Package 'KLexp'

October 12, 2022

Title Kernel_lasso Expansion

Version 1.0.0

Maintainer Zongrui Dai <dzr17723980497@gmail.com>

Description Provides the function to calculate the kernel-lasso expansion, Z-score, and max-min-scale standardization.It can increase the dimension of existed dataset and remove abundant features by lasso. Z Dai, L Jiayi, T Gong, C Wang (2021) <doi:10.1088/1742-6596/1955/1/012047>.

License GPL-2

URL https://github.com/Zongrui-Dai/Kernel-lasso-feature-expansion

Encoding UTF-8

RoxygenNote 7.1.1.9001

Depends glmnet (>= 4.1-2)

Imports graphics, stats

NeedsCompilation no

Author Zongrui Dai [aut, cre] (<https://orcid.org/0000-0002-7893-5004>)

Repository CRAN

Date/Publication 2021-08-21 09:30:08 UTC

R topics documented:

gauss *Gauss function*

Description

Gauss function

Usage

```
gauss(d1, d2, sigma = 0.5)
```

Arguments

d1	vector1
d2	vector2
sigma	The hyperparameter of RBF kernel function, which indicates the width.

Value

Calculate the Gauss function

Author(s)

Zongrui Dai

Source

https://github.com/Zongrui-Dai/Kernel-lasso-feature-expansion

Examples

```
##
data(iris,package = 'datasets')
w<-gauss(iris[,1],iris[,2])
print(w)
```

kernel_lasso_expansion *kernel_lasso_expansion*

Description

Kernel_lasso is one feature selection method, which combines the feature expansion and lasso regression together. Kernel function will increase the dimensions of the existed data and then reduce the features by lasso. 'glmnet' package should be higher than 4.1-2.

Arguments

x	Your input features, which have to be data.frame with at least two variables.
y	The dependent variable
sigma	The hyperparameter of RBF kernel function, which indicates the width.
dataframe	Wether the data is dataframe. The default is TURE
standard	Using 'max_min_scale' or 'Z_score' method to standardize the data. NULL means no standardization

Value

The result is stored in one list which contains the orignial dataset, amplified dataset, final features, and lasso output.

Author(s)

Zongrui Dai

Source

https://github.com/Zongrui-Dai/Kernel-lasso-feature-expansion

References

Z. Dai, J. Li, T. Gong, C. Wang (2021), Kernel_lasso feature expansion method: boosting the prediction ability of machine learning in heart attack," 2021 IEEE. About Kernel-lasso feature expansion method: boosting the prediction ability of machine learning in heart attack" 2021 IEEE.

Examples

```
##Regression (MSE)
data(attenu,package = 'datasets')
result<-kernel_lasso_expansion(x=attenu[,-c(3,5)],y=attenu[,5],
standard = 'max_min',sigma=0.01,control = lasso.control(nfolds=3,type.measure = 'mse'))
summary(result)

#Plot the lasso
```

```
plot(result$lasso)

#Result
result$original ##The original feature space
result$expansion  ##The feature space after expansion
result$final_feature  ##The name of the final feature
result$final_data  ##The dataframe of final feature
```

`lasso.control` *lasso.control*

Description

The same function from glmnet, which controls the training of lasso.

Usage

```
lasso.control(nfolds = 10, trace.it = 1, type.measure = "auc")
```

Arguments

`nfolds`	n-fold cross-validation.
`trace.it`	Whether to plot the training process
`type.measure`	Choose the loss funcrion.

Value

Will return the lasso training setting

Author(s)

Zongrui Dai

Source

https://github.com/Zongrui-Dai/Kernel-lasso-feature-expansion

Examples

```
##10-fold Cross-validation with MSE as loss function
c<-lasso.control(nfolds=10,type.measure='mse')
```

max_min_scale *max_min_scale*

Description

max_min_scale is used to calculate the standardization value of data.The formula is (x-min(x))/(max(x)-min(x)). It can compress the data into the (0,1).

Arguments

data	Your input data, which can be numerci or data.frame
dataframe	Wether the data is dataframe. The default is False(numeric)

Value

Calculate the max-min standardization of the dataset by the formula: (max(x)-x)/(max(x)-min(x))

Author(s)

Zongrui Dai

Source

https://github.com/Zongrui-Dai/Kernel-lasso-feature-expansion

Examples

```
##For the numeric data
data(iris,package = 'datasets')
w<-max_min_scale(iris[,1])
print(w)

##For the data.frame data
w1<-max_min_scale(iris[,-5],dataframe=TRUE)
print(w1)
```

Z_score *Z_score standardization*

Description

Z-score method is used to calculate the standardization value of data.The formula is (x-mean(x))/var(x). It can compress the data into the (0,1).

Usage

```
Z_score(data, dataframe = FALSE)
```

Arguments

data	Your input data, which can be numerci or data.frame
dataframe	Wether the data is dataframe. The default is False(numeric)

Value

Calculate the Z_score standardization of the dataset by the formula: (x-mean(x))/var(x)

Author(s)

Zongrui Dai

Source

https://github.com/Zongrui-Dai/Kernel-lasso-feature-expansion

Examples

```
##For the numeric data
data(iris,package = 'datasets')
w<-Z_score(iris[,1])
print(w)

##For the data.frame data
w1<-Z_score(iris[,-5],dataframe=TRUE)
print(w1)
```

Index