Package 'rinet'

January 29, 2026

Type Package

Title Clinical Reference Interval Estimation with Reference Interval Network (RINet)

Version 0.1.0

Description Predicts statistics of a reference distribution from a mixture of raw clinical measurements (healthy and pathological). Uses pretrained CNN models to estimate the mean, standard deviation, and reference fraction from 1D or 2D sample data. Methods are described in LeBien, Velev, and Roche-Lima (2026) ``RINet: synthetic data training for indirect estimation of clinical reference distributions'' <doi:10.1016/j.jbi.2026.104980>.

License MIT + file LICENSE

Encoding UTF-8

Depends R (>= 3.5.0)

Imports reticulate

SystemRequirements Python (>= 3.8), TensorFlow (>= 2.16), Keras (>= 3.0), scikit-learn

RoxygenNote 7.3.3

NeedsCompilation no

Author Jack LeBien [aut, cre]

Maintainer Jack LeBien <`jackgl4124@gmail.com`>

Repository CRAN

Date/Publication 2026-01-29 21:50:02 UTC

Contents

rinet-package *RINet: Predict Clinical Reference Intervals from Mixture Distributions*

Description

Predict the statistics of an underlying reference distribution from a mixture distribution of raw clinical measurements (healthy and pathological patients). Uses pre-trained convolutional neural networks to estimate means, standard deviations, correlations, and reference fractions from 1D or 2D sample data.

Details

The main functions in this package are:

- `predict_rinet()`: Automatically detect and predict (recommended)
- `predict_rinet_1d()`: Predict statistics from 1D mixture samples
- `predict_rinet_2d()`: Predict statistics from 2D mixture samples

Models are automatically loaded on first use and cached for efficiency. Model files should be placed in `inst/models/` with names:

- `rinet_1d.keras` - 1D model
- `rinet_2d.keras` - 2D model
- `scaler_1d.pkl` - 1D scaler
- `scaler_2d.pkl` - 2D scaler

Author(s)

Jack

Maintainer: Jack <your.email@example.com>

predict_rinet *Predict statistics of the underlying reference distribution from mixture distributions using RINet*

Description

Automatically detects whether input data is 1D or 2D and calls the appropriate prediction function. This is the main user-facing function. It estimates the statistics of a "healthy" reference population from a mixture of healthy and pathological measurements.

Usage

```
predict_rinet(
  data,
  feature_grid_range = c(-4, 4),
  feature_grid_nbins = 100,
  verbose = 0,
  log_scale = TRUE,
  percentiles = c(0.025, 0.975),
  n_bootstrap = 0,
  confidence_level = 0.95
)
```

Arguments

`data` A numeric vector, matrix, or list. For 1D: vector or matrix with 1 column. For 2D: matrix with 2 columns. Can also be a list of such objects.

`feature_grid_range` Numeric vector of length 2 specifying the range for histogram binning. Default is c(-4, 4).

`feature_grid_nbins` Integer specifying the number of histogram bins. Default is 100.

`verbose` Integer controlling verbosity (0 = silent). Default is 0.

`log_scale` Logical indicating whether to log-transform the data before prediction. If TRUE (default), returns log-scale statistics and calculates reference intervals in the original scale. Default is TRUE.

`percentiles` Numeric vector of length 2 specifying the lower and upper percentiles for the reference interval. Default is c(0.025, 0.975).

`n_bootstrap` Integer specifying the number of bootstrap resamples for confidence intervals. Default is 0 (no bootstrap). When > 0, confidence intervals are computed for all predicted statistics using batch inference.

`confidence_level` Numeric specifying the confidence level for bootstrap intervals. Default is 0.95.

Value

A list of predictions. Each element contains:

`mean` Predicted mean(s) (log-scale if log_scale=TRUE)

`std` Predicted standard deviation(s) (log-scale if log_scale=TRUE)

`covariance` Predicted covariance matrix

`correlation` Predicted correlation (NA for 1D)

`reference_fraction` Predicted reference component fraction

`reference_interval` Reference interval in original scale (if log_scale=TRUE)

`log_scale` Logical indicating whether log-scaling was used

`bootstrap_ci` List of bootstrap confidence intervals (if n_bootstrap > 0)

Examples

```
## Not run:
  # 1D sample (using positive data for log-scale)
  sample_1d <- exp(rnorm(1000, mean = 2, sd = 0.5))
  result <- predict_rinet(sample_1d)

  # 2D sample (using positive data for log-scale)
  sample_2d <- exp(matrix(rnorm(2000, mean = 2, sd = 0.5), ncol = 2))
  result <- predict_rinet(sample_2d)

  # Multiple samples (automatically detected)
  samples <- list(exp(rnorm(1000, mean = 2, sd = 0.5)),
                  exp(rnorm(1000, mean = 2, sd = 0.5)))
  results <- predict_rinet(samples)

## End(Not run)
```

---

predict_rinet_1d *Predict statistics of the underlying reference distribution from 1D mixture distributions using RINet*

Description

Takes one or more 1D samples and predicts the underlying reference population statistics (mean, std, reference fraction) from a mixture of healthy and pathological measurements.

Usage

```
predict_rinet_1d(
  data,
  feature_grid_range = c(-4, 4),
  feature_grid_nbins = 100,
  verbose = 0,
  log_scale = TRUE,
  percentiles = c(0.025, 0.975),
  n_bootstrap = 0,
  confidence_level = 0.95
)
```

Arguments

`data` A numeric vector, matrix, or list of vectors. Each sample should contain observations from a 1D mixture distribution.

`feature_grid_range` Numeric vector of length 2 specifying the range for histogram binning. Default is c(-4, 4).

`feature_grid_nbins` Integer specifying the number of histogram bins. Default is 100.

`verbose` Integer controlling verbosity (0 = silent). Default is 0.

`log_scale` Logical indicating whether to log-transform the data before prediction. If TRUE (default), returns log-scale statistics and calculates reference intervals in the original scale. Default is TRUE.

`percentiles` Numeric vector of length 2 specifying the lower and upper percentiles for the reference interval. Default is c(0.025, 0.975).

`n_bootstrap` Integer specifying the number of bootstrap resamples for confidence intervals. Default is 0 (no bootstrap). When > 0, confidence intervals are computed for all predicted statistics.

`confidence_level`
Numeric specifying the confidence level for bootstrap intervals. Default is 0.95.

Value

A list of predictions. Each element contains:

`mean` Predicted mean (scalar, log-scale if log_scale=TRUE)

`std` Predicted standard deviation (scalar, log-scale if log_scale=TRUE)

`covariance` Covariance matrix (1x1 matrix)

`correlation` Always NA for 1D

`reference_fraction`
Predicted reference component fraction

`reference_interval`
Reference interval in original scale (if log_scale=TRUE)

`log_scale` Logical indicating whether log-scaling was used

`bootstrap_ci` List of bootstrap confidence intervals (if n_bootstrap > 0): mean_ci, std_ci, reference_fraction_ci, reference_interval_lower_ci, reference_interval_upper_ci

Examples

```
## Not run:
  # Single sample (using positive data for log-scale)
  sample1 <- exp(rnorm(1000, mean = 2, sd = 0.3))
  result <- predict_rinet_1d(sample1)
  print(result[[1]]$mean)

  # Multiple samples
  samples <- list(exp(rnorm(1000, 2, 0.3)), exp(rnorm(1000, 1.5, 0.4)))
  results <- predict_rinet_1d(samples)

## End(Not run)
```

predict_rinet_2d *Predict statistics of the underlying reference distribution from 2D mixture distributions using RINet*

Description

Takes one or more 2D samples and predicts the underlying reference population statistics (means, stds, correlation, reference fraction) from a mixture of healthy and pathological measurements.

Usage

```
predict_rinet_2d(
  data,
  feature_grid_range = c(-4, 4),
  feature_grid_nbins = 100,
  verbose = 0,
  log_scale = TRUE,
  percentiles = c(0.025, 0.975),
  n_bootstrap = 0,
  confidence_level = 0.95
)
```

Arguments

`data` A matrix or list of matrices. Each sample should be a matrix with 2 columns representing observations from a 2D mixture distribution.

`feature_grid_range` Numeric vector of length 2 specifying the range for histogram binning. Default is c(-4, 4).

`feature_grid_nbins` Integer specifying the number of histogram bins. Default is 100.

`verbose` Integer controlling verbosity (0 = silent). Default is 0.

`log_scale` Logical indicating whether to log-transform the data before prediction. If TRUE (default), returns log-scale statistics and calculates reference intervals in the original scale. Default is TRUE.

`percentiles` Numeric vector of length 2 specifying the lower and upper percentiles for the reference interval. Default is c(0.025, 0.975).

`n_bootstrap` Integer specifying the number of bootstrap resamples for confidence intervals. Default is 0 (no bootstrap). When > 0, confidence intervals are computed for all predicted statistics.

`confidence_level` Numeric specifying the confidence level for bootstrap intervals. Default is 0.95.

Value

A list of predictions. Each element contains:

`mean` Predicted means (vector of length 2, log-scale if log_scale=TRUE)

`std` Predicted standard deviations (vector of length 2, log-scale if log_scale=TRUE)

`covariance` Predicted covariance matrix (2x2 matrix)

`correlation` Predicted correlation coefficient (scalar)

`reference_fraction` Predicted reference component fraction

`reference_interval` Reference region ellipse vertices (100x2 matrix) in original scale (if log_scale=TRUE)

`log_scale` Logical indicating whether log-scaling was used

`bootstrap_ci` List of bootstrap confidence intervals (if n_bootstrap > 0): mean_ci (2x2 matrix), std_ci (2x2 matrix), correlation_ci, reference_fraction_ci

Examples

```
## Not run:
  # Single 2D sample (using positive data for log-scale)
  sample1 <- exp(matrix(rnorm(2000, mean = 2, sd = 0.3), ncol = 2))
  result <- predict_rinet_2d(sample1)
  print(result[[1]]$mean)
  print(result[[1]]$covariance)

  # Multiple samples
  samples <- list(exp(matrix(rnorm(2000, mean = 2, sd = 0.3), ncol = 2)),
                  exp(matrix(rnorm(2000, mean = 2, sd = 0.3), ncol = 2)))
  results <- predict_rinet_2d(samples)

## End(Not run)
```

Index