SPLUNK® MACHINE LEARNING TOOLKIT (MLTK)

Available with any Splunk Cloud or Splunk Enterprise license, the MLTK extends the value of the Splunk platform by enabling Splunk users to apply machine learning to their 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
- Use machine learning to **guide investigations** and discover hidden meaningful patterns in your data

The Splunk Machine Learning Toolkit App delivers the capability to operationalize machine learning (ML) models on your data in Splunk. Explore a variety of machine learning concepts using new SPL commands, custom visualizations, guided assistants, and several examples spanning IT, Security, IoT and Business Operations. Use guided model assistants in MLTK that include end-to-end examples with datasets, plus the ability to apply visualizations and SPL commands to one's own data. The guided assistant panels' underlying code is made available for experienced users to see how it all works behind the easy-to-use UI.

**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 Guided Model Assistants:** The MLTK makes machine learning easier for a broad range of users through guided assistants with a point-and-click approach that writes SPL for you. Assistants help set up outlier and anomaly detection, predictive analytics and data clustering across IT, Security, IoT and Business Operations use cases. More on the guided assistants below.
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 includes seven guided machine learning assistant panels to walk users through model definition, training, review and operationalization into production. These include:

- Predict Numeric Fields (linear regression): e.g. predict median house values
- Predict Categorical Fields (logistic regression): e.g. predict customer churn
- Detect Numeric Outliers (distribution statistics): e.g. detect outliers in IT Ops data
- Detect Categorical Outliers (probabilistic measures): e.g. detect outliers in diabetes patient records
- Forecast Time Series: e.g. forecast data center growth and capacity planning
- Cluster Numeric Events: e.g. cluster hard drives by SMART Metrics
- Smart Forecasting: e.g. predict and automatically retrain models at scale

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
  - KerasNet
  - Lasso
  - LinearRegression
  - RandomForestRegressor
  - Ridge
  - SGDRegressor
  - System Identification

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

- **Cluster Numeric Events**
  - Birch
  - DBCSCAN
  - 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

- **Preprocessing**
  - HashingVectorizer
  - ICA
  - KernelPCA
  - NPR
  - PCA
  - TRAF

- **Feature Extraction**
  - FieldSelector
  - HashingVectorizer
  - ICA
  - KernelPCA
  - NPR
  - PCA
  - TRAF

Download the MLTK app: splk.it/MLTK and join our beta program: splk.it/MLBeta.

Want to learn more? Read splk.it/MLcheatsheet and visit our youtube page splk.it/MLvideos.