

Package ‘ExNRuleEnsemble’

December 19, 2022

Type Package

Title A k Nearest Neighbour Ensemble Based on Extended Neighbourhood Rule

Version 0.1.1

Description The extended neighbourhood rule for the k nearest neighbour ensemble where the neighbours are determined in k steps. Starting from the first nearest observation of the test point, the algorithm identifies a single observation that is closest to the observation at the previous step. At each base learner in the ensemble, this search is extended to k steps on a random bootstrap sample with a random subset of features selected from the feature space. The final predicted class of the test point is determined by using a majority vote in the predicted classes given by all base models. Amjad Ali, Muhammad Hamraz, Naz Gul, Dost Muhammad Khan, Saeed Aldahmani, Zardad Khan (2022) <[doi:10.48550/arXiv.2205.15111](https://doi.org/10.48550/arXiv.2205.15111)>.

License GPL (>= 3)

Encoding UTF-8

Imports FNN

LazyData true

Depends R (>= 2.10)

NeedsCompilation no

Author Amjad Ali [aut, cre, cph],
Muhammad Hamraz [aut],
Saeed Aldahmani [aut],
Zardad Khan [aut]

Maintainer Amjad Ali <Amjad.ali@awkum.edu.pk>

Repository CRAN

Date/Publication 2022-12-19 15:00:11 UTC

R topics documented:

ExNRule	2
ILPD	3
Index	5

 ExNRule

The Extended Neighbourhood Rule for k Nearest Neighbour Ensemble

Description

The function `ExNRule()` gives predictions and class probabilities for binary classification data using the extended neighbourhood rule (ExNRule) for k nearest neighbour (kNN) ensemble.

Usage

```
ExNRule(xtrain, xtest, ytrain, k=3, r=500, p=round(sqrt(ncol(xtrain))))
```

Arguments

<code>xtrain</code>	The features space of the training dataset.
<code>xtest</code>	The feature space of the testing data.
<code>ytrain</code>	The response variable of the training dataset.
<code>k</code>	The nearest observation in the extended pattern.
<code>r</code>	The total number of base kNN models using ExNRule.
<code>p</code>	The number of features used for each base kNN model using ExNRule.

Value

<code>class</code>	Predicted classes for the test data.
<code>class.prob</code>	Predicted class probabilities for the test data.

Author(s)

Amjad Ali, Muhammad Hamraz, Saeed Aldahmani, Zardad Khan.

Maintainer: Amjad Ali <Amjad.ali@awkum.edu.pk>

References

Ali, Amjad, et al. "An Optimal k Nearest Neighbours Ensemble for Classification Based on Extended Neighbourhood Rule with Features subspace." arXiv preprint arXiv:2211.11278 (2022).

Examples

```
data(ILPD)

X <- ILPD[, -11]
Y <- as.factor(ILPD[, 11])

n <- nrow(ILPD)
train <- sample(1:n, 0.7*n, replace = FALSE)
```

```
xtrain <- X[train,]
xtest <- X[-train,]

ytrain <- Y[train]
ytest <- Y[-train]

mymod <- ExNRule(xtrain, xtest, ytrain, k=3, r=500, p=round(sqrt(ncol(xtrain))))
pred.class <- mymod$class
class.prob <- mymod$class.prob

Accuracy <- mean(pred.class==ytest)
ytestnum <- as.numeric(as.factor(ytest))-1
BrierScore <- mean((ytestnum-class.prob)^2)

cbind(Accuracy, BrierScore)
```

ILPD

Indian Liver Patient Dataset

Description

This data set contains 416 liver patient records and 167 non liver patient records. The data set was collected from north east of Andhra Pradesh, India. The class label divides the patients into 2 groups i.e. liver patient or not (1, 0). This data set contains 441 male patient records and 142 female patient records.

Usage

```
data("ILPD")
```

Format

A data frame with 582 observations on the following 11 variables.

v1 Age of the patient. Any patient whose age exceeded 89 is listed as being of age "90".

v2 Gender of the patients (male = 1, female = 0).

v3 Total Bilirubin.

v4 Direct Bilirubin.

v5 Alkphos Alkaline Phosphatase.

v6 Sgpt Alanine Aminotransferase.

v7 Sgot Aspartate Aminotransferase.

v8 Total Proteins.

v9 Albumin.

v10 A/G Ratio Albumin and Globulin Ratio.

Class Liver patient or not (1, 0).

Source

<https://www.openml.org/search?type=data&sort=runs&id=1480&status=active>

References

1. Bendi Venkata Ramana, Prof. M. S. Prasad Babu and Prof. N. B. Venkateswarlu, A Critical Comparative Study of Liver Patients from USA and INDIA: An Exploratory Analysis., International Journal of Computer Science Issues, ISSN:1694-0784, May 2012.
2. Bendi Venkata Ramana, Prof. M. S. Prasad Babu and Prof. N. B. Venkateswarlu, A Critical Study of Selected Classification Algorithms for Liver Disease Diagnosis, International Journal of Database Management Systems (IJDMS), Vol.3, No.2, ISSN : 0975-5705, PP 101-114, May 2011.

Examples

```
data(ILPD)
str(ILPD)
```

Index

- * **Bootstrapping**
 - ExNRule, 2
 - * **Classification**
 - ExNRule, 2
 - * **Ensemble**
 - ExNRule, 2
 - * **ExNRule**
 - ExNRule, 2
 - * **ILPD**
 - ILPD, 3
- ExNRule, 2
- ILPD, 3