

<https://runawayhorse001.github.io/LearningApacheSpark/pyspark.pdf>

A
accuracy() (pys park.ml.classification.LogisticRegressionSummary property), 441
addGrid() (pyspark.ml.tuning.ParamGridBuilder method), 475
AFTSurvivalRegression (class in pys park.ml.regression), 411
AFTSurvivalRegressionModel (class in pys park.ml.regression), 412
aic() (pyspark.ml.regression.GeneralizedLinearRegressionSummary property), 420
ALS (class in pyspark.ml.recommendation), 465
ALSModel (class in pyspark.ml.recommendation), 469
areaUnderROC() (pys park.ml.classification.BinaryLogisticRegressionSummary property), 430
assignClusters() (pys park.ml.clustering.PowerIterationClustering method), 463
avgMetrics (pys park.ml.tuning.CrossValidatorModel attribute), 474

B
baseOn() (pyspark.ml.tuning.ParamGridBuilder method), 475
bestModel (pys park.ml.tuning.CrossValidatorModel attribute), 474
bestModel (pys park.ml.tuning.TrainValidationSplitModel attribute), 477
BinaryClassificationEvaluator (class in pyspark.ml.evaluation), 477
BinaryLogisticRegressionSummary (class in pyspark.ml.classification), 430
BinaryLogisticRegressionTrainingSummary (class in pyspark.ml.classification), 431
BisectingKMeans (class in pys park.ml.clustering), 450
BisectingKMeansModel (class in pys park.ml.clustering), 451
BisectingKMeansSummary (class in pys park.ml.clustering), 452
boundaries() (pys park.ml.regression.IsotonicRegressionModel property), 423
build() (pyspark.ml.tuning.ParamGridBuilder method), 475

C
ChiSquareTest (class in pyspark.ml.stat), 405
clusterCenters() (pys park.ml.clustering.BisectingKMeansModel method), 451
clusterCenters() (pys park.ml.clustering.KMeansModel method), 457
ClusteringEvaluator (class in pys park.ml.evaluation), 478
coefficientMatrix() (pys park.ml.classification.LogisticRegressionModel property), 440
coefficients() (pys park.ml.classification.LinearSVCModel property), 437
coefficients() (pys park.ml.classification.LogisticRegressionModel property), 440
coefficients() (pys park.ml.regression.AFTSurvivalRegressionModel 413
coefficients() (pys park.ml.regression.GeneralizedLinearRegressionModel property), 419
coefficients() (pys park.ml.regression.LinearRegressionModel property), 424
coefficientStandardErrors() (pys park.ml.regression.GeneralizedLinearRegressionTrainingSummary property), 421
coefficientStandardErrors() (pys park.ml.regression.LinearRegressionSummary property), 425
computeCost() (pys park.ml.clustering.BisectingKMeansModel method), 452
computeCost() (pys park.ml.clustering.KMeansModel method), 457
Configure Spark on Mac and Ubuntu, 17
copy() (pyspark.ml.classification.OneVsRest method), 447
copy() (pyspark.ml.classification.OneVsRestModel method), 448
copy() (pyspark.ml.pipeline.Pipeline method), 470
copy() (pyspark.ml.pipeline.PipelineModel method), 471
copy() (pyspark.ml.tuning.CrossValidator method), 473
copy() (pyspark.ml.tuning.CrossValidatorModel method), 474
copy() (pyspark.ml.tuning.TrainValidationSplit method), 476
copy() (pyspark.ml.tuning.TrainValidationSplitModel method), 477
corr() (pyspark.ml.stat.Correlation static method), 406
Correlation (class in pyspark.ml.stat), 406
count() (pyspark.ml.stat.Summarizer static method), 409
CrossValidator (class in pyspark.ml.tuning), 472
CrossValidatorModel (class in pys park.ml.tuning), 473

D
DecisionTreeClassificationModel (class in pyspark.ml.classification), 431 DecisionTreeClassifier (class in
pys park.ml.classification), 432 DecisionTreeRegressionModel (class in pyspark.ml.regression), 413 DecisionTreeRegressor
(class in pys park.ml.regression), 413 degreesOfFreedom() (pys park.ml.regression.GeneralizedLinearRegressionSummary
property), 420 degreesOfFreedom() (pys park.ml.regression.LinearRegressionSummary property), 425 describeTopics()
(pys park.ml.clustering.LDAModel method), 462 deviance() (pys park.ml.regression.GeneralizedLinearRegressionSummary
property), 420 devianceResiduals() (pys park.ml.regression.LinearRegressionSummary property), 425 dispersion()
(pys park.ml.regression.GeneralizedLinearRegressionSummary property), 420 DistributedLDAModel (class in
pys park.ml.clustering), 452 E estimatedDocConcentration() (pys park.ml.clustering.LDAModel method), 462 evaluate()
(pys park.ml.classification.LogisticRegressionModel method), 440 evaluate() (pyspark.ml.evaluation.Evaluator method), 479
evaluate() (pys park.ml.regression.GeneralizedLinearRegressionModel method), 419 evaluate()
(pys park.ml.regression.LinearRegressionModel method), 424 evaluateEachIteration()
(pys park.ml.classification.GBTClassificationModel method), 433

raw
<p>A accuracy() (pys park.ml.classification.LogisticRegressionSummary property), 441 addGrid() (pyspark.ml.tuning.ParamGridBuilder method), 475 AFTSurvivalRegression (class in pys park.ml.regression), 411 AFTSurvivalRegressionModel (class in pys park.ml.regression), 412 aic() (pyspark.ml.regression.GeneralizedLinearRegressionSummary property), 420 ALS (class in pyspark.ml.recommendation), 465 ALSModel (class in pyspark.ml.recommendation), 469 areaUnderROC() (pys park.ml.classification.BinaryLogisticRegressionSummary property), 430 assignClusters() (pys park.ml.clustering.PowerIterationClustering method), 463 avgMetrics (pys park.ml.tuning.CrossValidatorModel attribute), 474</p> <p>B baseOn() (pyspark.ml.tuning.ParamGridBuilder method), 475 bestModel (pys park.ml.tuning.CrossValidatorModel attribute), 474 bestModel (pys park.ml.tuning.TrainValidationSplitModel attribute), 477 BinaryClassificationEvaluator (class in pyspark.ml.evaluation), 477 BinaryLogisticRegressionSummary (class in pyspark.ml.classification), 430 BinaryLogisticRegressionTrainingSummary (class in pyspark.ml.classification), 431 BisectingKMeans (class in pys park.ml.clustering), 450 BisectingKMeansModel (class in pys park.ml.clustering), 451 BisectingKMeansSummary (class in pys park.ml.clustering), 452 boundaries() (pys park.ml.regression.IsotonicRegressionModel property), 423 build() (pyspark.ml.tuning.ParamGridBuilder method), 475</p> <p>C ChiSquareTest (class in pyspark.ml.stat), 405 clusterCenters() (pys park.ml.clustering.BisectingKMeansModel method), 451 clusterCenters() (pys park.ml.clustering.KMeansModel method), 457 ClusteringEvaluator (class in pys park.ml.evaluation), 478 coefficientMatrix() (pys park.ml.classification.LogisticRegressionModel property), 440 coefficients() (pys park.ml.classification.LinearSVCModel property), 437 coefficients() (pys park.ml.classification.LogisticRegressionModel property), 440 coefficients() (pys park.ml.regression.AFTSurvivalRegressionModel</p>