Package: ClassificationEnsembles (via r-universe)

October 26, 2026

Type Package

Title Automated Multi-Class Stacking Ensembles and Predictive Diagnostics

Version 1.1.0

Description Automated multi-class classification ensemble pipeline executing concurrent parameter optimizations across logistic, tree, kernel, and neural configurations. Implements dynamic out-of-sample blending, automated feature reduction via iterative Variance Inflation Factors (VIF), and native S3 diagnostic dashboards.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

Depends R (>= 4.1.0)

Imports car, caret, doParallel, dplyr, foreach, ggplot2, ggrepel, patchwork, pROC, rstudioapi, shiny, tidyr, utils, stats, grDevices

Suggests testthat (>= 3.0.0), knitr, quarto, rmarkdown

Config/testthat/edition 3

VignetteBuilder knitr

Config/roxygen2/version 8.0.0

URL https://github.com/InfiniteCuriosity/ClassificationEnsembles

BugReports https://github.com/InfinteCuriosity/ClassificationEnsembles/issues

Config/pak/sysreqs cmake make libicu-dev libuv1-dev zlib1g-dev

Repository https://infinitecuriosity.r-universe.dev

Date/Publication 2026-06-28 00:54:13 UTC

RemoteUrl https://github.com/infinitecuriosity/classificationensembles

RemoteRef HEAD

RemoteSha 5810a92a55eef40f3f89e1ccbb42e5d68fe2ba71

Contents

Carseats . . . 2
Classification . . . 3
ClassificationEnsemblesConfig . . . 4
ClassificationEnsemblesDemo . . . 5
ClassificationEnsemblesFastConfig . . . 5
Cleveland_heart . . . 6
Dry_Beans . . . 7
load_pipeline . . . 8
Maternal_Health_Risk . . . 8
plot.classification_pipeline . . . 9
predict.classification_pipeline . . . 9
predict_production . . . 10
print.classification_pipeline . . . 11
save_pipeline . . . 11

Index **12**

`Carseats` *Carseats data*

Description

This is the Carseats data as shown in the ISLR package.

Usage

```
Carseats
```

Format

Carseats A simulated data set with 400 observations and 11 rows

Sales Unit sales (in thousands) at each location

CompPrice Price charged by competitor at each location

Income Community income level (in thousands of dollars)

Advertising Local advertising budget for company at each location (in thousands of dollars)

Population Population size in region (in thousands)

Price Price company charges for car seats at each site

ShelveLoc A factor with levels Bad, Good and Medium indicating the quality of the shelving location for the car seats at each site

Age Average age of the local population

Urban A factor with levels No and Yes to indicate whether the store is in an urban or rural location

US A factor with levels No and Yes to indicate whether the store is in the US or not

Source

ISLR data set, https://www.rdocumentation.org/packages/ISLR/versions/1.4/topics/Carseats

---

`Classification` *Comprehensive Multi-Class Classification Ensemble Pipeline Engine*

Description

Comprehensive Multi-Class Classification Ensemble Pipeline Engine

Usage

```
Classification(
  dataset = NULL,
  target_col = NULL,
  palette_style = c("standard", "viridis", "modern"),
  config = ClassificationEnsemblesConfig(),
  verbose = TRUE
)
```

Arguments

`dataset`	A data frame containing the predictors and target variable.
`target_col`	Character string specifying the name of the target categorical variable factor.
`palette_style`	Character choice layout. Visual color scheme for plot vectors: "standard", "viridis", or "modern".
`config`	A pre-configured architecture configuration parameter matrix list from `ClassificationEnsemblesConfi`
`verbose`	Logical. If TRUE, logs operational pipeline milestones to console. Default is TRUE.

Value

An object of class `classification_pipeline` containing evaluation metrics matrices, models, data dictionaries, and plots.

`ClassificationEnsemblesConfig` *Configuration Parameters Matrix for Classification Pipeline*

Description

Configuration Parameters Matrix for Classification Pipeline

Usage

```
ClassificationEnsemblesConfig(
  cv_folds = 5,
  train_pct = 0.75,
  vif_threshold = 5,
  transform_steps = c("nzv", "medianImpute", "center", "scale"),
  rf_tune = 5,
 glmnet_grid = expand.grid(alpha = seq(0, 1, length = 4), lambda = seq(0.001, 0.1,
    length = 5)),
  svm_tune_length = 5,
  nnet_tune_length = 5,
  tree_tune_length = 5,
  leverage_threshold = 999
)
```

Arguments

`cv_folds`	Integer. The number of cross-validation folds. Default is 5.
`train_pct`	Numeric. The training data proportion as a decimal between 0 and 1. Default is 0.75.
`vif_threshold`	Numeric. Variance Inflation Factor threshold for filtering multicollinear predictors. Default is 5.
`transform_steps`	Character vector. Preprocessing transforms to pass to caret. Default is c("nzv", "medianImpute", "center", "scale").
`rf_tune`	Integer. Tuning length parameter for Random Forest/Ranger models. Default is 5.
`glmnet_grid`	Optional custom tuning grid for Elastic Net models.
`svm_tune_length`	Integer. Tuning length parameter for SVM models. Default is 5.
`nnet_tune_length`	Integer. Tuning length parameter for Neural Network models. Default is 5.
`tree_tune_length`	Integer. Tuning length parameter for tree-based models (Rpart, C5.0, gbm). Default is 5.
`leverage_threshold`	Numeric. Outlier removal multiplier relative to mean hatvalues. 999 deactivates filter. Default is 999.

Value

A structured list containing isolated operational tuning parameters.

`ClassificationEnsemblesDemo`
Run Fast Validation Classification Demo Suite

Description

Run Fast Validation Classification Demo Suite

Usage

```
ClassificationEnsemblesDemo()
```

Value

An invisible trained `classification_pipeline` object run over a standard medical diagnostic set.

`ClassificationEnsemblesFastConfig`
Fast-Execution Configuration Matrix for Classification Verification

---

Description

Fast-Execution Configuration Matrix for Classification Verification

Usage

```
ClassificationEnsemblesFastConfig()
```

Value

A fast-track parameter subset list optimized for rapid package verification checks.

`Cleveland_heart` *Cleveland Heart*

Description

Posted by John Gennari, 3/13/90, This is Dr. Detrano's database modified to be a real MIXED dataset.

Usage

```
Cleveland_heart
```

Format

Cleveland_heart These are the original attributes: Attributes: 8 symbolic, 6 numeric. Age; sex; chest pain type (angina, abnang, notang, asympt) Trestbps (resting blood pres); cholesteral; fasting blood sugar < 120 (true or false); resting ecg (norm, abn, hyper); max heart rate; exercise induced angina (true or false); oldpeak; slope (up, flat, down) number of vessels colored (???); thal (norm, fixed, rever). Finally, the class is either healthy (buff) or with heart-disease (sick).

Source

https://archive-beta.ics.uci.edu/dataset/45/heart+disease/files?path=cleve.mod

The column names were corrected to be usable by R ('Chest_pain_type' instead of 'chest pain type'), removed the '<' symbol from a column name because the '<' symbol causes some errors reading the column names (such as tree models), removed three columns due to very high number of missing cells (noted as '?' in the original file.) Those three columns are 'Number_of_vessels_colored', 'Thal', and 'Resting_ECG'.

Age Age of the subject

Sex Sex of the subject, either male or female

Chest_pain_type One of angina, abnang, notang, or asympt

Resting_blood_pressure The resting blood pressure for the subject

Cholesteral The patient's cholesterol

Fasting_blood_sugar Binary, whether the fasting blood sugar is <120

Max_heart_rate The maximum measured heart rate for the patient

Exercise_induced_angina Binary, whether angina was induced due to exercise

Old_peak Numeric value

Slope Three levels: Down, flat, up

Sick_or_buff Binary, is the patient sick or buff

Class 0 is healthy, 1,2,3,4 is sick.

Dry_Beans *Dry Beans*

Description

A data set of dry beans

Usage

```
Dry_Beans
```

Format

Dry_Beans

Source

The full data set is on UCI: https://archive.ics.uci.edu/dataset/602/dry+bean+dataset

Area The area of a bean zone and the number of pixels within its boundaries

Perimeter Bean circumference is defined as the length of its border

MajorAxisLength The distance between the ends of the longest line that can be drawn from a bean

MinorAxisLength The longest line that can be drawn from the bean while standing perpendicular to the main axis

AspectRatio Defines the relationship between MajorAxisLength and MinorAxisLength

Eccentricity Eccentricity of the ellipse having the same moments as the region

ConvexArea Number of pixels in the smallest convex polygon that can contain the area of a bean seed

EquivDiameter Equivalent diameter: The diameter of a circle having the same area as a bean seed area

Extent The ratio of the pixels in the bounding box to the bean area

Solidity Also known as convexity. The ratio of the pixels in the convex shell to those found in beans.

Roundness Calculated with the following formula: (4piA)/(P^2)

Compactness Measures the roundness of an object

ShapeFactor1 Continuous value

ShapeFactor2 Continuous value

ShapeFactor3 Continuous value

ShapeFactor4 Continuous value

Class (Seker, Barbunya, Bombay, Cali, Dermosan, Horoz and Sira)

load_pipeline *Load and Decompress Package Pipeline Footprints*

Description

Load and Decompress Package Pipeline Footprints

Usage

```
load_pipeline(file_path)
```

Arguments

file_path	Character path mapping to a serialized rds file.

Value

A deserialized `classification_pipeline` object.

Maternal_Health_Risk *Maternal Health Risk*

Description

Data has been collected from different hospitals, community clinics, maternal health cares from the rural areas of Bangladesh through the IoT based risk monitoring system.

Usage

```
Maternal_Health_Risk
```

Format

Maternal_Health_Risk Age, Systolic Blood Pressure as SystolicBP, Diastolic BP as DiastolicBP, Blood Sugar as BS, Body Temperature as BodyTemp, HeartRate and RiskLevel. All these are the responsible and significant risk factors for maternal mortality, that is one of the main concern of SDG of UN.

Source

https://archive.ics.uci.edu/dataset/602/dry+bean+dataset

Age Any ages in years when a women during pregnant.

SystolicBP Upper value of Blood Pressure in mmHg, another significant attribute during pregnancy.

DiastolicBP Lower value of Blood Pressure in mmHg, another significant attribute during pregnancy.

BS Blood glucose levels is in terms of a molar concentration

BodyTemp Body temperature in Farenheit

HeartRate A normal resting heart rate

RiskLevel Predicted Risk Intensity Level during pregnancy considering the previous attribute.

`plot.classification_pipeline` *Plot Classification Pipeline Diagnostics Curves Canvas*

Description

Plot Classification Pipeline Diagnostics Curves Canvas

Usage

```
## S3 method for class 'classification_pipeline'
plot(x, pace_output = TRUE, ...)
```

Arguments

`x`	A classification_pipeline object generated by the Classification() engine.
`pace_output`	Logical. If TRUE, paces chart rendering frames via devAskNewPage. Default is TRUE.
`...`	Additional arguments passed to underlying plot methods.

`predict.classification_pipeline` *Predict with Classification Pipeline Object*

Description

Predict with Classification Pipeline Object

Usage

```
## S3 method for class 'classification_pipeline'
predict(object, newdata, model_name = "best", type = c("class", "prob"), ...)
```

Arguments

object	A trained classification_pipeline asset.
newdata	A data frame containing incoming data rows to score.
model_name	Character string specifying the target model from the leaderboard to score. Default is "best".
type	Character choice layout. Type of prediction to return: "class" for factor labels, "prob" for a probability assignment matrix. Default is "class".
...	Additional arguments passed to underlying predict methods.

Value

A factor vector or numeric probability matrix of classifications.

---

predict_production *Executive Production Projections and Probabilities Matrix*

Description

Evaluates champion architectures against un-indexed datasets, appending classification projection labels and probability assignments.

Usage

```
predict_production(object, newdata)
```

Arguments

object	A trained `classification_pipeline` asset.
newdata	A data frame containing incoming raw feature data templates.

Details

Production Prediction Wrapper with Probability Metrics

Value

A data frame matching input dimensions appended with prediction classes and matching row probability matrices for the top 3 models.

print.classification_pipeline *Print Classification Pipeline Summary Report*

Description

Print Classification Pipeline Summary Report

Usage

```
## S3 method for class 'classification_pipeline'
print(x, ...)
```

Arguments

x	A classification_pipeline object generated by the Classification() engine.
...	Additional arguments passed to underlying print methods.

save_pipeline *Compress and Save Trained Classification Pipeline Assets*

Description

Compress and Save Trained Classification Pipeline Assets

Usage

```
save_pipeline(object, file_path)
```

Arguments

object	A trained classification_pipeline asset.
file_path	Character path mapping where to save the binary output.

Value

Invisible NULL.

Index

∗ **datasets**
Carseats, 2
Cleveland_heart, 6
Dry_Beans, 7
Maternal_Health_Risk, 8

Carseats, 2
Classification, 3
ClassificationEnsemblesConfig, 4
ClassificationEnsemblesDemo, 5
ClassificationEnsemblesFastConfig, 5
Cleveland_heart, 6

Dry_Beans, 7

load_pipeline, 8

Maternal_Health_Risk, 8

plot.classification_pipeline, 9
predict.classification_pipeline, 9
predict_production, 10
print.classification_pipeline, 11

save_pipeline, 11