Package: long2lstmarray (via r-universe)

September 18, 2024

Title Longitudinal Dataframes into Arrays for Machine Learning Training

Version 0.2.0

Maintainer Luis Garcez <luisgarcez1@gmail.com>

Description An easy tool to transform 2D longitudinal data into 3D arrays suitable for Long short-term memory neural networks training. The array output can be used by the 'keras' package. Long short-term memory neural networks are described in: Hochreiter, S., & Schmidhuber, J. (1997) <doi:10.1162/neco.1997.9.8.1735>.

Imports abind, dplyr

License GPL (>= 3)

Encoding UTF-8

LazyData true

Roxygen list(markdown = TRUE)

RoxygenNote 7.1.1

URL https://github.com/luisgarcez11/long2lstmarray

BugReports https://github.com/luisgarcez11/long2lstmarray/issues

Suggests knitr, rmarkdown, testthat

Depends R (>= 2.10)

VignetteBuilder knitr

Repository https://luisgarcez11.r-universe.dev

RemoteUrl https://github.com/luisgarcez11/long2lstmarray

RemoteRef HEAD

RemoteSha bcfb6246cd2945384ab07ff7c45dcdf7b38633cb

Contents

`alsfrs_data` *Clinical scale example data*

Description

An example dataset containing Amyotrophic Lateral Sclerosis Functional Rating Scale - Revised.

Usage

```
alsfrs_data
```

Format

A data frame with 100 rows and 15 variables:

subjid Subject ID

visdy Visit day

p1 Scale items

p2 Scale items

p3 Scale items

p4 Scale items

p5 Scale items

p6 Scale items

p7 Scale items

p8 Scale items

p9 Scale items

p10 Scale items

x1r Scale items

x2r Scale items

x3r Scale items

Source

https://pubmed.ncbi.nlm.nih.gov/10540002/

get_var_array *Generate a matrix with various lags from a variable in the dataframe*

Description

Generate a matrix with various lags from a variable in the dataframe

Usage

```
get_var_array(
  data,
  subj_var,
  var,
  time_var,
  lags,
  label_length = 1,
  label_output = FALSE
)
```

Arguments

data	A data frame, data frame extension (e.g. a tibble).
subj_var	A character string referring to the variable that specifies the "subject" variable.
var	A character string referring to the variable that contains the variable values.
time_var	A character string referring to the variable that contains the time variable values (e.g. visit day, minutes, years).
lags	The length of each sliced sequence.
label_length	How many values after are considered to be the label? Default to 1. If label_length = 1, the label value´ is always the value following the sliced sequence.
label_output	logical. if TRUE a list including the matrix with the sliced sequences and a vector with the label is returned.

Value

If label_output is FALSE, a matrix with the sliced sequences is returned. If label_output is TRUE, a list with the matrix and vector with the labels from the same variable is returned.

Examples

```
get_var_array(alsfrs_data, "subjid",
"p2", "visdy", lags = 3,
label_output = FALSE)
```

get_var_sequence *Get variable values from subject/variable name pair*

Description

Get variable values from subject/variable name pair

Usage

```
get_var_sequence(data, subj_var, subj, var)
```

Arguments

data	A data frame, data frame extension (e.g. a tibble).
subj_var	A character string referring to the variable that specifies the "subject" variable.
subj	Any value that the "subject" variable can take.
var	A character string referring to the variable that contains the variable values.

Value

A vector of values from variable var which subj_var equal to subj.

Examples

```
get_var_sequence(sleep, subj_var = "ID", 1, "extra")
```

longitudinal_array *Generate a matrix with various lags from a dataframe*

Description

Generate a matrix with various lags from a dataframe

Usage

```
longitudinal_array(
  data,
  subj_var,
  vars,
  time_var,
  lags,
  label_length = 1,
  label_var = NULL,
  label_output = FALSE,
  time_var_output = FALSE
)
```

Arguments

data	A data frame, data frame extension (e.g. a tibble).
subj_var	A character string referring to the variable that specifies the "subject" variable.
vars	A character string referring to the variables that contain the variable values.
time_var	A character string referring to the variable that contains the time variable values (e.g. visit day, minutes, years). Important to get the sequences in the right order.
lags	The length of each sliced sequence.
label_length	How many values after are considered to be the label? Default to 1. If label_length = 1, the label value is always the value following the sliced sequence.
label_var	A character string referring to the variables that contain the label variable values.
label_output	logical. if TRUE a list including the matrix with the sliced sequences and a vector with the label is returned.
time_var_output	logical. Is time_var to be included in the final output. Default to FALSE.

Value

If label_output is FALSE, a 3D array with the sliced sequences is returned. The array dimensions are subject, time and variable. If label_output is TRUE, a list with the array and vector with the labels is returned.

Examples

```
longitudinal_array(alsfrs_data, "subjid", c("p1", "p2", "p3"),
                   "visdy", lags = 3, label_output = FALSE)
longitudinal_array(alsfrs_data, "subjid", c("p1", "p2", "p3"),
                   "visdy", lags = 3, label_output = FALSE)[1,,]
longitudinal_array(alsfrs_data, "subjid", c("p1", "p2", "p3"),
                  "visdy", lags = 3, label_output = FALSE)[,1,]
longitudinal_array(alsfrs_data, "subjid", c("p1", "p2", "p3"),
                  "visdy", lags = 3, label_output = FALSE)[,,1]
```

---

slice_var_sequence *Generate a matrix with various lags from a sequence*

Description

Generate a matrix with various lags from a sequence

Usage

```
slice_var_sequence(sequence, lags, label_length = 1, label_output = TRUE)
```

Arguments

sequence	A vector representing the sequence to be sliced into many rows.
lags	The length of each sliced sequence.
label_length	How many values after are considered to be the label? Default to 1. If label_length = 1, the label value is always the value following the sliced sequence.
label_output	logical. if TRUE a list including the matrix with the sliced sequences and a vector with the labels is returned.

Value

If label_output is FALSE, a matrix with the sliced sequences is returned. If label_output is TRUE, a list with the matrix and vector with the labels is returned.

Examples

```
slice_var_sequence(sequence = 1:30,
 lags = 3, label_length = 1,
 label_output = TRUE)

slice_var_sequence(sequence = 1:30,
lags = 3, label_length = 1,
 label_output = FALSE)

slice_var_sequence(sequence = 1:30,
 lags = 3, label_length = 2,
  label_output = FALSE)
```

Index