Package 'SFHNV'

September 30, 2025

Type Package

Title Structural Forest for the Heterogeneous Newsvendor Model

Version 0.1.0

Description Implements the structural forest methodology for the heterogeneous newsvendor model. The package provides tools to prepare data, fit honest newsvendor trees and forests, and obtain point and distributional predictions for demand decisions under uncertainty.

URL https://github.com/MurphyLiCN/SFHNV

License MIT + file LICENSE

Depends R (>= 3.6)

Imports stats

Suggests future.apply, testthat (>= 3.0.0)

Encoding UTF-8

Config/testthat/edition 3

RoxygenNote 7.3.3

NeedsCompilation no

Author Mengfei Li [aut, cre]

Maintainer Mengfei Li <murphy.li.cn@gmail.com>

Repository CRAN

Date/Publication 2025-09-30 08:50:14 UTC

Contents

build_random_forest *Fit an SFHNV random forest*

Description

Fit an SFHNV random forest

Usage

```
build_random_forest(
  data,
  honest_ratio = 1,
  min_size = 50,
  max_depth = 50,
  num_trees = 100,
  feature_choose = "sqrt",
  parallel = TRUE,
  approximate = FALSE,
  max_candidates = 256,
  leaf_round_digits = 1L
)
```

Arguments

data	A data frame or output from NW_prepare().
honest_ratio	Ratio of the estimation subsample to the splitting subsample.
min_size	Minimum number of observations in each child node.
max_depth	Maximum depth of each tree.
num_trees	Number of trees to build.
feature_choose	Strategy for selecting features at each split. One of "sqrt", "log2", "third", or "all".
parallel	Logical; if TRUE and future.apply is available, build trees in parallel.
approximate	Logical; if TRUE, limit candidate split points for speed.
max_candidates	Maximum candidate split points per feature when approximate = TRUE.
leaf_round_digits	Rounding control for demand samples in leaf CDF estimation.

Value

A list of SFHNV trees.

Examples

```
data <- data.frame(x1 = rnorm(200), x2 = rnorm(200), D = rnorm(200), Q = rnorm(200))
forest <- build_random_forest(data, num_trees = 5, min_size = 20)
```

NW_prepare *Prepare data for SFHNV trees and forests*

Description

Converts a data frame with outcome quantities into a numeric matrix representation used by the Structural Forest for the Heterogeneous Newsvendor (SFHNV) estimators.

Usage

```
NW_prepare(data)
```

Arguments

`data`	A `data.frame` containing demand `D`, quantile `Q`, and feature columns.

Value

A list with prepared matrices (`X`), outcomes (`D`, `Q`), binary indicators (`z`), feature names, and the dimensions `n` and `p`.

Examples

```
data <- data.frame(x1 = rnorm(100), x2 = rnorm(100), D = rnorm(100), Q = rnorm(100))
prep <- NW_prepare(data)
str(prep)
```

NW_Tree *Fit a Structural Forest Heterogeneous Newsvendor tree*

Description

Builds an honest tree that estimates the structural parameter of the heterogeneous newsvendor model using the SFHNV algorithm.

Usage

```
NW_Tree(
  data,
  honest_ratio = 1,
  min_size = 50,
  max_depth = 50,
  features = NULL,
  approximate = FALSE,
  max_candidates = 256,
  leaf_round_digits = 1L
)
```

Arguments

data — A data frame or output from NW_prepare().

honest_ratio — Ratio of the estimation subsample to the splitting subsample.

min_size — Minimum number of observations in each child node.

max_depth — Maximum depth of the tree.

features — Optional subset of features (names or indices) to consider at each split.

approximate — Logical; if TRUE, limit candidate split points for speed.

max_candidates — Maximum candidate split points per feature when approximate = TRUE.

leaf_round_digits — Control the rounding of demand samples when fitting leaf CDFs. Use negative values to disable rounding.

Value

A list representing the fitted tree.

Examples

```
data <- data.frame(x1 = rnorm(200), x2 = rnorm(200), D = rnorm(200), Q = rnorm(200))
tree <- NW_Tree(data, min_size = 20, max_depth = 5)
preds <- predict_tree(tree, data)
```

predict_cdf_forest *Predict SFHNV random forest conditional CDF values*

Description

Predict SFHNV random forest conditional CDF values

Usage

```
predict_cdf_forest(
  forest,
  observations,
  d_values,
  parallel = TRUE,
  agg = "mean",
  trim_prop = 0.05
)
```

Arguments

`forest` A list of trees produced by `build_random_forest()`.

`observations` Data frame of new observations.

`d_values` Scalar or vector of demand thresholds.

`parallel` Logical; if `TRUE` and `future.apply` is available, predict in parallel.

`agg` Aggregation strategy across trees (`"mean"`, `"median"`, or `"trimmed"`).

`trim_prop` Trimming proportion when `agg = "trimmed"`.

Value

Numeric vector of CDF estimates.

Examples

```
data <- data.frame(x1 = rnorm(100), x2 = rnorm(100), D = rnorm(100), Q = rnorm(100))
forest <- build_random_forest(data, num_trees = 3, min_size = 15)
predict_cdf_forest(forest, data, d_values = 0)
```

`predict_cdf_tree` *Predict conditional CDF values from an SFHNV tree*

Description

Predict conditional CDF values from an SFHNV tree

Usage

```
predict_cdf_tree(tree, observations, d_values)
```

Arguments

`tree` An object produced by `NW_Tree()`.

`observations` Data frame of new observations containing the same features as the training data.

`d_values` Either a scalar demand threshold applied to all observations, or a numeric vector with one value per observation.

Value

Numeric vector of CDF values.

Examples

```
data <- data.frame(x = rnorm(50), D = rnorm(50), Q = rnorm(50))
tree <- NW_Tree(data, min_size = 10, max_depth = 3)
predict_cdf_tree(tree, data, d_values = 0)
```

`predict_forest` *Predict SFHNV random forest point estimates*

Description

Predict SFHNV random forest point estimates

Usage

```
predict_forest(forest, observations, trim_prop = 0.05, parallel = TRUE)
```

Arguments

`forest` A list of trees produced by `build_random_forest()`.

`observations` Data frame of new observations.

`trim_prop` Optional trimming proportion used in the robust aggregation.

`parallel` Logical; if `TRUE` and `future.apply` is available, predict in parallel.

Value

Numeric vector of aggregated predictions.

Examples

```
data <- data.frame(x1 = rnorm(100), x2 = rnorm(100), D = rnorm(100), Q = rnorm(100))
forest <- build_random_forest(data, num_trees = 3, min_size = 15)
predict_forest(forest, data)
```

`predict_tree` *Predict SFHNV tree point estimates*

Description

Predict SFHNV tree point estimates

Usage

```
predict_tree(tree, observations)
```

Arguments

`tree` An object produced by `NW_Tree()`.

`observations` Data frame of new observations containing the same features as the training data.

Value

Numeric vector of predicted structural parameters.

Examples

```
data <- data.frame(x = rnorm(50), D = rnorm(50), Q = rnorm(50))
tree <- NW_Tree(data, min_size = 10, max_depth = 3)
predict_tree(tree, data)
```

Index