

Tutorials

Clustering

[Beginner](#)

Anomaly

[Beginner](#)

Association Rules

[Beginner](#)

NLP

[Beginner](#)

[Intermediate](#)

Regression

[Beginner](#)

[Intermediate](#)

Classification

[Binary \(Beginner\)](#)

[Binary \(Intermediate\)](#)

[Multiclass \(Beginner\)](#)

Time Series

[Beginner](#)

Unsupervised Learning

Supervised Learning

Time Series Analysis

Clustering	Anomaly Detection	Association Rule	Regression & Classification		Time Series	
<div>setup(data=df)</div> <div>create_model(*)</div> <div>assign_model(*)</div> <div>plot_model(**)</div> <div>evaluate_model()</div> <div>tune_model()</div> <div>predict_model()</div> <div>deploy_model()</div> <div>save_model()</div> <div>load_model()</div> <div>pull()</div> <div>models()</div> <div>get_metrics()</div> <div>add_metric()</div> <div>remove_metric()</div> <div>get_logs()</div> <div>get_config()</div> <div>set_config()</div> <div>save_config()</div> <div>load_config()</div> <div>get_clusters()</div> <div>* model:</div> <div>'kmeans'</div> <div>'ap'</div> <div>'meanshift'</div> <div>'sc'</div> <div>'hclust'</div> <div>'dbscan'</div> <div>'optics'</div> <div>'birch'</div> <div>'kmodes'</div> <div>** plot =</div> <div>'cluster'</div> <div>'tsne'</div> <div>'elbow'</div> <div>'silhouette'</div> <div>'distance'</div> <div>'distribution'</div>	<div>setup(data=df)</div> <div>create_model(*)</div> <div>assign_model(*)</div> <div>plot_model(**)</div> <div>evaluate_model()</div> <div>tune_model()</div> <div>predict_model()</div> <div>deploy_model()</div> <div>save_model()</div> <div>load_model()</div> <div>models()</div> <div>get_logs()</div> <div>get_config()</div> <div>set_config()</div> <div>save_config()</div> <div>load_config()</div> <div>get_outliers()</div> <div>* model:</div> <div>'abod'</div> <div>'cluster'</div> <div>'histogram'</div> <div>'knn'</div> <div>'lof'</div> <div>'svm'</div> <div>'pca'</div> <div>'mcd'</div> <div>'sod'</div> <div>'sos'</div> <div>** plot=</div> <div>'tsne'</div> <div>'umap'</div>	<div>setup(data=df,</div> <div>transaction_id='column',</div> <div>item_id='column')</div> <div>create_model()</div> <div>plot_model(**)</div> <div>get_rules()</div> <div>** plot=</div> <div>'2d'</div> <div>'3d'</div> <div>NLP</div> <div>setup(data=df)</div> <div>create_model(*)</div> <div>assign_model()</div> <div>plot_model()</div> <div>tune_model()</div> <div>evaluate_model()</div> <div>save_model()</div> <div>load_model()</div> <div>models()</div> <div>get_logs()</div> <div>get_config()</div> <div>set_config()</div> <div>get_topics()</div> <div>* model:</div> <div>'lda'</div> <div>'lsi'</div> <div>'hdp'</div> <div>'rp'</div> <div>'nmf'</div> <div>'frequency'</div> <div>'distribution'</div> <div>'bigram'</div> <div>'trigram'</div> <div>'sentiment'</div> <div>'pos'</div> <div>'tsne'</div> <div>'topic_model'</div> <div>'topic_distribution'</div> <div>'wordcloud'</div> <div>'umap'</div>	<div>setup(data=df,</div> <div>target='column')</div> <div>models(*)</div> <div>compare_models()</div> <div>create_model(*)</div> <div>tune_model()</div> <div>ensemble_model()</div> <div>blend_models()</div> <div>stack_models()</div> <div>plot_model(**)</div> <div>evaluate_model()</div> <div>interpret_model()</div> <div>(1) calibrate_model()</div> <div>(1) optimize_threshold()</div> <div>predict_model()</div> <div>finalize_model()</div> <div>deploy_model()</div> <div>save_model()</div> <div>load_model()</div> <div>automl()</div> <div>pull()</div> <div>models()</div> <div>get_metrics()</div> <div>add_metric()</div> <div>remove_metric()</div> <div>get_logs()</div> <div>get_config()</div> <div>set_config()</div> <div>save_config()</div> <div>load_config()</div> <div>get_leaderboard()</div> <div>(1) classification only</div>	<div>* model:</div> <div>(classification)</div> <div>'lr'</div> <div>'knn'</div> <div>'nb'</div> <div>'dt'</div> <div>'svm'</div> <div>'rbfsvm'</div> <div>'gpc'</div> <div>'mlp'</div> <div>'ridge'</div> <div>'rf'</div> <div>'qda'</div> <div>'ada'</div> <div>'gbc'</div> <div>'lda'</div> <div>'et'</div> <div>'xgboost'</div> <div>'lightgbm'</div> <div>'catboost'</div> <div>** plot=</div> <div>'auc'</div> <div>'threshold'</div> <div>'pr'</div> <div>'error'</div> <div>'class_report'</div> <div>'boundary'</div> <div>'rfe'</div> <div>'learning'</div> <div>'manifold'</div> <div>'calibration'</div> <div>'vc'</div> <div>'dimension'</div> <div>'feature'</div> <div>'feature_all'</div> <div>'parameter'</div> <div>'lift'</div> <div>'gain'</div> <div>'tree'</div> <div>'ks'</div> <div>'confusion_matrix'</div>	<div>* model:</div> <div>(regression)</div> <div>'lr'</div> <div>'lasso'</div> <div>'ridge'</div> <div>'en'</div> <div>'lar'</div> <div>'llar'</div> <div>'omp'</div> <div>'br'</div> <div>'ard'</div> <div>'par'</div> <div>'ransac'</div> <div>'tr'</div> <div>'huber'</div> <div>'kr'</div> <div>'svm'</div> <div>'knn'</div> <div>'dt'</div> <div>'rf'</div> <div>'et'</div> <div>'ada'</div> <div>'gbr'</div> <div>'mlp'</div> <div>'xgboost'</div> <div>'lightgbm'</div> <div>'catboost'</div> <div>** plot=</div> <div>'residuals'</div> <div>'error'</div> <div>'cooks'</div> <div>'rfe'</div> <div>'learning'</div> <div>'vc'</div> <div>'manifold'</div> <div>'feature'</div> <div>'feature_all'</div> <div>'residuals_</div> <div>interactive'</div> <div>'parameter'</div> <div>'tree'</div>	<div>setup(data=df)</div> <div>models(*)</div> <div>compare_models()</div> <div>create_model(*)</div> <div>tune_model()</div> <div>blend_models()</div> <div>plot_model()</div> <div>finalize_model()</div> <div>deploy_model()</div> <div>save_model()</div> <div>load_model()</div> <div>pull()</div> <div>models()</div> <div>predict_model()</div> <div>get_metrics()</div> <div>add_metric()</div> <div>remove_metric()</div> <div>get_logs()</div> <div>get_config()</div> <div>set_config()</div> <div>save_config()</div> <div>load_config()</div> <div>** plot=</div> <div>'ts'</div> <div>'cv'</div> <div>'acf'</div> <div>'pacf'</div> <div>'decomp_stl'</div> <div>'diagnostics'</div> <div>'forecast'</div> <div>'insample'</div> <div>'residuals'</div> <div>'train_test_split'</div> <div>'decomp_classical'</div> <div>* model:</div> <div>'naïve'</div> <div>'grand_means'</div> <div>'snaive'</div> <div>'polytrend'</div> <div>'arima'</div> <div>'auto_arima'</div> <div>'exp_smooth'</div> <div>'ets'</div> <div>'theta'</div> <div>'tbats'</div> <div>'bats'</div> <div>'prophet'</div> <div>'lr_cds_dt'</div> <div>'en_cds_dt'</div> <div>'ridge_cds_dt'</div> <div>'lasso_cds_dt'</div> <div>'lar_cds_dt'</div> <div>'llar_cds_dt'</div> <div>'br_cds_dt'</div> <div>'huber_cds_dt'</div> <div>'par_cds_dt'</div> <div>'omp_cds_dt'</div> <div>'knn_cds_dt'</div> <div>'dt_cds_dt'</div> <div>'rf_cds_dt'</div> <div>'et_cds_dt'</div> <div>'gbr_cds_dt'</div> <div>'ada_cds_dt'</div> <div>'lightgbm_cds_dt'</div>

Clustering

```
data,
preprocess = True,
imputation_type = 'simple',
iterative_imputation_iters = 5,
categorical_features = None,
categorical_imputation = 'mode',
categorical_iterative_imputer = 'lightgbm',
ordinal_features = None,
high_cardinality_features = None,
high_cardinality_method = 'frequency',
numeric_features = None,
numeric_imputation = 'mean',
numeric_iterative_imputer = 'lightgbm',
date_features = None,
ignore_features = None,
normalize = False,
normalize_method = 'zscore',
transformation = False,
transformation_method = 'yeo-johnson',
handle_unknown_categorical = True,
unknown_categorical_method = 'least_frequent',
pca = False,
pca_method = 'Linear',
pca_components = None,
ignore_low_variance = False,
combine_rare_levels = False,
rare_level_threshold = 0.1,
bin_numeric_features = None,
remove_multicollinearity = False,
multicollinearity_threshold = 0.9,
remove_perfect_collinearity = False,
group_features = None,
group_names = None,
n_jobs = -1,
use_gpu = False,
custom_pipeline = None,
html = True,
session_id = None,
system_log = True,
log_experiment = False,
experiment_name = None,
log_plots = False,
log_profile = False,
log_data = False,
silent = False,
verbose = True,
profile = False,
profile_kwargs = None
```

Color code

required

optional

Anomaly Detection

```
data,
Preprocess = True,
imputation_type = 'simple',
iterative_imputation_iters = 5,
categorical_features = None,
categorical_imputation = 'mode',
categorical_iterative_imputer = 'lightgbm',
ordinal_features = None,
high_cardinality_features = None,
high_cardinality_method = 'frequency',
numeric_features = None,
numeric_imputation = 'mean',
numeric_iterative_imputer = 'lightgbm',
date_features = None,
ignore_features = None,
normalize = False,
normalize_method = 'zscore',
transformation = False,
transformation_method = 'yeo-johnson',
handle_unknown_categorical = True,
unknown_categorical_method = 'least_frequent',
pca = False,
pca_method = 'Linear',
pca_components = None,
ignore_low_variance = False,
combine_rare_levels = False,
rare_level_threshold = 0.1,
bin_numeric_features = None,
remove_multicollinearity = False,
multicollinearity_threshold = 0.9,
remove_perfect_collinearity = False,
group_features = None,
group_names = None,
n_jobs = -1,
use_gpu = False,
custom_pipeline = None,
html = True,
session_id = None,
system_log = True,
log_experiment = False,
experiment_name = None,
log_plots = False,
log_profile = False,
log_data = False,
silent = False,
verbose = True,
profile = False,
profile_kwargs = None
```

Regression & Classification

```
data = DataFrame, target = 'column_name',
train_size = 0.7,
test_data = None,
preprocess = True,
imputation_type = 'simple',
iterative_imputation_iters = 5,
categorical_features = None,
categorical_imputation = 'constant',
categorical_iterative_imputer = 'lightgbm',
ordinal_features = None,
high_cardinality_features = None,
high_cardinality_method = 'frequency',
numeric_features = None,
numeric_imputation = 'mean',
numeric_iterative_imputer = 'lightgbm',
date_features = None,
ignore_features = None,
normalize = False,
normalize_method = 'zscore',
transformation = False,
transformation_method = 'yeo-johnson',
handle_unknown_categorical = True,
unknown_categorical_method = 'least_frequent',
pca = False,
pca_method = 'Linear',
pca_components = None,
ignore_low_variance = False,
combine_rare_levels = False,
rare_level_threshold = 0.1,
bin_numeric_features = None,
remove_outliers = False,
outliers_threshold = 0.05,
remove_multicollinearity = False,
multicollinearity_threshold = 0.9,
remove_perfect_collinearity = True,
```

```
create_clusters = False,
cluster_iter = 20,
polynomial_features = False,
polynomial_degree = 2,
trigonometry_features = False,
polynomial_threshold = 0.1,
group_features = None,
group_names = None,
feature_selection = False,
feature_selection_threshold = 0.8,
feature_selection_method = 'classic',
feature_interaction = False,
feature_ratio = False,
interaction_threshold = 0.01,
transform_target = False,
transform_target_method = 'box-cox',
data_split_shuffle = True,
data_split_stratify = False,
fold_strategy = 'kfold',
fold = 10,
fold_shuffle = False,
fold_groups = None,
n_jobs = -1,
use_gpu = False,
custom_pipeline = None,
html = True,
session_id = None,
log_experiment = False,
experiment_name = None,
log_plots = False,
log_profile = False,
log_data = False,
silent = False,
verbose = True,
profile = False,
profile_kwargs = None
```

Time Series

```
data = [Series, DataFrame],
preprocess = True,
imputation_type = 'simple',
fold_strategy = 'expanding',
fold = 3,
fh = 1,
seasonal_period = None,
enforce_pi = False,
n_jobs = -1,
use_gpu = False,
custom_pipeline = None,
html = True,
session_id = None,
system_log = True,
log_experiment = False,
experiment_name = None,
log_plots = False,
log_profile = False,
log_data = False,
verbose = True,
profile = False,
profile_kwargs = None
```

Association Rule

```
data,
transaction_id = 'column_name',
item_id = 'column_name',
ignore_items = None,
session_id = None
```

NLP

```
data,
Target = 'column_name',
custom_stopwords = None,
Html = True,
session_id = None,
log_experiment = False,
experiment_name = None,
log_plots = False,
log_data = False,
Verbose = True
```