Package: RadialVisGadgets (via r-universe)

September 12, 2024

Type Package

Title Interactive Gadgets for Radial Visualization Approaches

Version 0.2.0

URL <https://github.com/jmatute/RadialShinyGadgets>

Imports import, ggplot2, tidyr, dplyr, miniUI, shiny, shinyjs, caret, rlang, shinyscreenshot

Maintainer José Matute <jmatuteflores@gmail.com>

Description Shiny-based interactive gadgets of radial visualization methods and extensions thereof.

License MIT + file LICENCE

Encoding UTF-8

LazyData true

RoxygenNote 7.1.1

Suggests knitr, rmarkdown, datasets, clValid

VignetteBuilder knitr

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

RemoteUrl https://github.com/jmatute/radialshinygadgets

RemoteRef HEAD

RemoteSha 6f31b7c0b0ed1c0f0401c8dc5d04fce61d70f757

Contents

StarCoordinates   *Star Coordinates Gadget*

Description

Creates a RShiny Gadget for Star Coordinates

Usage

```
StarCoordinates(
  df,
  color = NULL,
  approach = "Standard",
  numericRepresentation = TRUE,
  meanCentered = TRUE,
  projMatrix = NULL,
  clusterFunc = NULL
)
```

Arguments

df	A dataframe with the data to explore. It should contain only numeric or factor columns.
color	column where labels from the data are extracted.
approach	Standard approach as defined by Kandogan, or Orthographic Star Coordinates (OSC) with a recondition as defined by Lehmann and Thiesel
numericRepresentation	
	if true attempt to convert all factors to numeric representation, otherwise used mixed representation as defined in Hinted Star Coordinates
meanCentered	center the projection at the mean of the values. May allow for easier value estimation
projMatrix	a pre-defined projection matrix as an initial configuration. Should be defined in the same fashion as the output
clusterFunc	function to define hints, assume increase in value of the function is an increase in quality of the projection. The function will be called with two parameters (points, labels)

Details

Star Coordinate's (SC) goal is to generate a configuration which reveals the underlying nature of the data for cluster analysis, outlier detection, and exploratory data analysis, e.g., by investigating the effect of specific dimensions on the separation of the data. Traditional SC are defined for multi-dimensional numerical data sets $X = \{\mathbf{p}_1, \ldots, \mathbf{p}_N\}$, for N data points $\mathbf{x}_i \in \mathbf{R}^d$ of dimensionality d. Let $A = \{\mathbf{a}_1, \ldots, \mathbf{a}_d\}$, be a set of (typically 2D) vectors, each corresponding to one of the d

dimensions. The projection $\mathbf{p}'_i \in \mathbf{R}^2$, of a multidimensional point $\mathbf{p}_i = (p_{i1}, \ldots, p_{id}) \in \mathbf{R}^d$, in SC is then defined as:

$$\mathbf{x}'_i = \sum_{j=1}^{d} \mathbf{a}_j g_j(\mathbf{p}_i),$$

with

$$g_j(\mathbf{p}_i) = \frac{p_{ij} - min_j}{max_j - min_j},$$

and (min_j, max_j), denoting the value range of dimension j.

In the case of categorical dimensions, the values when numericRepresentation= TRUE are mapped into numerical type i.e. as.numeric() However equally spaced categorical points may not reflect the true nature of the data. Instead, a frequency-based representation may be applied for individual data points. Assuming a categorical dimension j, we calculate the frequency f_{jk}, of each category k of dimension j. The respective axis vector $\mathbf{a}_j$, is then divided into according blocks, whose size represent the relative frequency (or probability) $\frac{f_{jk}}{\sum_{l=1}^{m} f_{jl}}$, of each of the m categories of dimension j.

In summary, given an order for each categorical dimension, the Equation $g()$, above can be extended to SC for mixed data by:

$$g_j(\mathbf{x}_i) = F_j(x_{ij}) - \frac{P_j(x_{ij})}{2},$$

if categorical/ordinal

$$g_j(\mathbf{x}_i) = \frac{x_{ij} - min_j}{max_j - min_j},$$

if numerical

where F_j, is the cumulative density function for (categorical/ordinal) dimension j and P_j, its probability function.

Value

A list with the projection matrix, coordinates of the projected samples and a logical vector with the selected samples

References

Kandogan, E. (2001, August). Visualizing multi-dimensional clusters, trends, and outliers using star coordinates. In Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 107-116).

Lehmann, D. J., & Theisel, H. (2013). Orthographic star coordinates. IEEE Transactions on Visualization and Computer Graphics, 19(12), 2615-2624.

Rubio-Sánchez, M., & Sanchez, A. (2014). Axis calibration for improving data attribute estimation in star coordinates plots. IEEE transactions on visualization and computer graphics, 20(12), 2013-202

Matute, J., & Linsen, L. (2020, February). Hinted Star Coordinates for Mixed Data. In Computer Graphics Forum (Vol. 39, No. 1, pp. 117-133).

Examples

```
if (interactive()) {
 library(RadialVisGadgets)
 library(datasets)
 data(iris)
 StarCoordinates(iris, "Species")
}
```

Index