ML-SPL Quick Reference Guide

http://tiny.cc/mlcheatsheet

Machine Learning Toolkit

Search Commands for Machine Learning

The Machine Learning Toolkit provides custom search commands for applying machine learning to your data.

Command	Description	Syntax
fit	Fit and apply a machine learning model to search results.	`... | fit algorithm y from x params into model_name as output_field`
apply	Apply a machine learning model that was learned using the fit command.	`... | apply model_name as output_field`
summary	Return a summary of a machine learning model that was learned using the fit command.	`| summary model_name`
listmodels	Return a list of machine learning models that were learned using the fit command.	`| listmodels`
deletemodel	Delete a machine learning model that was learned using the fit command.	`| deletemodel model_name`
sample	Randomly sample or partition events.	`... | sample options by split_by_field`

Feature Extraction

Feature extraction algorithms transform fields for better prediction accuracy.

Algorithm	Examples
FieldSelector	`...| fit FieldSelector type=categorical SLA_violation from *`
PCA	`...| fit PCA * k=3`
KernelPCA	`...| fit KernelPCA * k=3 gamma=0.001`
TFIDF	`...| fit TFIDF Reviews into user_feedback_model max_def=0.6 min_def=0.2`

Anomaly Detection

Find events that contain unusual combinations of values.

Algorithm	Examples
OneClassSVM	`... | fit OneClassSVM * kernel="poly" nu=0.5coef0=0.5 gamma=0.5 tol=1 degree=3 shrinking=f into TESTMODEL_OneClassSVM`

Preprocessing

Preprocessing algorithms are used for preparing data and help with prediction accuracy.

Algorithm	Examples
StandardScaler	`... | fit StandardScaler *`

Cluster Numeric

Partition events with multiple numeric fields into clusters.

Algorithm	Examples
KMeans	`... | fit KMeans * k=3`
DBSCAN	`... | fit DBSCAN *`
BIRCH	`... | fit Birch * k=3`
SpectralClustering	`... | fit SpectralClustering * k=3`

Forecasting

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

Algorithm	Examples
ARIMA	`... | fit ARIMA Voltage order=4-0-1`

Predict Numeric

Predict the value of a numeric field using the values of other fields in that event.

Algorithm	Examples
LinearRegression	`... | fit LinearRegression temperature from date_month date_hour into temperature_model`
Lasso	`... | fit Lasso temperature from date_month date_hour`
Ridge	`... | fit Ridge temperature from date_month date_hour normalize=true alpha=0.5`
ElasticNet	`... | fit ElasticNet temperature from date_month date_hour normalize=true alpha=0.5`
KernelRidge	`... | fit KernelRidge temperature from date_month date_hour into temperature_model`
SGDRegressor	`... | fit SGDRegressor temperature from date_month date_hour into temperature_model`
DecisionTreeRegressor	`... | fit DecisionTreeRegressor temperature from date_month date_hour into temperature_model`
RandomForestRegressor	`... | fit RandomForestRegressor temperature from date_month date_hour into temperature_model`

Predict Categorical

Predict the value of a categorical field using the values of other fields in that event.

Algorithm	Examples
LogisticRegression	`... | fit LogisticRegression SLA_violation from IO_wait_time into sla_model`
SVM	`... | fit SVM SLA_violation from * into sla_model`
BernoulliNB	`... | fit BernoulliNB type from * into TESTMODEL_BernoulliNB alpha=0.5 binarize=0 fit_prior=f`
GaussianNB	`... | fit GaussianNB species from * into TESTMODEL_GaussianNB`
SGDClassifier	`... | fit SGDClassifier SLA_violation from * into sla_model`
DecisionTreeClassifier	`... | fit DecisionTreeClassifier SLA_violation from * into sla_model`
RandomForestClassifier	`... | fit RandomForestClassifier SLA_violation from * into sla_model`