Package: vahtian.epinet (via r-universe)

July 3, 2026

Title R Interface to the EpiNet Toolkit

Version 0.1.0

Description An R interface to EpiNet, a transparent toolkit for honestly evaluated outcome models on tabular and graph-shaped data. It wraps the tested Python 'vahtian.epinet' package through 'reticulate', so the algorithms are single-sourced and cannot diverge across languages, and returns calibrated, caveated results: discrimination, calibration, bootstrap intervals, a label-permutation null, and feature importance. A research and education demonstrator, not clinical decision support.

License MIT + file LICENSE

URL https://github.com/heidihelena/epinet

BugReports https://github.com/heidihelena/epinet/issues

Encoding UTF-8

SystemRequirements Python (>= 3.10) with the 'vahtian-epinet' package

Imports reticulate

Suggests testthat (>= 3.0.0), igraph

Config/testthat/edition 3

RoxygenNote 7.3.1

Config/pak/sysreqs libpng-dev python3

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

Date/Publication 2026-07-03 10:51:32 UTC

RemoteUrl https://github.com/heidihelena/epinet

RemoteRef HEAD

RemoteSha b1076a36a1f6291fefdd1c00a8c1bc886b35987a

RemoteSubdir r/vahtian.epinet

Contents

`epinet` *Fit EpiNet's honestly-evaluated outcome model on a data frame*

Description

Builds a design matrix from the named predictors (non-numeric predictors are one-hot encoded) and fits EpiNet's outcome model with its honest-evaluation defaults: imbalance-aware tuning, calibration, a percentile bootstrap interval, an optional label-permutation null, and permutation feature importance. The computation runs in the tested Python core via reticulate.

This is a research and education demonstrator, not clinical decision support.

Usage

```
epinet(data, outcome, predictors = NULL, n_iterations = 1L,
  n_permutations = 0L, n_bootstrap = 1000L, test_size = 0.2,
  random_state = 42L, tune_threshold = FALSE)

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

## S3 method for class 'epinet'
summary(object, ...)

## S3 method for class 'epinet'
plot(x, top = 10L, ...)
```

Arguments

`data`	A data frame: one row per subject.
`outcome`	Name of the outcome (label) column.
`predictors`	Character vector of predictor columns. Defaults to every column except `outcome`.
`n_iterations`	Number of repeated train/test splits (default 1).
`n_permutations`	Label-permutation null draws; 0 disables (default 0).
`n_bootstrap`	Bootstrap resamples for the primary-split interval; 0 disables (default 1000).
`test_size`	Held-out fraction per split (default 0.2).
`random_state`	Integer seed (default 42).

tune_threshold	Tune the decision threshold on out-of-bag training scores instead of 0.5 (binary outcomes; default FALSE).
x, object	An "epinet" object.
top	Number of top features to show in plot() (default 10).
...	Unused.

Value

An object of class "epinet": a list with outcome, predictors, features_used, n, metrics, and importance.

Examples

```
## Not run:
fit <- epinet(data, outcome = "copd", predictors = c("age", "sex", "smoking"))
summary(fit)
plot(fit)
print(fit)

## End(Not run)
```

---

epinet_contestability *Score contestability against the outcome-class centroids*

Description

For every row, computes the closed-form flip-distance (how far it must move in standardized feature space to flip its nearest-centroid class), the runner-up class, and a per-feature value-of-information ranking. The lowest contest_quantile of flip-distances are flagged as the most contestable. plot() draws the contestability lens: the flip-distance distribution with the contested tail shaded, beside the value-of-information ranking.

A research and education demonstrator, not clinical decision support.

Usage

```
epinet_contestability(data, outcome, predictors = NULL,
  metric = "euclidean", contest_quantile = 0.1)

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

## S3 method for class 'epinet_contestability'
summary(object, ...)

## S3 method for class 'epinet_contestability'
plot(x, top = 10L, ...)
```

Arguments

data	A data frame: one row per subject.
outcome	Name of the outcome (label) column.
predictors	Character vector of predictor columns (default: all but outcome). Non-numeric predictors are one-hot encoded.
metric	Distance metric: "euclidean" (default) or "mahalanobis".
contest_quantile	Fraction flagged as most contestable (default 0.1).
x, object	An "epinet_contestability" object.
top	Number of value-of-information features to show in plot() (default 10).
...	Unused.

Value

An object of class "epinet_contestability": a list with per-row flip_distance and contested vectors, the contest_threshold, a flip_summary, a feature_voi ranking, and the full assignments table.

Examples

```
## Not run:
cst <- epinet_contestability(data, outcome = "copd",
                             predictors = c("age", "sex", "smoking"))
summary(cst)
plot(cst)

## End(Not run)
```

---

epinet_federated *Federate the fit across sites and check it reconstructs the centralized run*

Description

Partitions the rows across sites (by the site column if given, otherwise into n_sites balanced random groups) and runs EpiNet's federated reconstruction: only per-site aggregates cross, never rows. The reported differences from the centralized fit should be at floating-point level, demonstrating the federation is exact. plot() shows the per-site sizes and the reconstruction error against the centralized run.

A research and education demonstrator, not clinical decision support.

Usage

```
epinet_federated(data, outcome, predictors = NULL, site = NULL,
  n_sites = 2L, metric = "euclidean", contest_quantile = 0.1,
  random_state = 42L)

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

## S3 method for class 'epinet_federated'
summary(object, ...)

## S3 method for class 'epinet_federated'
plot(x, ...)
```

Arguments

data	A data frame: one row per subject.
outcome	Name of the outcome (label) column.
predictors	Character vector of predictor columns (default: all but outcome and, if used, site). Non-numeric predictors are one-hot encoded.
site	Optional column name giving each row's site. If NULL, rows are split into n_sites random groups.
n_sites	Number of synthetic sites when site is NULL (default 2).
metric	Distance metric for the contestability round-trip.
contest_quantile	Contested fraction for the contestability round-trip.
random_state	Integer seed for the random site split.
x, object	An "epinet_federated" object.
...	Unused.

Value

An object of class "epinet_federated": a list with n, n_sites, per-site sites sizes, fit_diffs (max abs mean/sd/centroid differences vs the centralized fit), and, when computable, contestability_diffs plus runner_up_match/top_voi_match.

Examples

```
## Not run:
fed <- epinet_federated(data, outcome = "copd",
                        predictors = c("age", "sex", "smoking"), n_sites = 3)
summary(fed)
plot(fed)

## End(Not run)
```

epinet_graph *Build a graph, derive graph features, and fit the outcome model*

Description

Constructs the node/edge graph, computes graph features (degree, clustering, component size, optional centrality), joins them with node attributes, and fits EpiNet's honestly-evaluated outcome model. `plot()` draws the network natively in R (coloured by outcome, sized by degree) using **igraph** when available, otherwise a degree-distribution fallback.

A research and education demonstrator, not clinical decision support.

Usage

```
epinet_graph(nodes, edges, outcome, id_column = "ID",
  source_column = "SourceID", target_column = "TargetID",
  directed = FALSE, include_centrality = FALSE, n_iterations = 1L,
  n_bootstrap = 1000L, random_state = 42L)

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

## S3 method for class 'epinet_graph'
summary(object, ...)

## S3 method for class 'epinet_graph'
plot(x, ...)
```

Arguments

`nodes`	A data frame of nodes (must include `id_column` and `outcome`).
`edges`	A data frame of edges (must include `source_column` and `target_column`).
`outcome`	Name of the outcome column in `nodes`.
`id_column`	Node id column (default "ID").
`source_column, target_column`	Edge endpoint columns (defaults "SourceID"/"TargetID").
`directed`	Treat edges as directed (default FALSE).
`include_centrality`	Also compute betweenness/closeness/PageRank (default FALSE; slower on large graphs).
`n_iterations, n_bootstrap, random_state`	Passed to the outcome model.
`x, object`	An `"epinet_graph"` object.
`...`	Unused.

Value

An object of class "epinet_graph": a list with n_nodes, n_edges, metrics, importance, feature_columns, and the nodes/edges structure used for plotting.

Examples

```
## Not run:
g <- epinet_graph(nodes, edges, outcome = "Outcome")
summary(g)
plot(g)

## End(Not run)
```

Index