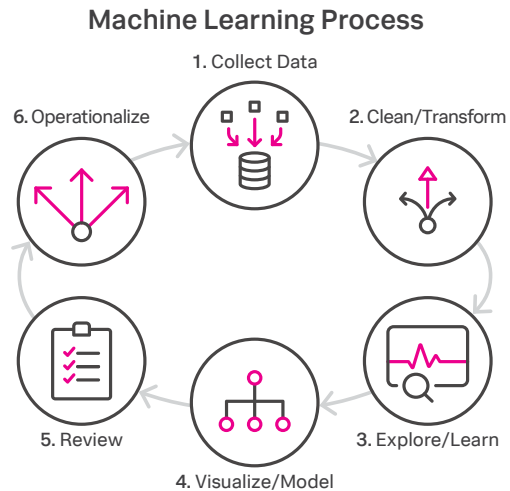# Splunk Machine Learning Toolkit (MLTK)

The MLTK is available to all Splunk Cloud or Splunk Enterprise customers and extends the value of the Splunk platform by enabling users to easily apply machine learning to their data.

- **Guide investigations** by using machine learning to discover hidden meaningful patterns in your data

- **Investigate** your expanding data universe and avoid costly downtime

- **Analyze and monitor** at machine speed with purpose-built machine learning algorithms

- **Automate** action with trained models for alerts in real time

Machine Learning Process

1. Collect Data
2. Clean/Transform
3. Explore/Learn
4. Visualize/Model
5. Review
6. Operationalize

The Splunk Machine Learning Toolkit app delivers the capability to operationalize machine learning (ML) models using your data in Splunk. Explore a variety of machine learning algorithms using guided workflows and an intuitive interface to support common IT, Security, and Business Operations needs and use cases. Our family of Smart Assistants provides step-by-step guidance to develop models, as well as the ability to apply custom visualizations and Splunk Search Processing Language (SPL) commands to one's own data. SPL code is also auto-generated and made available for experienced users to leverage in more sophisticated analyses.

## Let's dive into what the MLTK can do for you

Enables Actionable Intelligence: Splunk MLTK allows you to operationalize your machine learning. Collect and analyze, train models iteratively, and set alerts — in real time.
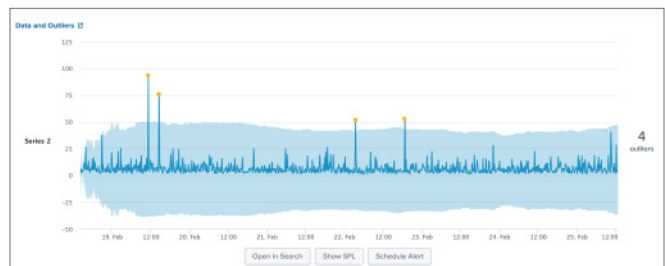
Easy-to-Use Smart Assistants: The MLTK makes machine learning easier and more accessible for a broad range of users with our family of Smart Assistants. Smart Assistant workflows can guide you with your forecasting, anomaly detection, and clustering needs, and by putting machine learning to work on your behalf, you don't need to be a data scientist to use these advanced capabilities. The Assistants provide a simplified interface and aim to remove complexities associated with machine learning exploration.
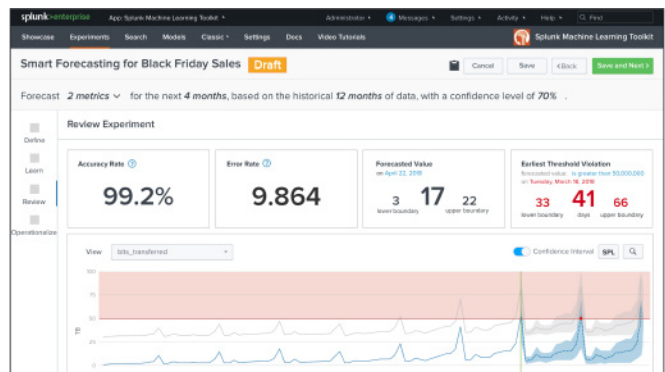
Extensible With Open Source: The MLTK is increasingly more extensible allowing connections with open source and third-party data science communities and tools through the Github community, Apache Spark connector, and container for Tensorflow.
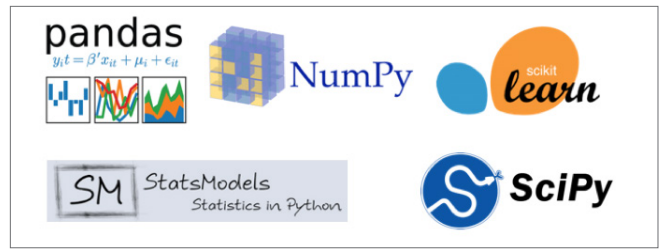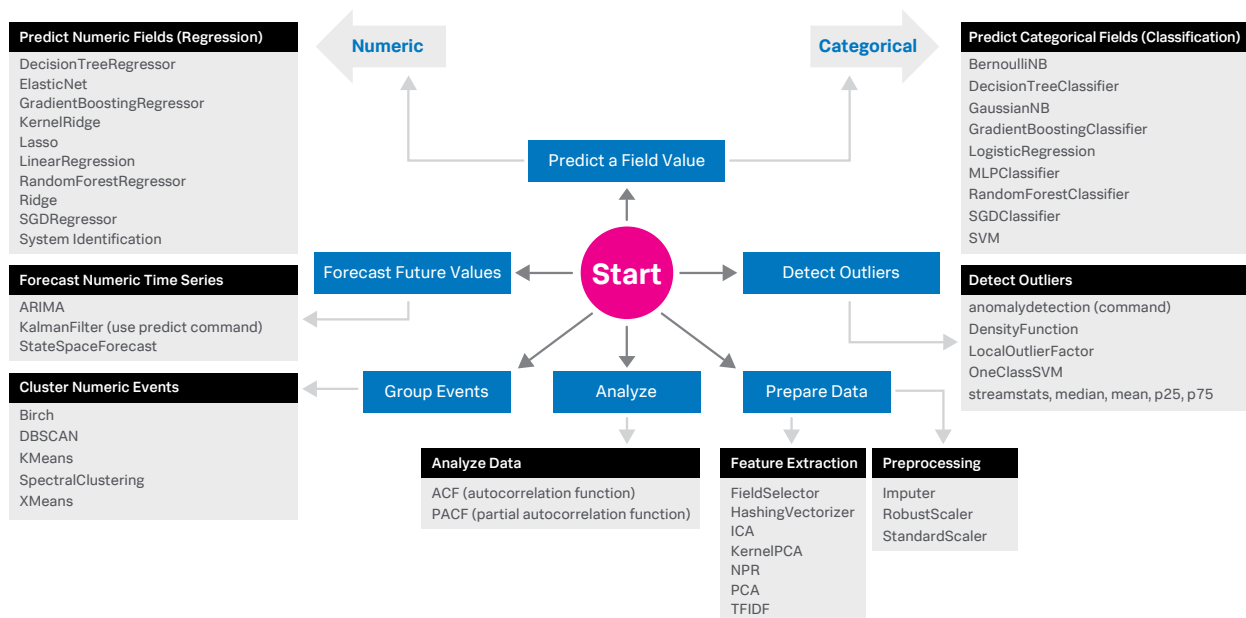

Guided by the Experts – ML Advisory Program: Develop your running production model with complimentary Splunk hands-on guidance and data science resources. Learn more about the ML Advisory Program today: **splk.it/MLAdvisory**.

The Splunk Machine Learning Toolkit has over 30 common algorithms available, as well as access to more than 300 popular open-source algorithms, and includes a Showcase of different sample datasets to help new users explore machine-learning concepts. The MLTK offers guided modeling Assistants for users to leverage, including four Smart Assistants (Smart Forecasting, Smart Outlier Detection, Smart Clustering, and Smart Prediction), as well as six Experiment Assistants, which include:

- Predict Numeric Fields (Linear Regression)
- Predict Categorical Fields (Logistic Regression)
- Detect Numeric Outliers (distribution statistics)
- Detect Categorical Outliers (probabilistic measures)
- Forecast Time Series
- Cluster Numeric Events

New guided assistant panels become available with each additional release of the Machine Learning Toolkit giving users easy access to algorithms like those below and more:

Predict Numeric Fields (Regression)
DecisionTreeRegressor
ElasticNet
GradientBoostingRegressor
KernelRidge
Lasso
LinearRegression
RandomForestRegressor
Ridge
SGDRegressor
System Identification

Forecast Numeric Time Series
ARIMA
KalmanFilter (use predict command)
StateSpaceForecast

Cluster Numeric Events
Birch
DBSCAN
KMeans
SpectralClustering
XMeans

Predict Categorical Fields (Classification)
BernoulliNB
DecisionTreeClassifier
GaussianNB
GradientBoostingClassifier
LogisticRegression
MLPClassifier
RandomForestClassifier
SGDClassifier
SVM

Detect Outliers
anomalydetection (command)
DensityFunction
LocalOutlierFactor
OneClassSVM
streamstats, median, mean, p25, p75

Numeric

Categorical

Predict a Field Value

Start

Forecast Future Values

Detect Outliers

Group Events

Analyze

Prepare Data

Analyze Data
ACF (autocorrelation function)
PACF (partial autocorrelation function)

Feature Extraction
FieldSelector
HashingVectorizer
ICA
KernelPCA
NPR
PCA
TFIDF

Preprocessing
Imputer
RobustScaler
StandardScaler

Download the MLTK app: **splk.it/MLTK**.
Want to learn more? Read **splk.it/MLcheatsheet** and visit our youtube page **splk.it/MLvideos**.

splunk>

Learn more: **www.splunk.com/asksales**

www.splunk.com