

Package ‘scatter’

October 14, 2022

Type Package

Title Generate Instance Space Based on Scagnostics

Version 0.0.1

Maintainer Janith C. Wanniarachchi <janithcwanni@gmail.com>

Description Generate bivariate data based on scatterplot features defined through scagnostics. Scagnostics is an exploratory graphical tool that defines nine features of a scatterplot based on the characteristics of three geometric graphs defined on the scatterplot. The exact calculation of these measurements are based on Wilkinson, L., Anand, A., & Grossman, R. (2005) <[doi:10.1109/INFVIS.2005.1532142](https://doi.org/10.1109/INFVIS.2005.1532142)>.

Set the required values for the scagnostic measurement type and the number of points that are needed to generate a bivariate dataset that gives the expected scagnostic measurements.

SystemRequirements Java 8 or higher

License MIT + file LICENSE

Encoding UTF-8

Imports scagnostics, GenSA, stats, rJava

Suggests testthat

RoxygenNote 7.1.2

NeedsCompilation no

Author Janith C. Wanniarachchi [aut, cre],
Thiyanga S. Talagala [aut] (<<https://orcid.org/0000-0002-5058-7503>>)

Repository CRAN

Date/Publication 2022-08-17 06:40:01 UTC

R topics documented:

scatterR	2
Index	3

 scatteR

Generate scatterplots based on the scagnostics measurement

Description

Generate scatterplots based on the scagnostics measurement

Usage

```
scatteR(
  measurements = c(Monotonic = 1, Outlying = 0.5),
  n_points = 50,
  init_points = NULL,
  global_min = 0.001,
  error_var = 0.001,
  epochs = 100,
  seed = 1835,
  verbose = TRUE,
  loss = "mae",
  ...
)
```

Arguments

measurements	A named vector containing the scagnostic measurements that the resulting scatterplot should have
n_points	The number of points that the resulting scatterplot should have.
init_points	The number of initial points to use to build the iterative scatterplots. Default is NULL in which case the factors of n_points will be calculated and the 25th quantile from the set of factors will be used.
global_min	the error that the resulting scagnostics can give
error_var	the variance of the random error that is added on to the existing points
epochs	number of epochs the optimization method should run
seed	The random number generation seed
verbose	Logical. TRUE means that messages from the optimization algorithm are shown. Default is TRUE
loss	The loss function to be used within the scatteR program. Can be either 'mae' or 'mse'. Default is 'mae'
...	Extra arguments to be used in the control argument of the GenSA function of the GenSA package

Value

A bivariate data.frame with two columns named x and y that gives a roughly similar scagnostic measurement to the 'measurements' argument

Index

scatter, [2](#)