Package 'SKNN'

June 11, 2022

Type Package

Title A Super K-Nearest Neighbor (SKNN) Classification Algorithm

Version 3.1

Date 2022-06-11

Maintainer Yi Ya <Yi.YA_yaya@hotmail.com>

Depends methods, stats

Description It's a Super K-Nearest Neighbor classification method with using kernel density to describe weight of the distance between a training observation and the testing sample.

Collate SKNN.R Dist.R PCAy.R AllClasses.R KNN.R

License GPL-2

NeedsCompilation no

Author Yi Ya [aut, cre],
Nader Ebrahimi [aut],
Yoram Rubin [aut],
Jacob Zhang [aut]

Repository CRAN

Date/Publication 2022-06-11 17:00:02 UTC

R topics documented:

SKNN-package *Super K-Nearest Neighbor (SKNN) Classification*

Description

It's a Super K-Nearest Neighbor classification method with using kernel density to describe the weight of the distance between a training observation and the sample to be classified.

Details

Package:	SKNN
Type:	Package
Version:	3.1
Date:	2022-06-11
License:	GPL-2

Author(s)

Yi Ya, Nader Ebrahimi, Yoram Rubin, and Jacob Zhang

References

Yarong Yang(Yi Ya).(2022) SKNN: A Super K-Nearest Neighbor Classification Algorithm. submitted to Journal of Statistical Software

Yarong Yang, Matt Over, and Yoram Rubin.(2012) Strategic Placement of Localization Devices (such as Pilot Points and Anchors) in Inverse Modeling Schemes. Water Resources Research, 48, W08519, doi:10.1029/2012WR011864.

B.B.W. Silverman.(1986) Density Estimation for Statistics and Data Analysis. London: Chapman and Hall.

Examples

```
Sepal.Length<-c(4.8, 5.1, 4.6, 5.3, 5.0, 5.7, 5.7, 6.2, 5.1, 5.7, 6.7, 6.3, 6.5, 6.2, 5.9)
Sepal.Width<-c(3.0, 3.8, 3.2, 3.7, 3.3, 3.0, 2.9, 2.9, 2.5, 2.8, 3.0, 2.5, 3.0, 3.4, 3.0)
Petal.Length<-c(1.4, 1.6, 1.4, 1.5, 1.4, 4.2, 4.2, 4.3, 3.0, 4.1, 5.2, 5.0, 5.2, 5.4, 5.1)
Petal.Width<-c(0.3, 0.2, 0.2, 0.2, 0.2, 1.2, 1.3, 1.3, 1.1, 1.3, 2.3, 1.9, 2.0, 2.3, 1.8)
Species<-as.factor(c(rep("red",5),rep("blue",5),rep("green",5)))
iris<-cbind(Sepal.Length,Sepal.Width,Petal.Length,Petal.Width)
Res<-length(nrow(iris))
k<-10
for(i in 1:nrow(iris))
     Res[i]<-SKNN(data=iris,Class=as.vector(Species),k=k,test=iris[i,])
accuracy<-length(which(Res==Species))/length(Species)
```

```
plot(x=1:15,y=rep(1,15),col=as.vector(Species),lwd=4,ylim=c(0,3),xlab="",ylab="",
yaxt = "n",xaxt="n")
par(new=TRUE)
plot(x=1:15,y=rep(2,15),col=Res,lwd=4,ylim=c(0,3),xlab="",ylab="",yaxt = "n",xaxt="n")
ind<-which(Res!=Species)
if(length(ind)>0) {
  for(j in 1:length(ind))
      lines(x=c(ind[j],ind[j]),y=c(1+0.05,2-0.05))
}
text(5,0.3,paste("SKNN Misclassified:",length(ind)))
axis(2,at=2,labels="SKNN",las=1)
text(10,2.5,paste("k: ",k))
```

`Dist` *Finding the distance between two observations.*

Description

It's a function of finding the distance between two observations.

Usage

```
Dist(x,y)
```

Arguments

`x`	Numeric. A vector denoting an observation.
`y`	Numeric. A vector denoting an observation.

Value

A list.

Author(s)

Yi Ya

KNN *K-Nearest Neighbor Classification*

Description

It's implementation of the K-Nearest Neighbor classification method for data of any number of dimentions.

Usage

```
KNN(data, Class, k, test)
```

Arguments

data	Numeric. The training data set, matrix.
Class	Character. Class of the training observations, vector.
k	Integer. The number of K to be used.
test	Numeric. The sample to be classified.

Value

A character.

Author(s)

Yi Ya

References

A.T. Covert and P. Hart. Nearest Neighbor Pattern Classification. IEEE Transactions on Information Theory, 13(1): 21-27, 1967.

Examples

```
Sepal.Length<-c(4.8, 5.1, 4.6, 5.3, 5.0, 5.7, 5.7, 6.2, 5.1, 5.7, 6.7, 6.3, 6.5, 6.2, 5.9)
Sepal.Width<-c(3.0, 3.8, 3.2, 3.7, 3.3, 3.0, 2.9, 2.9, 2.5, 2.8, 3.0, 2.5, 3.0, 3.4, 3.0)
Petal.Length<-c(1.4, 1.6, 1.4, 1.5, 1.4, 4.2, 4.2, 4.3, 3.0, 4.1, 5.2, 5.0, 5.2, 5.4, 5.1)
Petal.Width<-c(0.3, 0.2, 0.2, 0.2, 0.2, 1.2, 1.3, 1.3, 1.1, 1.3, 2.3, 1.9, 2.0, 2.3, 1.8)
Species<-as.factor(c(rep("red",5),rep("blue",5),rep("green",5)))
iris<-cbind(Sepal.Length,Sepal.Width,Petal.Length,Petal.Width)
Res<-length(nrow(iris))
k<-10
for(i in 1:nrow(iris))
     Res[i]<-KNN(data=iris,Class=as.vector(Species),k=k,test=iris[i,])
accuracy<-length(which(Res==Species))/length(Species)
plot(x=1:15,y=rep(1,15),col=as.vector(Species),lwd=4,ylim=c(0,3),xlab="",ylab="",
yaxt = "n",xaxt="n")
```

```
par(new=TRUE)
plot(x=1:15,y=rep(2,15),col=Res,lwd=4,ylim=c(0,3),xlab="",ylab="",yaxt = "n",xaxt="n")
ind<-which(Res!=Species)
if(length(ind)>0) {
  for(j in 1:length(ind))
      lines(x=c(ind[j],ind[j]),y=c(1+0.05,2-0.05))
}
text(5,0.3,paste("KNN Misclassified:",length(ind)))
axis(2,at=2,labels="KNN",las=1)
text(10,2.5,paste("k: ",k))
```

PCAy *Revised PCA analysis*

Description

It's a revised PCA analysis.

Usage

```
PCAy(data)
```

Arguments

data	Numeric. Data matrix for revised PCA analysis.

Value

An object of class "PCAyd".

Author(s)

Yi Ya, Yarong Yang, and Yoram Rubin

References

Yarong Yang, Matt Over, and Yoram Rubin.(2012) Strategic Placement of Localization Devices (such as Pilot Points and Anchors) in Inverse Modeling Schemes. Water Resources Research, 48, W08519, doi:10.1029/2012WR011864.

Yarong Yang(Yi Ya).(2022) SKNN: A Super K-Nearest Neighbor Classification Algorithm. submitted to Journal of Statistical Software

Examples

```
Sepal.Length<-c(4.8, 5.1, 4.6, 5.3, 5.0, 5.7, 5.7, 6.2, 5.1, 5.7, 6.7, 6.3,
 6.5, 6.2, 5.9)
Sepal.Width<-c(3.0, 3.8, 3.2, 3.7, 3.3, 3.0, 2.9, 2.9, 2.5, 2.8, 3.0, 2.5,
 3.0, 3.4, 3.0)
Petal.Length<-c(1.4, 1.6, 1.4, 1.5, 1.4, 4.2, 4.2, 4.3, 3.0, 4.1, 5.2, 5.0,
 5.2, 5.4, 5.1)
Petal.Width<-c(0.3, 0.2, 0.2, 0.2, 0.2, 1.2, 1.3, 1.3, 1.1, 1.3, 2.3, 1.9,
 2.0, 2.3, 1.8)
dat<-cbind(Sepal.Length,Sepal.Width,Petal.Length,Petal.Width)
Res<-PCAy(dat)
```

PCAyd *Class to contain the results from revised PCA analysis.*

Description

The function PCAy returns object of class PCAyd.

Objects from the Class

new("PCAyd",Var=new("numeric"),PC=new("matrix"),Scores=new("matrix"),IScores=new("numeric"))

Slots

`Var`: An numeric vector giving the variance of each PC.

`PC`: A numeric matrix about the coefficients of each PC.

`Scores`: A numeric matrix showing the loading coefficiens of each PC.

`IScores`: A numeric vector with each element being the rowsum of Scores.

Author(s)

Yi Ya, Yarong Yang, and Yoram Rubin

References

Yarong Yang, Matt Over, and Yoram Rubin.(2012) Strategic Placement of Localization Devices (such as Pilot Points and Anchors) in Inverse Modeling Schemes. Water Resources Research, 48, W08519, doi:10.1029/2012WR011864.

Yarong Yang(Yi Ya).(2022) SKNN: A Super K-Nearest Neighbor Classification Algorithm. submitted to Journal of Statistical Software

Examples

```
showClass("PCAyd")
```

SKNN *Super K-Nearest Neighbor (SKNN) Classification*

Description

It's a Super K-Nearest Neighbor classification method with using kernel density to describe weight of the distance between a training observation and the testing sample.

Usage

```
SKNN(data, Class, k, test)
```

Arguments

`data`	Numeric. The training data set, matrix.
`Class`	Character. Class of the training observations, vector.
`k`	Integer. The number of K to be used.
`test`	Numeric. The sample to be classified.

Value

A character.

Author(s)

Yi Ya, Nader Ebrahimi, and Yoram Rubin

References

Yarong Yang(Yi Ya).(2022) SKNN: A Super K-Nearest Neighbor Classification Algorithm. submitted to Journal of Statistical Software

Yarong Yang, Matt Over, and Yoram Rubin.(2012) Strategic Placement of Localization Devices (such as Pilot Points and Anchors) in Inverse Modeling Schemes. Water Resources Research, 48, W08519, doi:10.1029/2012WR011864.

B.B.W. Silverman. Density Estimation for Statistics and Data Analysis. London: Chapman and Hall, 1986.

Examples

```
Sepal.Length<-c(4.8, 5.1, 4.6, 5.3, 5.0, 5.7, 5.7, 6.2, 5.1, 5.7, 6.7, 6.3, 6.5, 6.2, 5.9)
Sepal.Width<-c(3.0, 3.8, 3.2, 3.7, 3.3, 3.0, 2.9, 2.9, 2.5, 2.8, 3.0, 2.5, 3.0, 3.4, 3.0)
Petal.Length<-c(1.4, 1.6, 1.4, 1.5, 1.4, 4.2, 4.2, 4.3, 3.0, 4.1, 5.2, 5.0, 5.2, 5.4, 5.1)
Petal.Width<-c(0.3, 0.2, 0.2, 0.2, 0.2, 1.2, 1.3, 1.3, 1.1, 1.3, 2.3, 1.9, 2.0, 2.3, 1.8)
Species<-as.factor(c(rep("red",5),rep("blue",5),rep("green",5)))
iris<-cbind(Sepal.Length,Sepal.Width,Petal.Length,Petal.Width)
Res<-length(nrow(iris))
```

```
k<-10
for(i in 1:nrow(iris))
     Res[i]<-SKNN(data=iris,Class=as.vector(Species),k=k,test=iris[i,])
accuracy<-length(which(Res==Species))/length(Species)
plot(x=1:15,y=rep(1,15),col=as.vector(Species),lwd=4,ylim=c(0,3),xlab="",ylab="",
yaxt = "n",xaxt="n")
par(new=TRUE)
plot(x=1:15,y=rep(2,15),col=Res,lwd=4,ylim=c(0,3),xlab="",ylab="",yaxt = "n",xaxt="n")
ind<-which(Res!=Species)
if(length(ind)>0) {
  for(j in 1:length(ind))
      lines(x=c(ind[j],ind[j]),y=c(1+0.05,2-0.05))
}
text(5,0.3,paste("SKNN Misclassified:",length(ind)))
axis(2,at=2,labels="SKNN",las=1)
text(10,2.5,paste("k: ",k))
```

Index