map.build

Description

Constructs a SOM, returns an object of class 'map'.

Usage

map.build(data, labels=NULL, xdim=10, ydim=5, alpha=0.3, train=1000, algorithm="vsom")

Arguments

data       a dataframe where each row contains an unlabeled training instance.
labels     a vector or dataframe with one label for each observation in data.
xdim       the x-dimension of the map. (default=10)
ydim       the y-dimension of the map. (default=5)
alpha      the learning rate, should be a positive non-zero real number. (default=0.3)
train      the number of training iterations. (default=1000)
algorithm  training algorithm selection switch. (default="vsom")

Value

object of type 'map'.

Note

You have a choice of training algorithms:

- "vsom" - vectorized stochastic learning, this is a highly optimized version of stochastic training written in FORTRAN 9X.
- "som" - the traditional stochastic learning algorithm written in C++/TNT.
- "batchsom" - batch version of the SOM training algorithm, written in C.
- "experimental" - experimental implementation of vectorized stochastic learning implemented directly in R.

Note

If your training data does not have any labels you can construct a simple label vector as follows: labels <- 1:nrow(training.data). If you let the labels default to the NULL value then no labels will be shown in the map visualization.
Examples

```r
data(iris)

## set data frame and labels
df <- subset(iris, select=-Species)
labels <- subset(iris, select=Species)

## build a map
m <- map.build(df, labels, xdim=15, ydim=10, train=1000)
```

Description

Evaluate the quality of a SOM using embedding and estimated topographic accuracy.

Usage

```r
map.convergence(map, conf.int=.95, k=50, verb=FALSE, ks=TRUE)
```

Arguments

- `map`: an object of type `map`
- `conf.int`: is the confidence interval of the quality assessment (default 95%)
- `k`: number of samples to use in the computation of the estimated topographical accuracy (default=50)
- `verb`: if true reports the two convergence components separately, otherwise it will report a linear combination of the two indices (default=FALSE)
- `ks`: if true uses the Kolmogorov-Smirnov convergence test otherwise a convergence test based on variance and means is performed (default=TRUE)

Value

A single value or a pair of values: 1) embedding accuracy 2) estimated topographic accuracy.

Author(s)

Lutz Hamel

References

Examples

```r
data(iris)

## set data frame and labels
df <- subset(iris, select=-Species)
labels <- subset(iris, select=Species)

## build a map
m <- map.build(df, labels, xdim=15, ydim=10, train=1000)

## map quality
map.convergence(m)
```

---

### `map.embed`

**Map Embedding Accuracy**

<table>
<thead>
<tr>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluates how well a map models the underlying training data distribution.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>map.embed(map, conf.int = 0.95, verb=FALSE, ks=FALSE)</code></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Arguments</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>map</code> an object of type 'map'.</td>
</tr>
<tr>
<td><code>conf.int</code> the confidence interval of the embedding test (default 95 percent).</td>
</tr>
<tr>
<td><code>verb</code> a switch controlling the structure of the output value (default=FALSE)</td>
</tr>
<tr>
<td><code>ks</code> if true uses the Kolmogorov-Smirnov convergence test otherwise a convergence test based on variance and means is performed (default=FALSE)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>The embedding accuracy of the map. If the switch <code>verb=TRUE</code> then a vector of the individual feature embedding accuracies are returned.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>map.embed</code> has been deprecated as of this release (version 4.3.0) of the popsom package for R. <code>map.embed</code> has also been flagged for removal upon release of the next version of the popsom package.</td>
</tr>
<tr>
<td>The use of <code>map.convergence</code>, also available in the popsom package, with the argument 'verb = True' is now encouraged to explicitly compute the embedding accuracy of maps in future popsom package releases.</td>
</tr>
</tbody>
</table>
map.marginal

Author(s)

Lutz Hamel, Benjamin Ott, Gregory Breard, Robert Tatoian

References


Examples

data(iris)

## set data frame and labels
df <- subset(iris, select=-Species)
labels <- subset(iris, select=Species)

## build a map
m <- map.build(df, labels, xdim=15, ydim=10, train=1000)

## display the embedding accuracy of the map
map.embed(m)

## display the embedding accuracies of the individual features
data.frame(names(df),map.embed(m,verb=TRUE))

map.marginal

Plot Marginal Distribution

Description

Generate a plot that shows the marginal probability distribution of the neurons and data.

Usage

map.marginal(map,marginal)

Arguments

map an object of type 'map'.
marginal is the name of a training data frame dimension or index.

Author(s)

Lutz Hamel, Robert Tatoian
map.neuron

Return a Map Neuron

Description

Returns the contents of a neuron at (x,y) on the map as a vector.

Usage

map.neuron(map, x, y)

Arguments

map  an object of type 'map'.
x   map x-coordinate of neuron.
y  map y-coordinate of neuron.

Value

A vector representing the neuron.

Author(s)

Lutz Hamel

Examples

data(iris)

## set data frame and labels
df <- subset(iris, select=-Species)
labels <- subset(iris, select=Species)

## build a map
m <- map.build(df, labels, xdim=15, ydim=10, train=1000)

## display marginal distribution of dimension 1
map.marginal(m,1)
## display the neuron at position (1,1)
map.neuron(m,1,1)

### Description
Prints the association of labels with map elements.

### Usage
```r
map.projection(map)
```

### Arguments
- **map**
  - an object of type 'map'.

### Value
a dataframe containing the projection onto the map for each training observation.

### Note
map.projection has been deprecated as of this release (version 4.3.0) of the popsom package for R. map.projection has also been flagged for removal upon release of the next version of the popsom package.

### Author(s)
Lutz Hamel, Benjamin Ott, Gregory Breard

### Examples
```r
data(iris)

## set data frame and labels
df <- subset(iris,select=-Species)
labels <- subset(iris,select=Species)

## build a map
m <- map.build(df, labels, xdim=15, ydim=10, train=1000)

## display the label association for the map
map.projection(m)
```
map.significance  Compute Significance Of Features

Description
Computes the relative significance of each feature and plots it.

Usage
map.significance(map, graphics = TRUE, feature.labels = TRUE)

Arguments
map an object of type 'map'.
graphics a switch that controls whether a plot is generated or not.
feature.labels a switch to allow the plotting of feature names vs feature indices.

Value
if graphics=FALSE a vector containing the significance for each feature is returned.

Note
We use a Bayesian approach to compute the relative significance of features based on variance.

Author(s)
Lutz Hamel, Benjamin Ott, Gregory Breard

References

Examples
data(iris)

## set data frame and labels
df <- subset(iris,select=-Species)
labels <- subset(iris,select=Species)

## build a map
m <- map.build(df, labels, xdim=15, ydim=10, train=1000)

## show the relative feature significance for each feature
data.frame(names(df),map.significance(m,graphics=FALSE))

## display the relative feature significance graphically
map.significance(m)
map.starburst

Generate Starburst For Map

Description
Computes and displays the starburst representation of clusters on a map.

Usage
map.starburst(map, explicit = FALSE, smoothing = 2, merge.clusters=FALSE, merge.range=.25)

Arguments
map
an object of type 'map'.

explicit
controls the shape of the connected components.

smoothing
controls the smoothing level of the map display (NULL, 0, >0).

merge.clusters
is a switch that controls if the starburst clusters are merged together

merge.range
is a range that is used as a percentage of a certain distance in the code to determine whether components are closer to their centroids or centroids closer to each other.

Author(s)
Lutz Hamel, Benjamin Ott, Gregory Breard, Robert Tatoian, Vishakh Gopu

References

Examples

data(iris)

## set data frame and labels
df <- subset(iris,select=-Species)
labels <- subset(iris,select=Species)

## build a map
m <- map.build(df, labels, xdim=15, ydim=10, train=1000)

## display the starburst for the map
map.starburst(m)
map.topo

Estimated Topographical Accuracy

Description
Evaluate the topological quality of a SOM using the estimated topographical accuracy.

Usage
map.topo(map,k=50,conf.int = 0.95,verb=FALSE,interval=TRUE)

Arguments

map an object of type ’map’.

k number of samples to use in the computation of the estimated topographical accuracy (default=50)

conf.int the confidence interval of the estimated topographical accuracy (default 95 percent).

verb a switch controlling the structure of the output value (default=FALSE)

interval a switch that controls whether the confidence interval is computed (default=TRUE)

Value
1) The value of the estimated topographical accuracy. 2) The low value of the confidence interval and the high value of the confidence interval ’conf.int’ if interval=TRUE. 3) If verb=TRUE then map.accuracy will return a vector with the accuracies of the individual k samples.

Note
map.topo has been deprecated as of this release (version 4.3.0) of the popsom package for R. map.topo has also been flagged for removal upon release of the next version of the popsom package.

The use of map.convergence, also available in the popsom package, with the argument ’verb = True’ is now encouraged to explicitly compute the estimated topographic accuracy of maps in future popsom package releases.

Author(s)
Lutz Hamel

References
Examples

data(iris)

## set data frame and labels
df <- subset(iris, select=-Species)
labels <- subset(iris, select=Species)

## build a map
m <- map.build(df, labels, xdim=15, ydim=10, train=1000)

## display estimated topographical accuracy of the map
map.topo(m)

---

Description

These functions have been deprecated as of this release (version 4.3.0) of the popsom package for R.
Said functions have been flagged for removal upon release of the next version of the popsom package.

Details

The functions currently included in the popsom that have been deprecated are as follows:

1) map.embed()
   – (NOTE: To explicitly compute the embedding accuracy of maps in future popsom package releases, call map.convergence() with the argument 'verb = True')

2) map.topo()
   – (NOTE: To explicitly compute the estimated topographic accuracy of maps in future popsom package releases, call map.convergence() with the argument 'verb = True')

3) map.projection()

Author(s)

Lutz Hamel, Benjamin Ott, Greg Breard, University of Rhode Island with Robert Tatoian, Vishakh Gopu, and Michael Eiger
Index

map.build, 2
map.convergence, 3
map.embed, 4
map.marginal, 5
map.neuron, 6
map.projection, 7
map.significance, 8
map.starburst, 9
map.topo, 10

popsom-deprecated, 11