test for MCMC

```
Warning: package 'bartMachine' was built under R version 4.3.3

Warning: package 'randomForest' was built under R version 4.3.3

Warning: package 'missForest' was built under R version 4.3.3

Warning: package 'dbarts' was built under R version 4.3.3

Warning: package 'BART' was built under R version 4.3.3

Warning: package 'bench' was built under R version 4.3.3

Warning: package 'ggplot2' was built under R version 4.3.3
```

```
data(benchmark_datasets)
```

```
set.seed(316)
cancer <- read.csv("C:/Users/pyk/Desktop/nus/RA/project/imodels-data-master/data_cleaned/brea
x <- boston[,-ncol(boston)]
y <- boston[,ncol(boston)]

data(benchmark_datasets)

train_indices <- createDataPartition(y, p = 0.8, list = FALSE)
train_x <- x[train_indices, ]
test_x <- x[-train_indices, ]
train_y <- y[train_indices]
test_y <- y[-train_indices]
```

```r
bart_model <- bartMachine(
  X = train_x,
  y = train_y,
  mh_prob_steps = c(0.5,0.5,0)
  )
```

```
bartMachine initializing with 50 trees...
bartMachine vars checked...
bartMachine java init...
bartMachine factors created...
bartMachine before preprocess...
bartMachine after preprocess... 14 total features...
bartMachine sigsq estimated...
bartMachine training data finalized...
Now building bartMachine for regression...Covariate importance prior ON.
evaluating in sample data...done
```

```r
pred <- predict(bart_model,test_x)
bart_model1 <- bartMachine(
  X = train_x,
  y = train_y,
  mh_prob_steps = c(0.25,0.25,0.5)
  )
```

```
bartMachine initializing with 50 trees...
bartMachine vars checked...
bartMachine java init...
bartMachine factors created...
bartMachine before preprocess...
bartMachine after preprocess... 14 total features...
bartMachine sigsq estimated...
bartMachine training data finalized...
Now building bartMachine for regression...Covariate importance prior ON.
evaluating in sample data...done
```

```r
pred1 <- predict(bart_model1,test_x)
bart_model2 <- bartMachine(train_x,train_y)
```

```
bartMachine initializing with 50 trees...
bartMachine vars checked...
```

```
bartMachine java init...
bartMachine factors created...
bartMachine before preprocess...
bartMachine after preprocess... 14 total features...
bartMachine sigsq estimated...
bartMachine training data finalized...
Now building bartMachine for regression...Covariate importance prior ON.
evaluating in sample data...done
```

```
pred2 <- predict(bart_model2,test_x)
```

```
mse <- mean((test_y - pred)^2)
mse1 <- mean((test_y - pred1)^2)
mse2 <- mean((test_y - pred2)^2)
print(mse)
```

```
[1] 13.10632
```

```
print(mse1)
```

```
[1] 13.97632
```

```
print(mse2)
```

```
[1] 15.14523
```

```
dbart <- bart(train_x,train_y,test_x,proposalprobs = c(0.25,0.25,0.4,0.1))
```

```
Running BART with numeric y

number of trees: 200
number of chains: 1, number of threads 1
tree thinning rate: 1
Prior:
    k prior fixed to 2.000000
    degrees of freedom in sigma prior: 3.000000
    quantile in sigma prior: 0.900000
    scale in sigma prior: 0.001983
```

```
    power and base for tree prior: 2.000000 0.950000
    use quantiles for rule cut points: false
    proposal probabilities: birth/death 0.50, swap 0.10, change 0.40; birth 0.50
data:
    number of training observations: 406
    number of test observations: 100
    number of explanatory variables: 13
    init sigma: 3.656196, curr sigma: 3.656196

Cutoff rules c in x<=c vs x>c
Number of cutoffs: (var: number of possible c):
(1: 100) (2: 100) (3: 100) (4: 100) (5: 100)
(6: 100) (7: 100) (8: 100) (9: 100) (10: 100)
(11: 100) (12: 100) (13: 100)
Running mcmc loop:
iteration: 100 (of 1000)
iteration: 200 (of 1000)
iteration: 300 (of 1000)
iteration: 400 (of 1000)
iteration: 500 (of 1000)
iteration: 600 (of 1000)
iteration: 700 (of 1000)
iteration: 800 (of 1000)
iteration: 900 (of 1000)
iteration: 1000 (of 1000)
total seconds in loop: 0.574540

Tree sizes, last iteration:
[1] 2 2 3 2 2 3 2 2 2 3 2 4 2 3 2 1 2 2
2 2 3 3 3 2 2 3 2 2 2 3 3 2 3 2 2 2 2 3
2 2 2 1 1 2 2 2 3 2 2 3 2 2 2 2 2 2 2 2
2 2 2 2 1 2 3 2 2 2 4 2 3 1 2 2 2 2 4 3
1 2 3 2 2 2 2 3 2 2 2 2 1 2 3 3 4 2 2 3
3 3 2 2 3 2 2 3 3 1 2 2 2 3 3 2 2 2 4 1
4 3 1 3 2 2 4 5 2 2 4 4 4 3 2 2 1 2 2 2
2 3 2 2 2 2 2 3 2 3 1 3 2 2 2 3 3 2 2 2
2 3 2 2 2 2 1 2 1 2 3 3 4 2 2 2 2 2 3 2
2 3 2 2 5 2 2 1 2 2 2 2 2 2 2 2 1 2 2 2
3 4

Variable Usage, last iteration (var:count):
(1: 23) (2: 12) (3: 16) (4: 15) (5: 26)
(6: 21) (7: 18) (8: 30) (9: 20) (10: 21)
```

```
(11: 21) (12: 18) (13: 17)
DONE BART
```

```r
pred <- colMeans(dbart$yhat.test)
mse_d <- mean((test_y - pred)^2)

dbart1 <- bart(train_x,train_y,test_x,proposalprobs = c(0.5,0.5,0,0))
```

```
Running BART with numeric y

number of trees: 200
number of chains: 1, number of threads 1
tree thinning rate: 1
Prior:
    k prior fixed to 2.000000
    degrees of freedom in sigma prior: 3.000000
    quantile in sigma prior: 0.900000
    scale in sigma prior: 0.001983
    power and base for tree prior: 2.000000 0.950000
    use quantiles for rule cut points: false
    proposal probabilities: birth/death 0.50, swap 0.10, change 0.40; birth 0.50
data:
    number of training observations: 406
    number of test observations: 100
    number of explanatory variables: 13
    init sigma: 3.656196, curr sigma: 3.656196

Cutoff rules c in x<=c vs x>c
Number of cutoffs: (var: number of possible c):
(1: 100) (2: 100) (3: 100) (4: 100) (5: 100)
(6: 100) (7: 100) (8: 100) (9: 100) (10: 100)
(11: 100) (12: 100) (13: 100)
Running mcmc loop:
iteration: 100 (of 1000)
iteration: 200 (of 1000)
iteration: 300 (of 1000)
iteration: 400 (of 1000)
iteration: 500 (of 1000)
iteration: 600 (of 1000)
iteration: 700 (of 1000)
iteration: 800 (of 1000)
```

```
iteration: 900 (of 1000)
iteration: 1000 (of 1000)
total seconds in loop: 0.586904

Tree sizes, last iteration:
[1] 2 2 2 2 2 2 3 2 2 2 2 4 2 2 2 2 2 2
3 2 2 3 1 3 2 2 2 2 1 3 3 2 2 2 4 2 4 2
3 2 2 2 2 3 2 2 3 2 3 3 2 3 3 2 2 2 2 2
2 2 2 2 2 3 3 2 3 2 2 3 3 2 2 2 2 1 2 1
2 2 2 2 3 2 2 2 2 2 2 3 2 3 3 2 3 1 3 1
1 2 4 2 2 2 4 3 2 2 2 2 2 2 2 2 2 3 2 2
2 3 3 2 3 2 4 3 1 2 2 2 3 2 1 2 2 2 2 4
2 2 3 3 2 2 3 2 3 2 2 1 3 3 2 2 2 3 3 3
3 2 4 3 2 2 2 2 2 2 2 2 2 3 2 4 2 5 2 2
2 2 2 3 3 2 3 3 3 2 2 3 3 2 2 2 2 2 3 2
5 2

Variable Usage, last iteration (var:count):
(1: 31) (2: 23) (3: 20) (4: 22) (5: 18)
(6: 18) (7: 29) (8: 21) (9: 27) (10: 14)
(11: 14) (12: 13) (13: 15)
DONE BART
```

```r
pred1 <- colMeans(dbart1$yhat.test)
mse_d1 <- mean((test_y - pred1)^2)

dbart2 <- bart(train_x,train_y,test_x)
```

```
Running BART with numeric y

number of trees: 200
number of chains: 1, number of threads 1
tree thinning rate: 1
Prior:
    k prior fixed to 2.000000
    degrees of freedom in sigma prior: 3.000000
    quantile in sigma prior: 0.900000
    scale in sigma prior: 0.001983
    power and base for tree prior: 2.000000 0.950000
    use quantiles for rule cut points: false
    proposal probabilities: birth/death 0.50, swap 0.10, change 0.40; birth 0.50
```

```
data:
    number of training observations: 406
    number of test observations: 100
    number of explanatory variables: 13
    init sigma: 3.656196, curr sigma: 3.656196

Cutoff rules c in x<=c vs x>c
Number of cutoffs: (var: number of possible c):
(1: 100) (2: 100) (3: 100) (4: 100) (5: 100)
(6: 100) (7: 100) (8: 100) (9: 100) (10: 100)
(11: 100) (12: 100) (13: 100)
Running mcmc loop:
iteration: 100 (of 1000)
iteration: 200 (of 1000)
iteration: 300 (of 1000)
iteration: 400 (of 1000)
iteration: 500 (of 1000)
iteration: 600 (of 1000)
iteration: 700 (of 1000)
iteration: 800 (of 1000)
iteration: 900 (of 1000)
iteration: 1000 (of 1000)
total seconds in loop: 0.575754

Tree sizes, last iteration:
[1] 2 2 2 4 4 2 6 2 2 2 3 1 2 3 2 2 1 3
2 2 2 3 2 3 1 2 2 2 2 3 4 2 2 2 2 4 2 3
1 2 4 2 2 2 4 3 1 2 2 3 2 3 3 2 2 3 2 3
2 2 2 1 2 3 2 2 3 2 2 4 5 2 2 2 2 2 2 3
3 2 2 1 2 2 1 3 2 3 2 2 4 2 2 2 2 2 2 3
4 5 2 4 2 3 3 2 2 2 4 3 3 2 2 2 4 3 2 2
2 2 3 2 2 2 2 2 2 2 3 2 2 3 1 2 3 2 3 3
2 3 3 2 2 2 4 3 2 4 3 3 2 2 2 2 3 2 2 3
2 2 2 3 2 2 2 3 2 3 3 2 2 2 2 4 3 2 2 2
3 4 3 2 2 2 4 2 3 3 2 2 3 2 4 2 3 2 2 2
1 2

Variable Usage, last iteration (var:count):
(1: 30) (2: 26) (3: 20) (4: 22) (5: 26)
(6: 20) (7: 32) (8: 17) (9: 15) (10: 18)
(11: 20) (12: 17) (13: 22)
DONE BART
```

```r
pred2 <- colMeans(dbart2$yhat.test)
mse_d2 <- mean((test_y - pred2)^2)

print(mse_d)
```

```
[1] 14.1365
```

```r
print(mse_d1)
```

```
[1] 14.70578
```

```r
print(mse_d2)
```

```
[1] 14.21109
```