



Industrial Assets



Consumer and
Mobile Devices



OT



Real Time

splunk>

Search

Alert



Flash lights



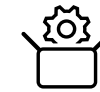
Email

Send an
email



Tickets

File a
ticket



Third-Party
Applications

Trigger
process flow

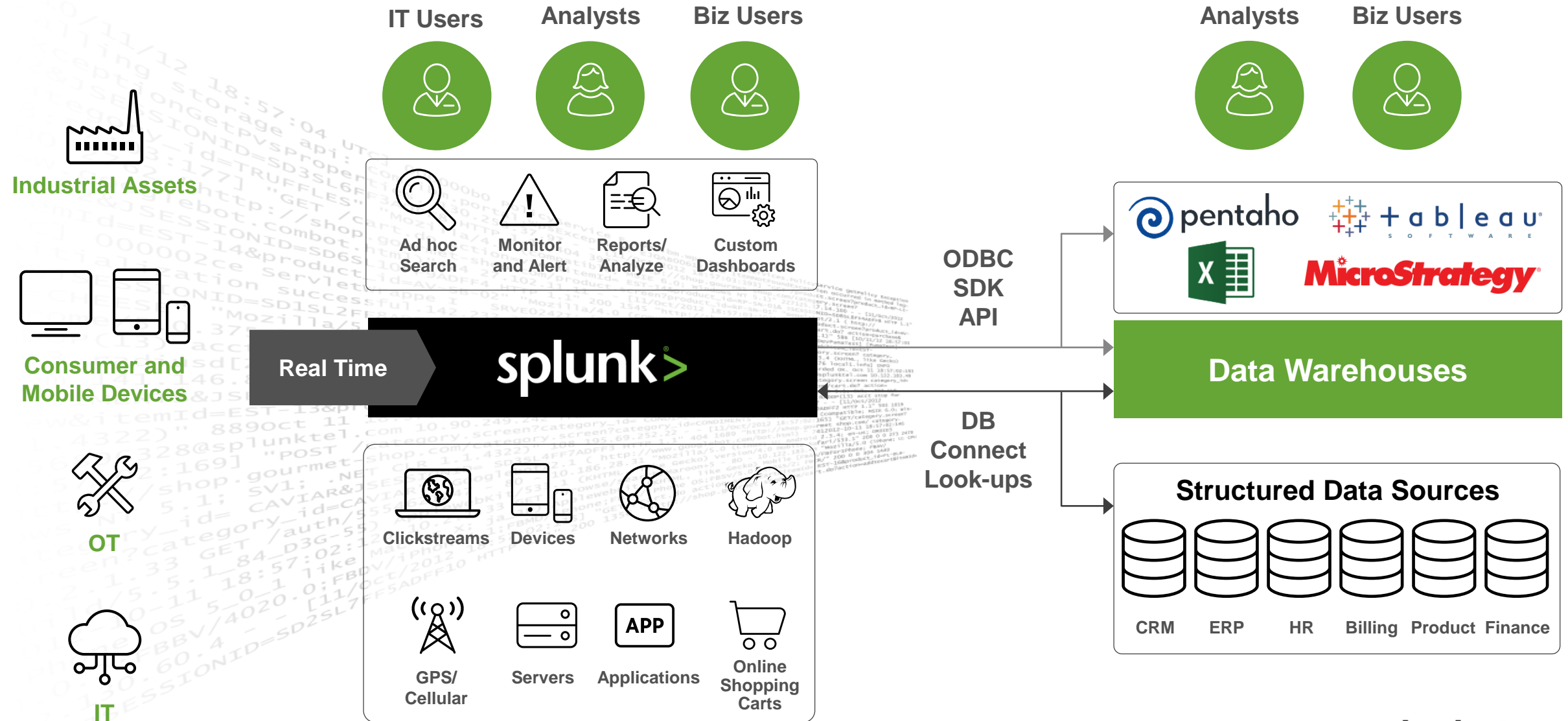


Smartphones
and Devices

Send a text

splunk>live!

Splunk: Data Fabric



Show me the ML!

ML Use Cases and Customer Stories

Machine Learning Customer Success



Network Incident Detection
Service Degradation Detection



Security / Fraud Prevention



Cell Tower Incident Detection
Optimize Repair Operations



Prioritize Website Issues
and Predict Root Cause



Entertainment
Company

Predict Gaming Outages
Fraud Prevention



Machine Learning
Consulting Services



Analytics App Built
on ML Toolkit

Optimizing operations and business results

ML Toolkit Customer Use Cases



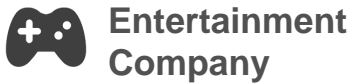
- ▶ Reducing customer service disruption with early identification of difficult-to-detect network incidents
- ▶ Minimizing cell tower degradation and downtime with improved issue detection sensitivity



- ▶ Speeding website problem resolution by automatically ranking actions for support engineers



- ▶ Ensuring mobile device security by detecting anomalies in ID authentication



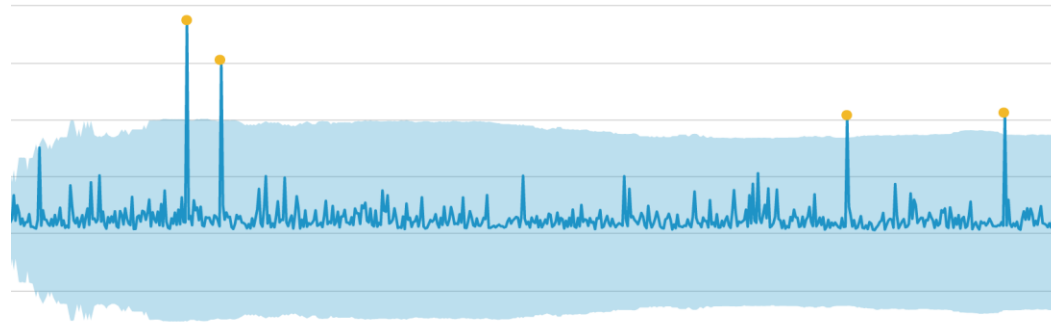
- ▶ Predicting and averting potential gaming outage conditions with finer-grained detection
- ▶ Preventing fraud by Identifying malicious accounts and suspicious activities



- ▶ Improving cell tower uptime and reducing repair truck roles with anomaly detection and root cause analysis

Detect Network Outliers

Reduced downtime + increased service availability = better customer satisfaction



ML Use Case

- ▶ Monitor noise rise for 20,000+ cell towers to increase service and device availability, reduce MTTR

Technical Overview

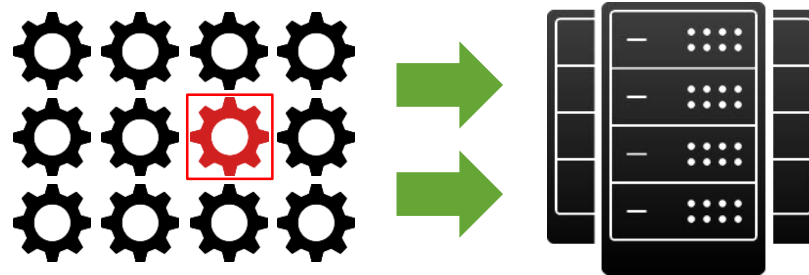
- ▶ A customized solution deployed in production based on outlier detection
- ▶ Leverage previous month data and voting algorithms

“The ability to model complex systems and alert on deviations is where IT and security operations are headed ... Splunk Machine Learning has given us a head start...”



Reliable Website Updates

Proactive website monitoring leads to reduced downtime



ML Use Case

- ▶ Very frequent code and config updates (1000+ daily) can cause site issues
- ▶ Find errors in server pools, then prioritize actions and predict root cause

Technical Overview

- ▶ Custom outlier detection built using ML Toolkit Outlier assistant
- ▶ Built by Splunk Architect with no Data Science background

“Splunk ML helps us rapidly improve end-user experience by ranking issue severity which helps us determine root causes faster thus reducing MTTR and improving SLA”

Wrap Up

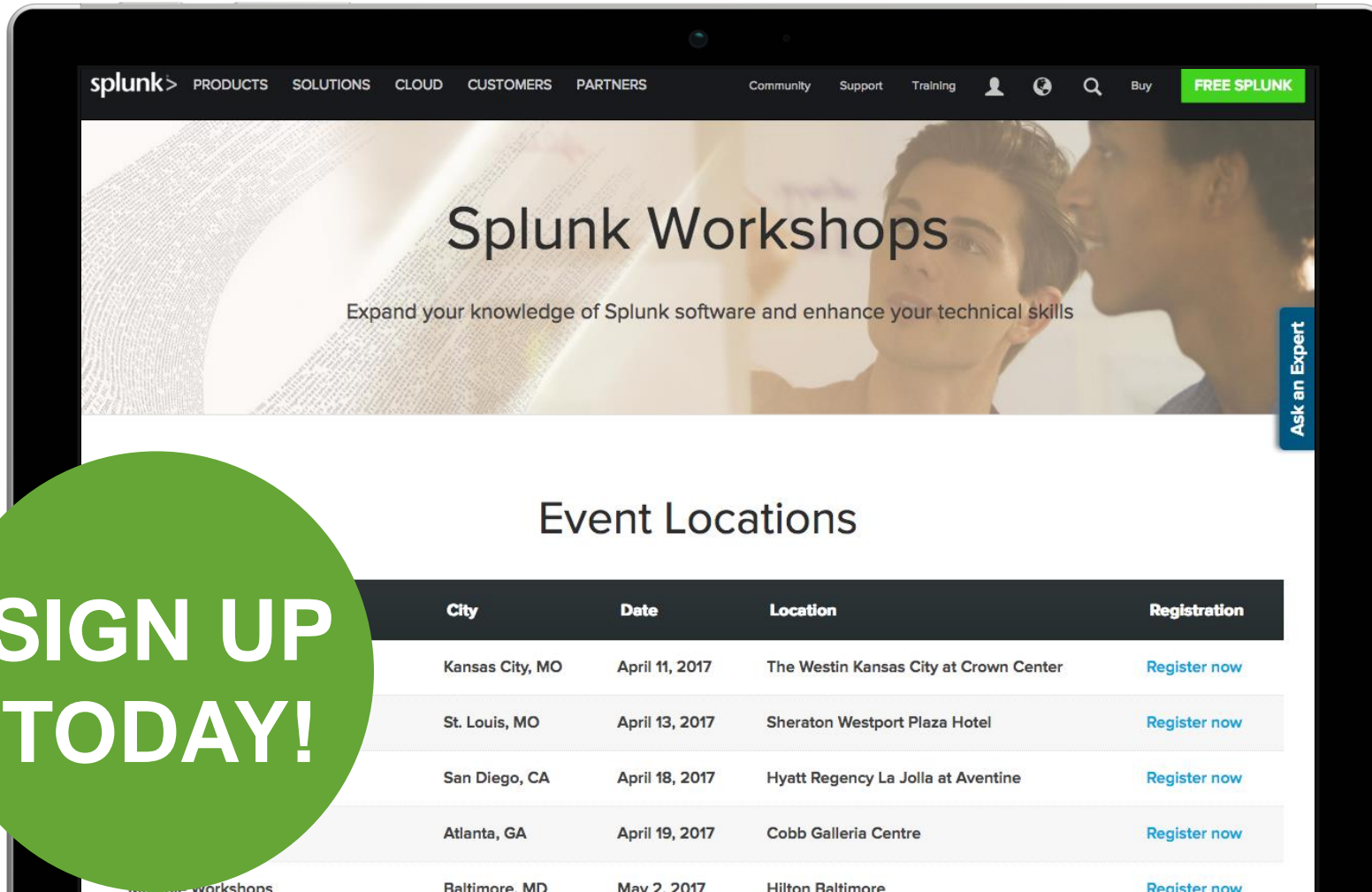
What Else?

- ▶ Get the Machine Learning Toolkit from Splunkbase
- ▶ Go watch Machine Learning Videos on Splunk YouTube Channel
<http://tiny.cc/splunkmlvideos>
- ▶ Go watch the Machine Learning talks from .conf2016:
 - Advanced Machine Learning in SPL with the Machine Learning Toolkit by Jacob Leverich
 - Extending SPL with Custom Search Commands and the Splunk SDK for Python by Jacob Leverich
- ▶ Early Adopter and Customer Advisory Program: mlprogram@splunk.com
- ▶ Field ML Architects: Andrew Stein (astein@), Brian Nash (bnash@)

Workshops: Get Splunk Hands-on Experience

Attend a Splunk Workshop

splunk.com/workshops



**SIGN UP
TODAY!**

May 23: San Francisco

- ▶ **Venue:** Sheraton Fisherman's Wharf
- ▶ **Time:** 8:30am
- ▶ Register Soon!

- splunk.com/workshops

May 25: Sacramento

- ▶ **Venue:** Hyatt Regency Sacramento
- ▶ **Time:** 8:30am
- ▶ Register Now!

- splunk.com/workshops

splunk>live!



.conf2017

.conf2017

The 8th Annual Splunk Conference

SEPT 25-28, 2017

Walter E. Washington Convention Center
Washington, D.C.

SAVE OVER \$450

You will receive an email after registration opens with a link to save over \$450 on the full conference rate.

You'll have 30 days to take advantage of this special promotional rate!

conf.splunk.com



Complete the survey for
your chance to win a
.conf2017 pass

Thank you!