Package 'carsAlgo'

May 8, 2026

Title Competitive Adaptive Reweighted Sampling (CARS) Algorithm

Version 0.5.0

Maintainer Md. Ashraful Haque <ashrafulhaque664@gmail.com>

Description Implements Competitive Adaptive Reweighted Sampling (CARS) algorithm for variable selection from high-dimensional dataset using Partial Least Squares (PLS) regression models. CARS algorithm iteratively applies the Monte Carlo sub-sampling and exponential variable elimination techniques to identify/select the most informative variables/features subjected to minimal cross-validated RMSE score. The implementation of CARS algorithm is inspired from the work of Li et al. (2009) <doi:10.1016/j.aca.2009.06.046>.
This algorithm is widely applied in near-infrared (NIR), mid-infrared (MIR), hyperspectral chemometrics areas, etc.

License MIT + file LICENSE

Encoding UTF-8

RoxygenNote 7.3.3

Depends R (>= 4.1.0)

Imports ggplot2, pls, rlang, stats, utils

URL https://github.com/mah-iasri/carsAlgo

BugReports https://github.com/mah-iasri/carsAlgo/issues

NeedsCompilation no

Author Md. Ashraful Haque [aut, cre],
Avijit Ghosh [aut],
Sayantani Karmakar [aut],
Harsh Sachan [aut],
Shalini Kumari [aut]

Repository CRAN

Date/Publication 2026-04-16 10:12:11 UTC

Contents

`CARSAlgorithm` *Create an object of the CARS algorithm*

Description

The `CARSAlgorithm()` function creates a configuration object for the Competitive Adaptive Reweighted Sampling (CARS) algorithm. Pass this object to `fit.CARSAlgorithm` to run variable selection on your high dimensional dataset.

Usage

```
CARSAlgorithm(max_iter = 100, N = 50, cv_folds = 5, random_state = 42)
```

Arguments

`max_iter`	Maximum number of CARS iterations. Default `100`.
`N`	Number of Monte Carlo sub-sampling runs per iteration. Default `50`.
`cv_folds`	Number of folds for k-fold cross-validation. Default `5`.
`random_state`	Integer seed for reproducibility. Default `42`.

Value

An object of class `"CARSAlgorithm"` - a named list of hyperparameters to be passed to `fit.CARSAlgorithm`.

See Also

`fit.CARSAlgorithm`

Examples

```
cars_obj <- CARSAlgorithm(max_iter = 20, N = 30, cv_folds = 5)
cars_obj
```

fit *Fit a Model Object to Data*

Description

Generic function for fitting model objects to data. Methods are dispatched based on the class of `x`.

Usage

```
fit(cars_obj, ...)
```

Arguments

`cars_obj`	A model configuration object (e.g., a `CARSAlgorithm` object).
`...`	Additional arguments passed to the specific method.

Value

Depends on the method. See `fit.CARSAlgorithm`.

See Also

`fit.CARSAlgorithm`

fit.CARSAlgorithm *Fits a CARS Object to any high dimensional dataset*

Description

Applies the CARS algorithm to a high-dimensional data matrix `X` and response vector `y`, iteratively selecting the optimal variable subset via Monte Carlo enabled PLS regression and adaptive reweighted sampling techniques.

Usage

```
## S3 method for class 'CARSAlgorithm'
fit(cars_obj, X, y, max_components = 10L, plot = TRUE, plot_path = NULL, ...)
```

Arguments

`cars_obj`	A `CARSAlgorithm` object created by `CARSAlgorithm`.
`X`	Numeric matrix of predictors (n_samples x n_features).
`y`	Numeric response vector of length n_samples.
`max_components`	Integer cap on PLS latent components. Default `10`.
`plot`	Logical. Whether to display and save the RMSECV curve. Default `TRUE`.
`plot_path`	File path for saving the RMSECV plot. Default `"../carsAlgo_rmsecv_curve.jpg"`.
`...`	Currently unused.

Details

This function iteratively:

1. Sub-samples the calibration set (Monte Carlo, `N` runs per iteration).
2. Fits a PLS model and extracts regression coefficients.
3. Selects variables by Adaptive Reweighted Sampling (ARS) proportional to absolute coefficient magnitude.
4. Evaluates the subset via k-fold cross-validation (RMSECV).
5. Retains the best subset and repeats with an exponentially shrinking variable set.

Value

A named list with:

`best_features` Sorted 1-based column indices of selected features.

`best_rmsecv` Lowest RMSECV achieved across all iterations.

`rmsecv_history` Numeric vector of best RMSECV per iteration.

`num_features_history` Integer vector of feature count per iteration.

`plot` A `ggplot2` object of the RMSECV curve.

See Also

`CARSAlgorithm`

Examples

```
set.seed(1)
X <- matrix(rnorm(100 * 200), nrow = 100)
y <- X[, 5] * 2 + X[, 50] * -1.5 + rnorm(100, sd = 0.5)

cars_obj <- CARSAlgorithm(max_iter = 15, N = 30, cv_folds = 5)
result   <- fit(cars_obj, X, y, max_components = 8)

cat("Best RMSECV      :", result$best_rmsecv, "\n")
cat("Selected features:", result$best_features, "\n")
```

`print.CARSAlgorithm` *Print method for CARSAlgorithm objects*

Description

Print method for CARSAlgorithm objects

Usage

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

Arguments

`x`	A `CARSAlgorithm` object.
`...`	Ignored.

Value

No return value, called for side effects

Index