

Task	Method	Package	Function
Classification	Decision Tree (CART)	rpart	rpart
	Decision Tree (C5.0)	C50	C5.0
	Random Forest	randomForest	randomForest
	K-Nearest Neighbor (kNN)	class	knn
	Support Vector Machines (SVM)	e1071	svm
		kernlab	ksvm
	Neural Network (NN)	nnet	nnet
	Logistic Regression	stats	glm
	Naïve Bayes Classifier	e1071	naiveBayes
Regression	Linear Discriminant Analysis	MASS	lda
	Decision Tree (CART)	rpart	rpart
	Random Forest	randomForest	randomForest
	Support Vector Machines	e1071	svm
		kernlab	ksvm
	Linear Regression	stats	lm
	Logistic Regression	stats	glm
	Stepwise Regression	stats	step
	Quantile Regression	quantreg	rq
	LASSO (Least Absolute Shrinkage and Selection Operator)	glmnet	glmnet
		lars	lars
		lasso2	gl1ce
	Ridge Regression	ridge	linearRidge logisticRidge
		MASS	lm.ridge
	Gaussian Process Regression	kernlab	gausspr
Clustering	K-Centroids Clustering	flexclust	kcca
	K-Means Clustering	stats	kmeans
	Kernel K-Means Clustering	kernlab	kkmeans
	Hierarchical Clustering	stats	hclust
	Agglomerative Hierarchical Clustering	cluster	agnes
	Divisive Hierarchical Clustering	cluster	diana
	Density-Based Clustering (DBSCAN)	fpc	dbscan
	Model-Based Clustering	mclust	MClust
	Self-Organizing Maps	kohonen	som
Frequent Pattern /Association Mining	Apriori Algorithm	arules	apriori
	Eclat Algorithm	arules	eclat
	cSPADE Algorithm	arulesSequences	cspade
Anomaly /Outlier/Change Point Detection	Outlier Detection	outliers	outlier
	Chi-Squared Test for Outliers	outliers	chisq.out.test
	Cochran's C Test for Outliers	outliers	cochran.test
	Dixon's Q Test for Outliers	outliers	dixon.test
	Grubb's Test for Outliers	outliers	grubbs.test
	Local Outlier Factor (LOF) Algorithm	DMwR	lofactor
	Changes in Mean Detection	changepoint	cpt.mean
	Changes in Mean and Variance Detection	changepoint	cpt.meanvar
	Changes in Variance Detection	changepoint	cpt.var
	E-Divisive with Medians (EDM) Algorithm	twitter/Breakout Detection	breakout
	Seasonal Hybrid ESD (S-H-ESD) Algorithm	twitter/AnomalyDetection	AnomalyDetectionTs AnomalyDetectionVec

Similarity Matching	Partial String Matching	base	pmatch
	Approximate String Matching	stringdist	amatch
	Pairwise String Distances Computation	stringdist	stringdist stringdistmatrix
	Cluster Similarity Computation (Rand Index, Jaccard Coefficient)	clusteval	cluster_similarity
	Vertex Similarity Computation (Jaccard Coefficient, Dice Coefficient)	igraph	similarity
	Dissimilarity Matrix Computation	cluster	daisy
	Fingerprint Matching	fingerprint	distance
	Flexible Maximum Common Substructure (FMCS) Algorithm	fmcsR	fmcs
Data Reduction/ Sampling	Random Sampling (with or without Replacement)	base	sample
		sampling	srswr srswor
	Balanced Sampling	sampling	balancedcluster balancedstratification balancedtwestage
	Multistage Sampling	sampling	mstage
	Stratified Sampling	sampling	strata
	Brewer Sampling	sampling	UPbrewer
Feature Extraction/ Dimensionality Reduction	Singular Value Decomposition (SVD)	base	svd
	Principal Component Analysis (PCA)	stats	prcomp princomp
	Kernel PCA	kernlab	kpca
	Independent Component Analysis (ICA)	fastICA	fastICA
Feature Selection	Correlation-Based Feature Selection (CFS)	FSelector	cfs
	Chi-Squared Filter	FSelector	chi.squared
	Consistency-Based Filter	FSelector	consistency
	Correlation Filter	FSelector	linear.correlation rank.correlation
	Entropy-Based Filter	FSelector	information.gain gain.ratio symmetrical.uncertainty
	Random Forest Filter	FSelector	random.forest.importance
	RReliefF Filter	FSelector	relief
	Feature Subset Selection	FSelector	best.first.search exhaustive.search greedy.search hill.climbing.search
Causal Modeling	PC Algorithm	pcalg	pc
	Fast Causal Inference (FCI) Algorithm	pcalg	fci
	Really Fast Causal Inference (RFCI) Algorithm	pcalg	rfci
	Greedy Equivalence Search (GES) Algorithm	pcalg	ges
	Greedy Interventional Equivalence Search (GIES) Algorithm	pcalg	gies
	Intervention calculus when the DAG is Absent (IDA) Algorithm	pcalg	ida