Splunk Your Brain

Analyze Brainwaves using Machine Learning to Predict Activity or Mental State

Brian Guilfoyle | Sr. Sales Engineer

October 2018 | Version 1.0

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.

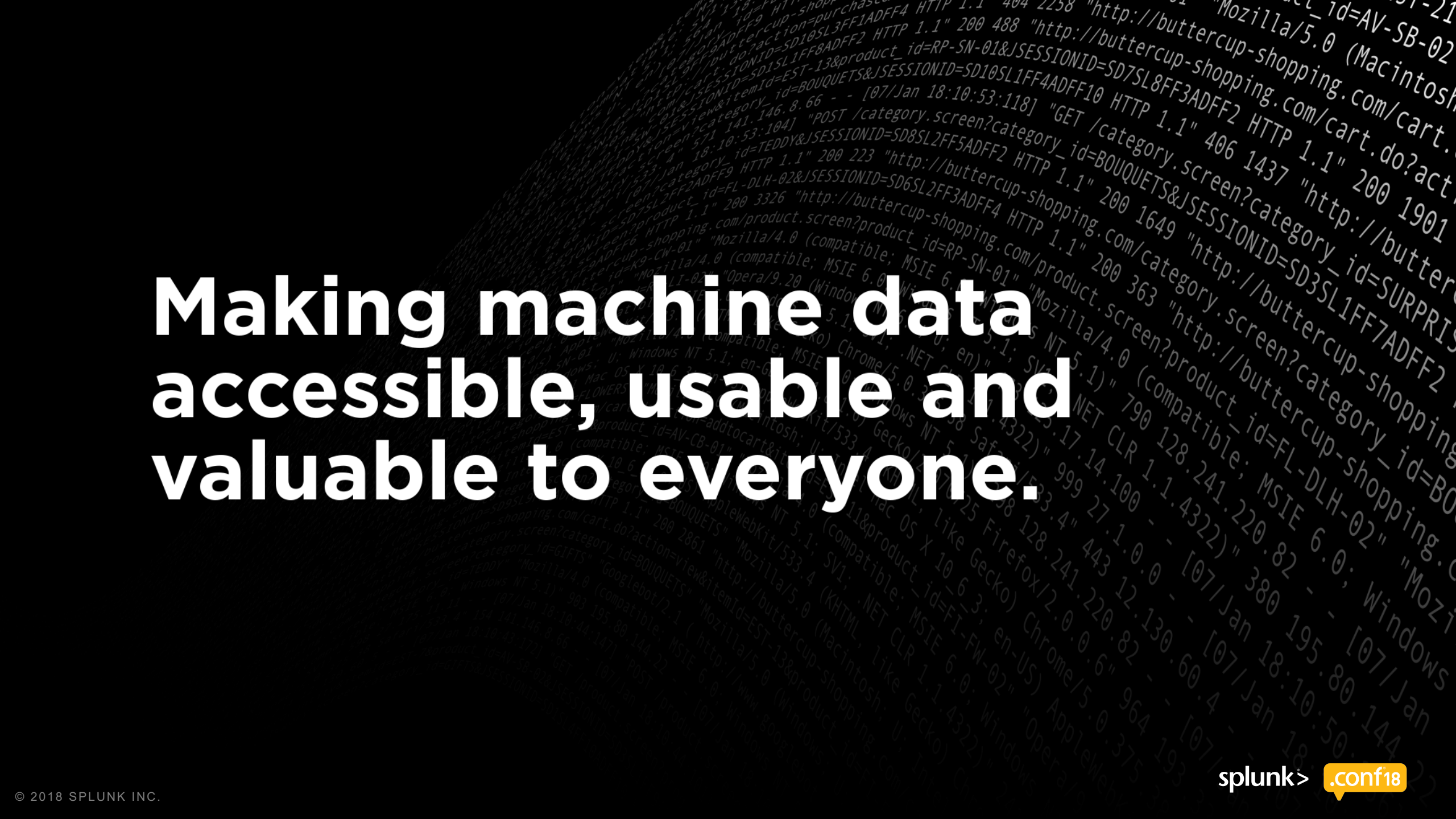
The forward-looking statements made in this presentation are being made as of the time and date of its live presentation. If reviewed after its live presentation, this presentation may not contain current or accurate information. We do not assume any obligation to update any forward-looking statements 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> .conf18

Data Acquisition and Analysis

How to Build an ML Model using EEG data

splunk> .conf18



Making machine data accessible, usable and valuable to everyone.

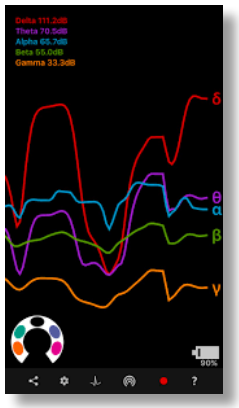
splunk> .conf18

Getting EEG Data In Splunk

Brainwave Data = Machine Data!

Muse EEG Headset

Bluetooth



Mobile Phone with
Muse Monitor App

UDP 5000

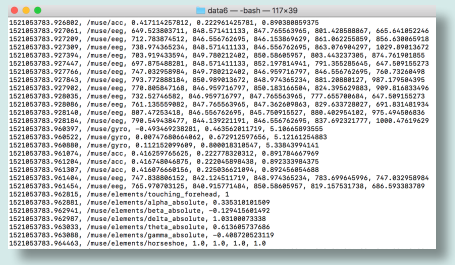
MacBook Pro

MusePlayer

Reroute, replay, and convert Muse data to and from multiple file formats

OSC Server

Splunk Indexer



Log File Output

splunk> .conf18

Exploring EEG Data

What does the raw data look like?

data6 — -bash — 117×39

```
1521053783.775765, /muse/eeg, 513.333312988, 848.571411133, 843.333312988, 790.952392578, 688.608032227
1521053783.775969, /muse/eeg, 561.282043457, 846.556762695, 844.139221191, 857.838806152, 854.212463379
1521053783.776125, /muse/eeg, 586.26373291, 847.765563965, 844.945068359, 871.538452148, 1056.48352051
1521053783.776216, /muse/eeg, 554.029296875, 852.197814941, 849.780212402, 812.710632324, 927.545776367
1521053783.776294, /muse/eeg, 522.600708008, 850.58605957, 848.168518066, 776.043945312, 699.084228516
1521053783.776545, /muse/eeg, 555.641052246, 844.139221191, 840.109863281, 819.157531738, 754.285705566
1521053783.776850, /muse/gyro, 2.03369140625, -0.530853271484, 8.54598999023
1521053783.777085, /muse/gyro, 1.44302368164, -0.95703125, 7.96279907227
1521053783.778281, /muse/gyro, 0.859832763672, -1.3981628418, 7.64129638672
1521053783.778405, /muse/acc, 0.401245117188, 0.237060546875, 0.895629882812
1521053783.778462, /muse/acc, 0.408508300781, 0.240112304688, 0.896118164062
1521053783.778508, /muse/acc, 0.416564941406, 0.242858886719, 0.898315429688
1521053783.842507, /muse/eeg, 602.783874512, 841.721618652, 836.886474609, 863.8828125, 1020.2197876
1521053783.842635, /muse/eeg, 582.637390137, 844.139221191, 839.706970215, 825.604370117, 1021.83148193
1521053783.842888, /muse/eeg, 548.388305664, 848.571411133, 839.304016113, 772.820495605, 753.076904297
1521053783.843505, /muse/eeg, 566.923095703, 846.153869629, 836.886474609, 803.040283203, 717.619018555
1521053783.843577, /muse/eeg, 617.692321777, 843.333312988, 837.289367676, 861.868103027, 952.52746582
1521053783.843628, /muse/eeg, 613.260070801, 846.153869629, 840.915771484, 844.945068359, 1007.72894287
1521053783.843696, /muse/eeg, 560.879150391, 847.765563965, 846.959716797, 784.908447266, 792.564086914
1521053783.843864, /muse/eeg, 562.087890625, 844.945068359, 846.556762695, 796.996337891, 712.783874512
1521053783.843937, /muse/eeg, 618.498168945, 845.347961426, 844.139221191, 855.421264648, 882.820495605
1521053783.843988, /muse/eeg, 646.300354004, 850.58605957, 847.362609863, 861.062255859, 1000.87915039
1521053783.844035, /muse/eeg, 598.754577637, 851.79486084, 846.959716797, 796.190490723, 840.512817383
1521053783.844207, /muse/eeg, 576.593383789, 852.197814941, 844.542114258, 776.043945312, 672.490844727
1521053783.845602, /muse/elements/touching_forehead, 1
1521053783.845673, /muse/elements/alpha_absolute, 0.336772918701
1521053783.847106, /muse/elements/beta_absolute, -0.136177688837
1521053783.847176, /muse/elements/delta_absolute, 1.00698816776
1521053783.847233, /muse/elements/theta_absolute, 0.590401768684
1521053783.847281, /muse/elements/gamma_absolute, -0.415095329285
1521053783.847322, /muse/elements/horseshoe, 1.0, 1.0, 1.0, 1.0
1521053783.847741, /muse/gyro, 0.254211425781, -1.72714233398, 7.55905151367
1521053783.847815, /muse/gyro, -0.441131591797, -1.57760620117, 7.46185302734
1521053783.847865, /muse/gyro, -1.08413696289, -1.06918334961, 7.04315185547
1521053783.847920, /muse/acc, 0.420288085938, 0.240417480469, 0.899353027344
1521053783.847965, /muse/acc, 0.418884277344, 0.235473632812, 0.898742675781
1521053783.848005, /muse/acc, 0.414916992188, 0.226623535156, 0.896545410156
1521053783.880194, /muse/eeg, 620.512817383, 847.362609863, 840.512817383, 838.901123047, 821.978027344
1521053783.880279, /muse/eeg, 655.164855957, 844.945068359, 840.915771484, 871.94140625, 1053.26013184
```

Transforms.conf

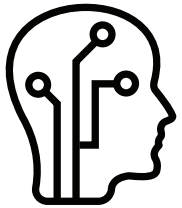
```
[elements_alpha_absolute_value]
REGEX = \/muse\/elements\/alpha_absolute\,\s([-]?[0-9]*\.?[0-9]+)
FORMAT = alpha_absolute::$1
WRITE_META = true

[elements_beta_absolute_value]
REGEX = \/muse\/elements\/beta_absolute\,\s([-]?[0-9]*\.?[0-9]+)
FORMAT = beta_absolute::$1
WRITE_META = true
```

- Granular updates – 100ms
- Clear field boundaries
- Pre-processed EEG metrics

Ideal for Splunk and the MLTK!

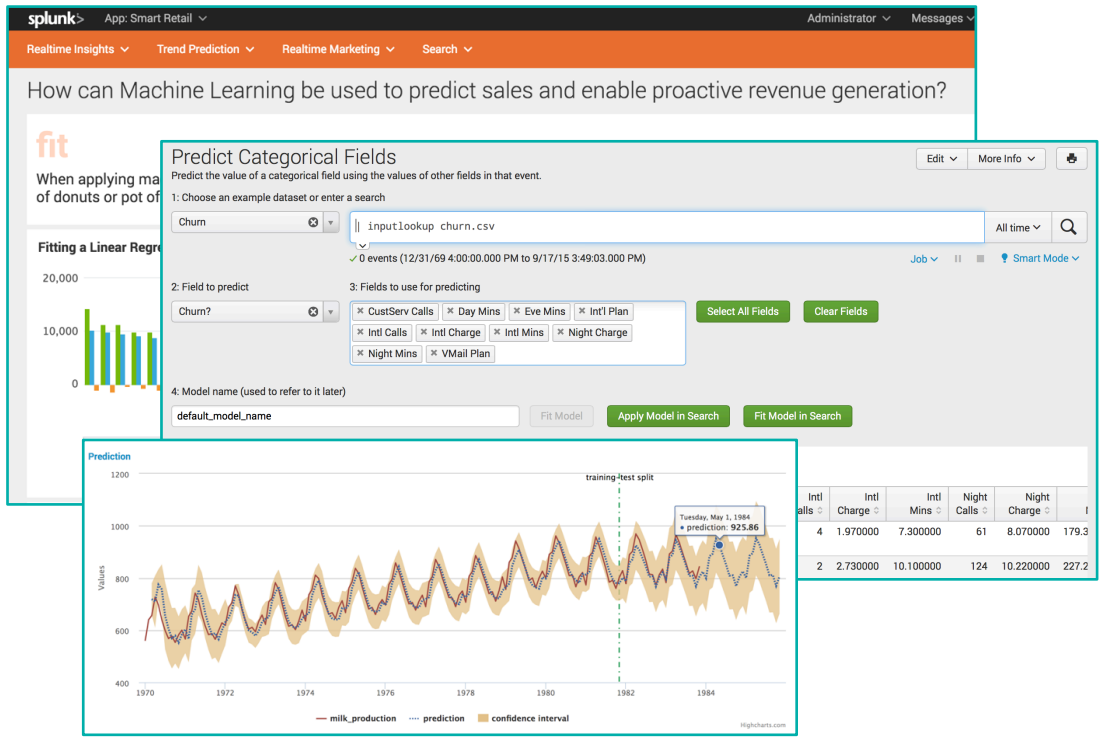
splunk> .conf18



Splunk Machine Learning Toolkit

Extends Splunk platform functions and provides a guided modeling environment

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



Build custom analytics for any use case

splunk> .conf18

Guided Model Building with the MLTK

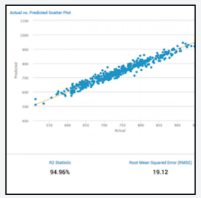
Showcase Provides Real-World Examples of ML Concepts

Showcase

Welcome to the Showcase, which exhibits some of the analytics enabled by this app. Click on one of the examples to see that Assistant applied to a real dataset. Please see the video tutorials for more information.

Select which examples to show

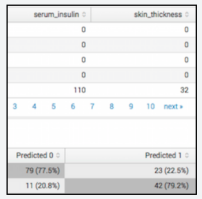
All Examples



Predict Numeric Fields

Predict the value of a numeric field using a weighted combination of the values of other fields in that event. A common use of these predictions is to identify anomalies: predictions that differ significantly from the actual value may be considered anomalous.

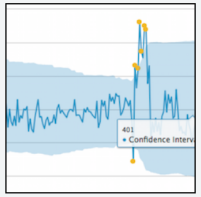
- Predict Server Power Consumption
- Predict VPN Usage
- Predict Median House Value
- Predict Power Plant Energy Output



Predict Categorical Fields

Predict the value of a categorical field using the values of other fields in that event. A common use of these predictions is to identify anomalies: predictions that differ significantly from the actual value may be considered anomalous.

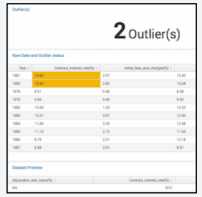
- Predict Hard Drive Failure
- Predict the Presence of Malware
- Predict Telecom Customer Churn
- Predict the Presence of Diabetes
- Predict Vehicle Make and Model



Detect Numeric Outliers

Find values that differ significantly from previous values.

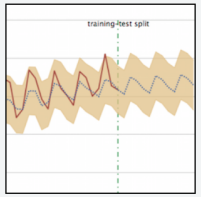
- Detect Outliers in Server Response Time
- Detect Outliers in Number of Logins (vs. Predicted Value)
- Detect Outliers in Supermarket Purchases
- Detect Outliers in Power Plant Humidity



Detect Categorical Outliers

Find events that contain unusual combinations of values.

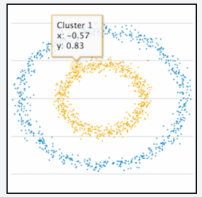
- Detect Outliers in Disk Failures
- Detect Outliers in Bitcoin Transactions
- Detect Outliers in Supermarket Purchases
- Detect Outliers in Mortgage Contracts
- Detect Outliers in Diabetes Patient Records
- Detect Outliers in Mobile Phone Activity



Forecast Time Series

Forecast future values given past values of a metric (numeric time series).

- Forecast Internet Traffic
- Forecast the Number of Employee Logins
- Forecast Monthly Sales
- Forecast the Number of Bluetooth Devices
- Forecast Exchange Rate TWI using ARIMA



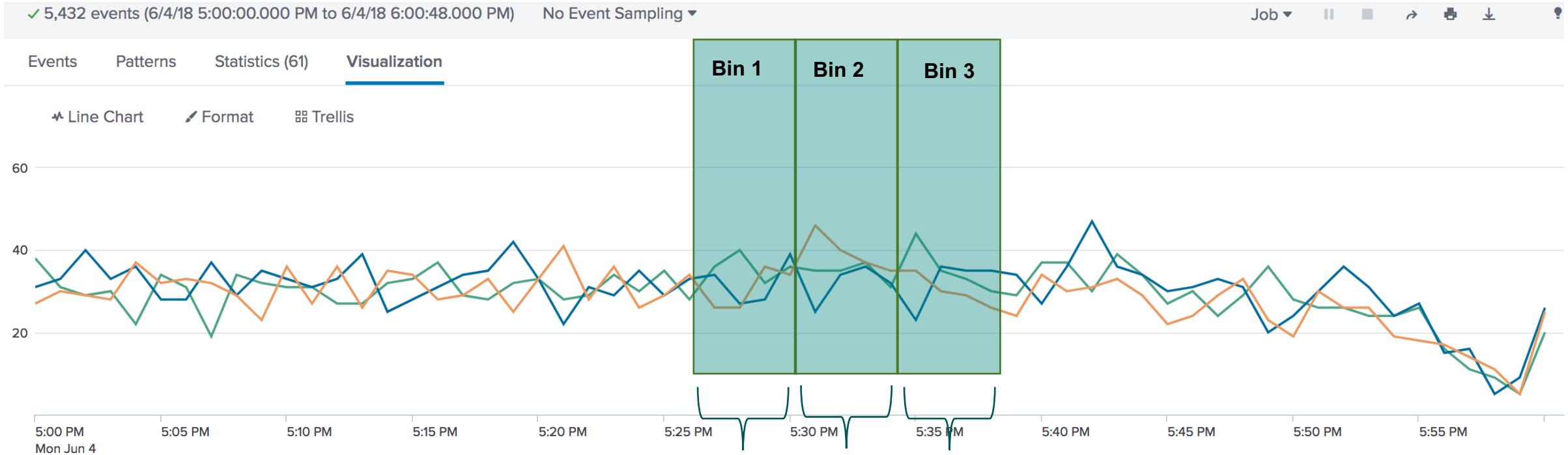
Cluster Numeric Events

Partition events with multiple numeric fields into clusters.

- Cluster Hard Drives by SMART Metrics
- Cluster Behavior by App Usage
- Cluster Neighborhoods by Properties
- Cluster Vehicles by Onboard Metrics
- Cluster Power Plant Operating Regimes

ML Approach

Predict Categorical Field (Mental State) using Time Series EEG Data



- Bucket time into 10 second bins
- Calculate stats on EEG data within bins
- Use calculated stats fields as features for the model

stdev	stdev	stdev
med	med	med
mod	mod	mod
var	var	var
avg	avg	avg

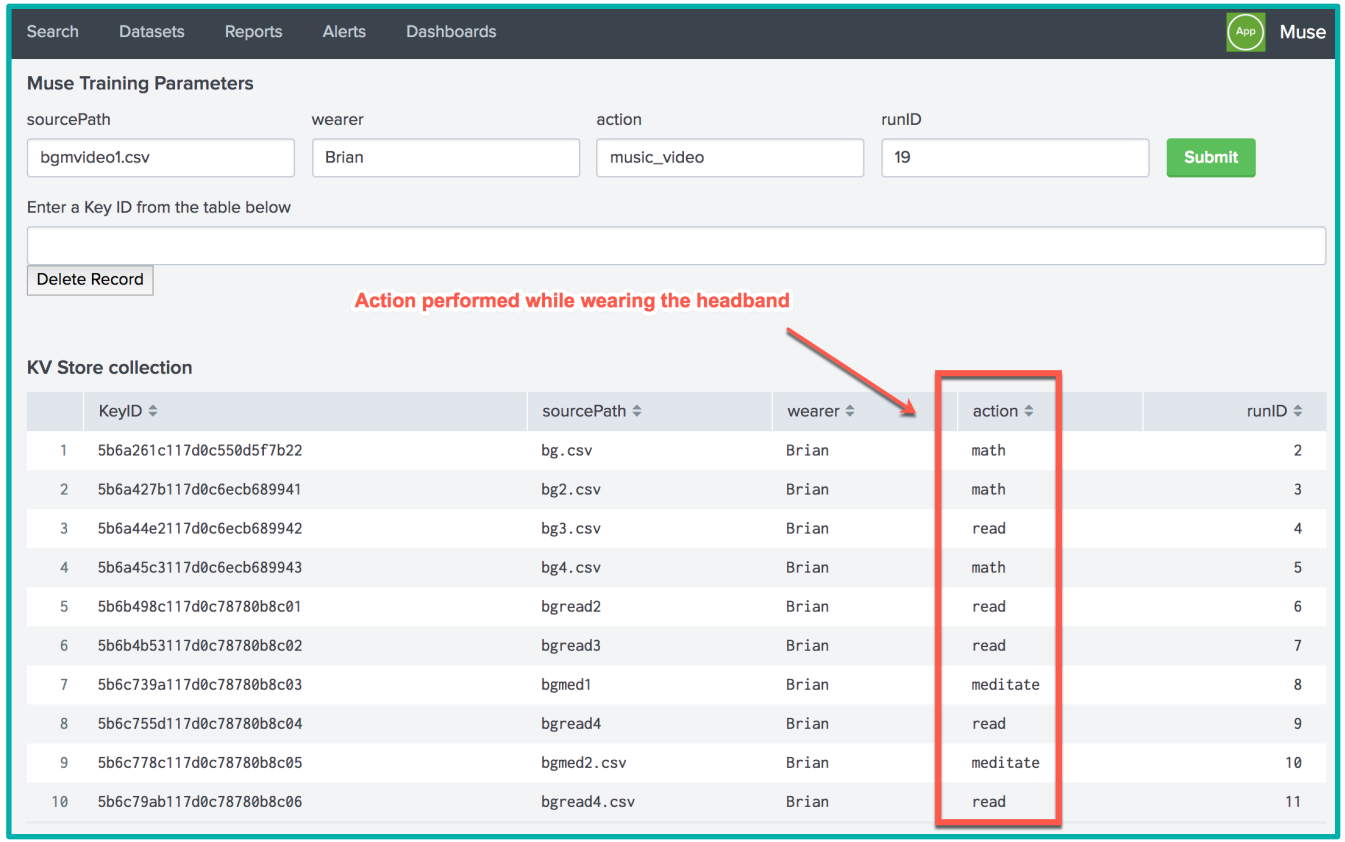
- Tag training data with “action” being performed by EEG wearer
- Use MLTK to test algorithms and model accuracy
- Apply model on real-time data for predictions

splunk> .conf18

Preparing Training Data for ML

Tag EEG Data with Wearer Info Prior to Analysis

- Use KV Store Lookup to tag EEG data with user “action” when wearing headband
- Create custom form to enter “action” field
- Create scheduled search to perform stat calculations on data, lookup “action” tags and save to summary index
- Build ML model from searches against summary index



splunk> .conf18

Machine Learning Process with Splunk

Fit Model

Collect Data

ML Toolkit

Train model

Clean Transform

props transforms

Save to summary index

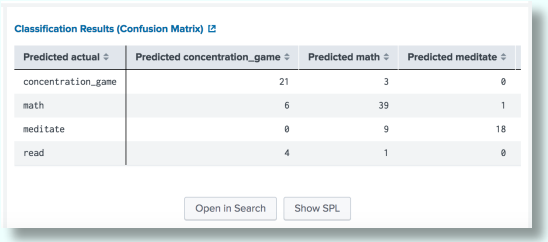
Add "action" field

KV Store

Perform stats calculation

Apply Model

Collect Data



Clean Transform

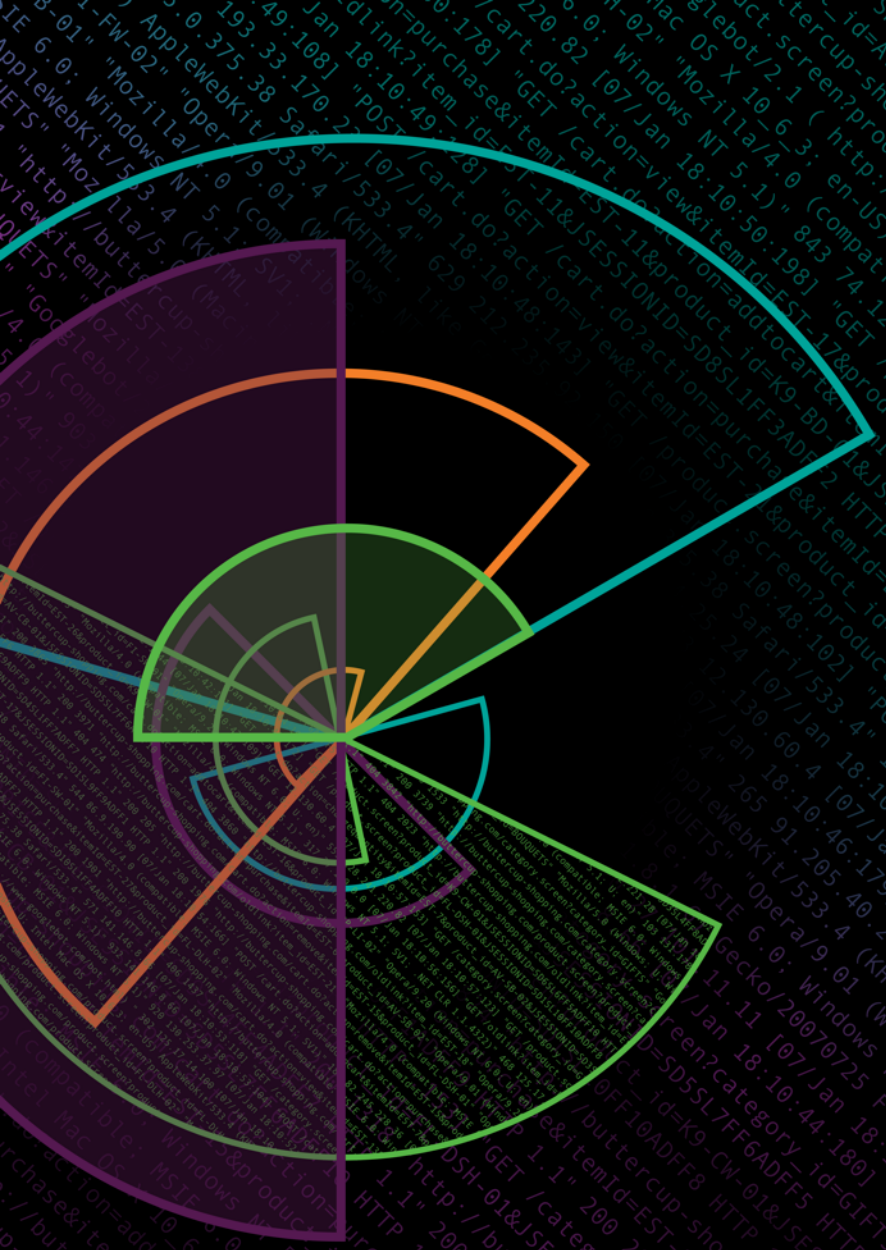
props transforms

Predict "action" field

Perform Stats calculation

Apply model

splunk> .conf18



EEG DEMO

Key Takeaways

1. The MLTK is like a box of legos. What type of model do you want to build?

2. Predict categorical fields on time series data by bucketing into bins, performing stats on bucketed data and using the calculated values as model features

3. Enrich training data with KVStore Lookups, pre-process and store data in a summary index

Q&A

Brian Guilfoyle | Sr Sales Engineer

splunk> .conf18

Thank You

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