Package 'RBBR'

May 7, 2026

Type Package

Title Regression-Based Boolean Rule Inference

Version 0.1.0

Description Tools for regression-based Boolean rule inference in artificial intelligence studies. The package fits ridge regression models on conjunction expansions and composes interpretable rule sets. Parallel execution is supported for multi-CPU environments.

License GPL-3

URL https://github.com/CompBioIPM/RBBR

Depends R (>= 4.4)

Imports doParallel, foreach, glmnet, parallel, stats, utils

Suggests testthat

Config/testthat/edition 3

Encoding UTF-8

LazyData true

LazyDataCompression xz

Language en-US

RoxygenNote 7.3.3

NeedsCompilation no

Author Seyed Amir Malekpour [aut, cre]

Maintainer Seyed Amir Malekpour <a.malekpour@ipm.ir>

Repository CRAN

Date/Publication 2026-02-17 15:20:02 UTC

Contents

`MAGIC_data` *MAGIC data*

Description

MAGIC data

Usage

```
MAGIC_data
```

Format

An object of class `data.frame` with 19020 rows and 11 columns.

`OR_data` *OR data*

Description

OR data

Usage

```
OR_data
```

Format

An object of class `data.frame` with 1000 rows and 5 columns.

rbbr_predictor *Predict Using a Trained RBBR Model*

Description

Make predictions for new datapoints by utilizing a trained RBBR model.

Usage

```
rbbr_predictor(
  trained_model,
  data_test,
  num_top_rules = 1,
  slope = 10,
  num_cores = 1,
  verbose = FALSE
)
```

Arguments

trained_model	Model returned by ‘rbbr_train()‘
data_test	The new dataset for which we want to predict the target class or label probability. Each sample is represented as a row, and features are in columns.
num_top_rules	Number of Boolean rules with the best Bayesian Information Criterion (BIC) scores to be used for prediction. The default value is 1.
slope	The slope parameter for the sigmoid activation function. Default is 10.
num_cores	Number of parallel workers to use for computation. Adjust according to your system. Default is NA (automatic selection).
verbose	Logical. If TRUE, progress messages are shown. Default is FALSE.

Value

Numeric vector of predicted probabilities (length = nrow(data_test))

Examples

```
# Load dataset
data(example_data)

# Inspect loaded data
head(XOR_data)

# For fast run, use the first three input features to predict target class in column 11
data_train <- XOR_data[1:800, c(1,2,3,11)]
data_test <- XOR_data[801:1000, c(1,2,3,11)]

# training model
```

```
trained_model <- rbbr_train(data_train,
                            max_feature = 2,
                            num_cores = 1, verbose = TRUE)

head(trained_model$boolean_rules)

# testing model
data_test_x <- data_test[ ,1:(ncol(data_test)-1)]
labels <- data_test[ ,ncol(data_test)]

predicted_label_probabilities <- rbbr_predictor(trained_model,
                                    data_test_x,
                                    num_top_rules = 1,
                                    num_cores = 1, verbose = TRUE)

head(predicted_label_probabilities)
head(labels) # true labels
```

`rbbr_scaling` *Scale features to [0,1] range*

Description

Scales input features to the [0,1] interval using the 97.5th percentile of each feature. The last column (target) is not scaled.

Usage

```
rbbr_scaling(data)
```

Arguments

`data`	A numeric dataset. Each row is a sample and each column a feature. The target variable is expected in the last column.

Value

A dataset with scaled features (all columns except the last), capped at 0.9999.

Examples

```
# Load dataset
data(example_data)

# Inspect loaded data
head(MAGIC_data)

# Scale features
data_scaled <- rbbr_scaling(MAGIC_data)
```

```
head(data_scaled)
```

`rbbr_train` *Trains RBBR to learn Boolean rules*

Description

Regression-Based Boolean Rule (RBBR) inference is performed on datasets where the input and target features are either binarized or continuous within the [0,1] range.

Usage

```
rbbr_train(
  data,
  max_feature = 3,
  mode = "1L",
  slope = 10,
  weight_threshold = 0,
  balancing = TRUE,
  num_cores = NA,
  verbose = FALSE
)
```

Arguments

`data`	The dataset with scaled features within the [0,1] interval. Each row represents a sample and each column represents a feature. The target variable must be in the last column.
`max_feature`	The maximum number of input features allowed in a Boolean rule. The default value is 3.
`mode`	Choose between "1L" for fitting 1-layered models or "2L" for fitting 2-layered models. The default value is "1L".
`slope`	The slope parameter used in the Sigmoid activation function. The default value is 10.
`weight_threshold`	Conjunctions with weights above this threshold in the fitted ridge regression models will be printed as active conjunctions in the output. The default value is 0.
`balancing`	Logical. This is for adjusting the distribution of target classes or categories within a dataset to ensure that each class is adequately represented. The default value is TRUE. Set it to FALSE, if you don't need to perform the data balancing.
`num_cores`	Number of parallel workers to use for computation. Adjust according to your system. Default is NA (automatic selection).
`verbose`	Logical. If TRUE, progress messages and a progress bar are shown. Default is FALSE.

Value

This function outputs the predicted Boolean rules with the best Bayesian Information Criterion (BIC).

Examples

```
# Load dataset
data(example_data)

# Example for training a two-layer model
head(OR_data)

# For fast run, use the first three input features to predict target class in column five
data_train   <- OR_data[1:800, c(1,2,3,5)]
data_test    <- OR_data[801:1000, c(1,2,3,5)]

# training model
trained_model <- rbbr_train(data_train,
                            max_feature = 2,
                            mode = "2L",
                            balancing = FALSE,
                            num_cores = 1, verbose = TRUE)

head(trained_model$boolean_rules)

# testing model
data_test_x  <- data_test[ ,1:(ncol(data_test)-1)]
labels       <- data_test[ ,ncol(data_test)]

predicted_label_probabilities <- rbbr_predictor(trained_model,
                                    data_test_x,
                                    num_top_rules = 10,
                                    num_cores = 1, verbose = TRUE)

head(predicted_label_probabilities)
```

XOR_data *XOR data*

Description

XOR data

Usage

```
XOR_data
```

Format

An object of class `data.frame` with 1000 rows and 11 columns.

Index