Package: ARISE (via r-universe)

November 7, 2026

Type Package

Title Adaptive Residual-Informed Stability Ensemble Feature Selection

Version 0.1.0

Author Zardad Khan [aut], Amjad Ali [aut, cre], Naz Gul [aut], Saeed Aldahmani [aut]

Maintainer Amjad Ali <amjadali@uaeu.ac.ae>

Description Provides the Adaptive Residual-Informed Stability Ensemble ('ARISE') method for supervised feature selection in small-sample biomedical omics. The method combines complementary relevance components, class-balanced subsample stability, residual-informed redundancy control, multiclass pair coverage, nested profile racing, and soft consensus. The main function returns only the column indices of the top-ranked features. Method details are described by Zardad Khan, Amjad Ali, Naz Gul, and Saeed Aldahmani (2026) <doi:10.48550/arXiv.2608.14866>.

License GPL-3

Encoding UTF-8

LazyData false

Depends R (>= 4.1.0)

Imports e1071, matrixStats, randomForest, stats, utils

NeedsCompilation no

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

Date/Publication 2026-11-07 20:58:50 UTC

RemoteUrl https://github.com/cran/ARISE

RemoteRef HEAD

RemoteSha f1cba1b0247455b79443d7449d6e481dcf514546

Contents

ARISE-package 2
ARISEFS 2
NovartisMultiTissue 4

Index **5**

`ARISE-package` *ARISE: Adaptive Residual-Informed Stability Ensemble Feature Selection*

Description

The ARISE package provides the Adaptive Residual-Informed Stability Ensemble method for supervised feature selection in biomedical omics.

Details

The package contains one feature selection function, `ARISEFS`, and one example dataset, `NovartisMultiTissue`. The function returns only the column indices of the selected features.

Author(s)

Zardad Khan, Amjad Ali, Naz Gul, and Saeed Aldahmani

References

Khan, Zardad, Amjad Ali, Naz Gul, and Saeed Aldahmani (2026). ARISE: An adaptive residual-informed stability ensemble for feature selection in small-sample biomedical omics. *arXiv preprint arXiv:2608.14866*. doi:10.48550/arXiv.2608.14866.

See Also

`ARISEFS`, `NovartisMultiTissue`

`ARISEFS` *Adaptive Residual-Informed Stability Ensemble Feature Selection*

Description

Selects the top-ranked features using the Adaptive Residual-Informed Stability Ensemble method.

Usage

```
ARISEFS(X, Y, topK = 10)
```

Arguments

X	A real numeric or logical matrix, or a data frame with real numeric or logical vector columns. Rows are samples and columns are features. If supplied, feature names must be unique, non-missing, and non-empty.
Y	An atomic response vector or factor with one non-missing class label for each row of X. At least two classes and at least three observations per class are required. Complex values and non-finite numeric labels are not supported.
topK	A positive whole number giving the maximum number of top-ranked features to return. Defaults to 10.

Details

ARISE combines seven relevance components with class-balanced subsample stability, within-class residual redundancy control, multiclass pair coverage, nested profile racing, and soft consensus. The feature path is ranked from strongest to weakest, and the requested leading feature indices are returned.

Missing and non-finite predictor values are replaced by column medians computed from the training data. Columns with no finite values or variance at most 1e-12 are removed. Preprocessing is fitted separately on each inner training split and on the complete input for final selection. An informative error is raised when no variable feature remains in a required training split.

Profile selection uses three stratified inner folds with fixed k-nearest neighbour, radial support vector machine, and random forest learners. The internal seed is fixed at 1; the caller's random number state is restored after internal randomized operations. Results are reproducible for the same input, R version, and dependency versions.

For an unbiased assessment of a prediction model, call this function separately on each outer training split. Select features before fitting the corresponding model and apply the returned column indices to its held-out data.

Value

An unnamed integer vector containing unique column indices of the selected features in X, ordered from strongest to weakest. The length is at most topK and may be smaller when fewer usable features remain. No fitted classifier is returned.

References

Khan, Zardad, Amjad Ali, Naz Gul, and Saeed Aldahmani (2026). ARISE: An adaptive residual-informed stability ensemble for feature selection in small-sample biomedical omics. *arXiv preprint arXiv:2608.14866*. doi:10.48550/arXiv.2608.14866.

Examples

```
X <- cbind(
  signal = c(-3, -2, -1, 1, 2, 3),
  noise = c(1, 3, 2, 3, 1, 2)
)
Y <- factor(rep(c("a", "b"), each = 3L))
selected <- ARISEFS(X, Y, topK = 1)
```

```
X[, selected, drop = FALSE]
```

`NovartisMultiTissue` *Novartis Multitissue Gene Expression Dataset*

Description

Gene expression data for 102 samples and 5,565 features from four tissue classes.

Usage

```
data(NovartisMultiTissue)
```

Format

A list with two components:

X A numeric matrix with 102 samples in rows and 5,565 features in columns.

Y A factor with 102 class labels. The levels `1`, `2`, `3`, and `4` contain 25, 26, 28, and 23 samples, respectively.

Details

The matrix contains processed expression values. Sample and feature identifiers are generated as `Sample1`, `Sample2`, ... and `Feature1`, `Feature2`, ..., respectively. The supplied file does not include mappings to tissue names, gene symbols, or probe identifiers. Two columns are constant and are removed during ARISE preprocessing.

Source

Processed benchmark supplied with the ARISE source package. The associated original study is Su et al. (2002), cited below. No additional normalization is performed when loading the data.

References

Su, A. I., et al. (2002). Large-scale analysis of the human and mouse transcriptomes. *Proceedings of the National Academy of Sciences of the United States of America*, 99(7), 4465–4470. doi:10.1073/pnas.012025199.

Examples

```
data(NovartisMultiTissue)
X <- NovartisMultiTissue$X
Y <- NovartisMultiTissue$Y
```

Index

∗ **datasets**
NovartisMultiTissue, 4
∗ **methods**
ARISEFS, 2
∗ **multivariate**
ARISEFS, 2
∗ **package**
ARISE-package, 2

ARISE *(*ARISE-package*)*, 2
ARISE-package, 2
ARISEFS, 2, *2*

NovartisMultiTissue, *2*, 4