Package 'BNN'

November 8, 2017

Type Package

Title Bayesian Neural Network for High-Dimensional Nonlinear Variable Selection

Version 1.0.0

Date 2017-11-08

Author Bochao Jia, Faming Liang

Maintainer Bochao Jia <jbc409@ufl.edu>

Depends R (>= 3.0.2)

Description This package is to perform Bayesian variable selection for high-dimensional nonlinear systems. The computation can be accelerated using multiple CPUs. The package can also be used to test nonlinearity for a general regression problem.

License GPL-2

LazyLoad true

NeedsCompilation yes

Repository CRAN

Date/Publication 2017-11-08 21:56:57 UTC

RoxygenNote 6.0.1

R topics documented:

BNN-package *Bayesian Neural Networks for High-Dimensional Nonlinear Variable Selection*

Description

This package is to perform Bayesian variable selection for high-dimensional nonlinear systems. The computation can be accelerated using multiple CPUs. The package can also be used to test nonlinearity for a general regression problem.

Details

Package:	BNN
Type:	Package
Version:	1.0.0
Date:	2017-11-08
License:	GPL-2

The Bayesian neural network used in the package is a one-hidden layer feedforward neural network with shortcut connections. The first module of the package is to calculate the prior probabilities assigned to the class of linear models (i.e., those networks with only shortcut connections) and the class of nonlinear models; and the second module is to perform nonlinear variable selection and calculate the posterior probabilities of the classes of linear and nonlinear models.

Author(s)

Bochao Jia, Faming Liang Maintainer: Bochao Jia<jbc409@ufl.edu>

References

Liang, F., Li, Q. and Zhou, L. (2017) Bayesian Neural Networks for Selection of Drug Sensitive Genes. Journal of the American Statistical Association, under revision.

Xue, J. and Liang, F. (2017). Robust model-free feature screening for ultrahigh dimensional data. Journal of Computational and Graphical Statistics, in press.

Examples

```
#library(BNN)
#BNNprior(50, 1, hid_num = 3,lambda=0.025,total_iteration = 1000000, popN = 20)
```

BNNprior *Prior Probability of Bayesian Neural Networks*

Description

Calculating the prior probability of linear and nonlinear classes of BNN models.

Usage

```
BNNprior(dimX, dimY, hid_num = 3,lambda=0.025,total_iteration = 1000000, popN = 20)
```

Arguments

dimX	Dimension of the input data.
dimY	The dimension of reponse data. It is restricted to 1 in the current version of the package.
hid_num	Number of hidden units. The default setting is 3.
lambda	The prior probability for each connection of the neural network being selected for the final model. The default setting is 0.025.
total_iteration	Number of total iterations, default of 1000,000.
popN	Number of Markov Chains, default of 20.

Value

prob	Prior probability assigned to the class of linear models.

Author(s)

Bochao Jia and Faming Liang

References

Liang, F., Li, Q. and Zhou, L. (2017) Bayesian Neural Networks for Selection of Drug Sensitive Genes. Journal of the American Statistical Association, under revision.

Examples

```
#library(BNN)
#BNNprior(50, 1, hid_num = 3,lambda=0.025,total_iteration = 1000000, popN = 20)
```

BNNsel *Bayesian Neural Network for Variable Selection*

Description

Perform variable selection and calculate posterior probabilities for the classes of linear and nonlinear of models.

Usage

```
BNNsel(X,Y,train_num, hid_num = 3, lambda=0.025,total_iteration = 1000000, popN = 20, nCPUs = 20)
```

Arguments

X	a *nxp* input data matrix.
Y	response vector.
train_num	Number of training samples. The default setting is 80% of the input samples.
hid_num	Number of hidden units. The default setting is 3.
lambda	The prior probability for each connection of the neural network being selected for the final model. The default setting is 0.025.
total_iteration	Number of iterations. The default setting is 1000,000.
popN	Number of Markov Chains in a parallel run. The default setting is 20.
nCPUs	Number of CPUs to be used in the simulation. The default setting is 20.

Value

A list of five elements:

net	Marginal inclusion probability of each connection of the neural network.
prob	Posterior probability of the class of linear models.
mar	Marginal inclusion probability of each input variable, which can be used for variable selection based on a multiple-hypothesis test or the median probability model criterion.
fit	Fitted value for the response vector of training data.
pred	Predicted value for the response vector of testing data.

Author(s)

Bochao Jia and Faming Liang

References

Liang, F., Li, Q. and Zhou, L. (2017) Bayesian Neural Networks for Selection of Drug Sensitive Genes. Journal of the American Statistical Association, under revision.

Examples

```
#library(BNN)
#data(Topotecan)
#X <- Topotecan$X
#Y <- Topotecan$Y
#BNNsel(X,Y, hid_num = 3,lambda=0.025,total_iteration = 1000000, popN = 20, nCPUs = 20)
```

Topotecan *Example dataset*

Description

A subset drug response data extracted from the cancer cell line encyclopedia (CCLE) database for the drug topotecan.

Usage

```
data(Topotecan)
```

Format

X a *nxp* data matrix; n=491, p=89

Y response vector.

References

Liang, F., Li, Q. and Zhou, L. (2017) Bayesian Neural Networks for Selection of Drug Sensitive Genes. Journal of the American Statistical Association, under revision.

Xue, J. and Liang, F. (2017). Robust model-free feature screening for ultrahigh dimensional data. Journal of Computational and Graphical Statistics, in press.

Index