Simulated examples for the survival ensemble methods

Select data type

W500

Time point for event prediction:

5

Random seed for calibration and validation

42

K_Outer loop CV (for validation)

3

K_ inner CV folds (model tuning)

3

Simulated data: random seed (generation):

4242

Sample size:

150

Observation time

5

Expected event prevalence by study end

0.5

Expected drop out rate

0.3

Custom data: path to data file

~/Desktop/Study_KCL/PhD P

Predictors to use in the model

"baseline_age_", "genderdun

Time variable name

time

Event indicator variable name

event

Sample statistics CoxPH SRF Ens1: CoxPH->SRF Ens2: CoxPH in clusters

Ens3: extended CoxPH Summary Conclusions

Internally cross-validated results:

Show 10 **entries** **Search:**

	AUCROC	BS	BS_scaled	C_score	Calib_slope	Calib_alpha	T
test	0.8271	0.217	0.232	0.772	0.9054	0.1111	5
train	0.8527	0.1781	0.3753	0.7909	1.0573	0.1	5

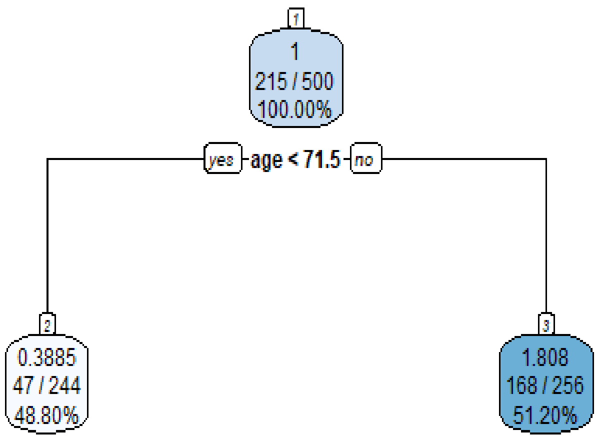
Showing 1 to 2 of 2 entries Previous 1 Next

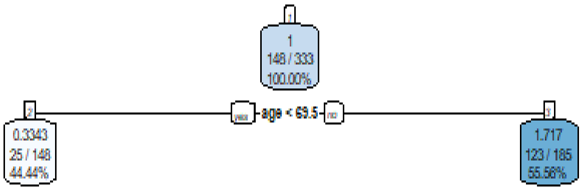
Internally cross-validated Test results for each CV fold:

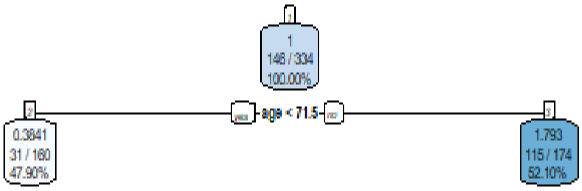
Show 10 **entries** **Search:**

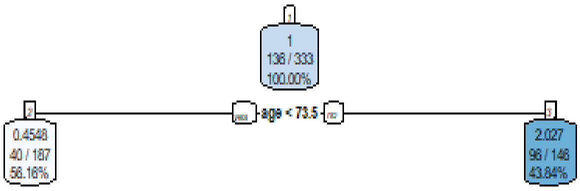
	AUCROC	BS	BS_scaled	C_score	Calib_slope	Calib_alpha	T
test.1	0.8489	0.271	0.0783	0.8009	1.0125	-0.0289	5
test.2	0.8103	0.1946	0.3014	0.7768	0.8511	0.1708	5
test.3	0.8222	0.1853	0.3163	0.7382	0.8526	0.1914	5

Showing 1 to 3 of 3 entries Previous 1 Next









Show 25 **entries** **Search:**

	coef	exp(coef)	se(coef)	z	Pr(>|z|)
age	0.0337	1.0342	0.0103	3.2654	0.0011
gender	-0.3189	0.727	0.1455	-2.1924	0.0283
hr	0.0105	1.0105	0.0031	3.334	0.0009
sysbp	0.001	1.001	0.0029	0.3546	0.7229
diasbp	-0.0126	0.9874	0.0049	-2.5611	0.0104
bmi	-0.0499	0.9513	0.0169	-2.9587	0.0031
cvd	-0.0583	0.9433	0.1828	-0.319	0.7497
afb	0.0351	1.0357	0.182	0.193	0.847
sho	1.2344	3.4363	0.2876	4.2924	0
chf	0.7446	2.1057	0.1554	4.792	0
av3	0.2402	1.2716	0.4292	0.5598	0.5756
miord	0.0741	1.0769	0.1517	0.4887	0.625
mitype	-0.1726	0.8414	0.1944	-0.8879	0.3746
los	-0.0078	0.9922	0.0164	-0.4759	0.6341
y1997	-0.4716	0.624	0.2012	-2.3439	0.0191

	coef	exp(coef)	se(coef)	z	Pr(>|z|)
y1999	-0.3397	0.712	0.1862	-1.8244	0.0681
cluster_tree1.808196	0.4818	1.619	0.2806	1.717	0.086

Showing 1 to 17 of 17 entries

Previous 1 Next

```
$test
  T    AUCROC        BS  BS_scaled   C_score Calib_slope
1 5 0.8489330 0.2710019 0.07828996 0.8008552   1.0124953
2 5 0.8102685 0.1946135 0.30143785 0.7768316   0.8510680
3 5 0.8221614 0.1853237 0.31629568 0.7382306   0.8525944
  Calib_alpha test cv_n
1 -0.02889638    1    1
2  0.17078132    1    2
3  0.19137805    1    3

$train
  T    AUCROC        BS BS_scaled   C_score Calib_slope
1 5 0.8294790 0.1960707 0.3245530 0.7749963   0.9679331
2 5 0.8790076 0.1640144 0.4197382 0.7900887   1.1769625
3 5 0.8495388 0.1743388 0.3815738 0.8075582   1.0268940
  Calib_alpha test cv_n
1  0.09576816    0    1
2  0.10958584    0    2
3  0.09450955    0    3

$testaverage
          T      AUCROC          BS   BS_scaled
  5.0000000   0.8271210   0.2169797   0.2320078
    C_score Calib_slope Calib_alpha        test
  0.7719725   0.9053859   0.1110877   1.0000000

$trainaverage
          T      AUCROC          BS   BS_scaled
 5.00000000  0.85267511  0.17814132  0.37528838
    C_score Calib_slope Calib_alpha        test
 0.79088108  1.05726318  0.09995452  0.00000000

$model_list
$model_list[[1]]
$model_list[[1]]$treemodel
n= 333

node), split, n, deviance, yval
      * denotes terminal node

1) root 333 465.2627 1.0000000
  2) age< 69.5 148 128.6917 0.3342506 *
  3) age>=69.5 185 258.2622 1.7171170 *

$model_list[[1]]$modcoxmodel
Call:
coxph(formula = as.formula(paste("Surv(df_train$time, df_train$event) ~",
    paste(predict.factors, collapse = "+"))), data = df_train,
    x = TRUE)

                         coef exp(coef)  se(coef)
age                  0.021342  1.021571  0.012863
gender              -0.298280  0.742094  0.177958
hr                   0.010339  1.010393  0.003887
sysbp                0.004023  1.004032  0.003693
```

```
diasbp               -0.015144  0.984970  0.006540
bmi                  -0.053344  0.948054  0.019792
cvd                  -0.101787  0.903222  0.222122
afb                   0.003267  1.003273  0.224159
sho                   1.029878  2.800724  0.374469
chf                   0.567875  1.764513  0.190091
av3                   0.399542  1.491142  0.498889
miord                -0.030185  0.970266  0.192159
mitype               -0.042932  0.957977  0.234652
los                  -0.032421  0.968099  0.024817
y1997                -0.244791  0.782868  0.245627
y1999                -0.253538  0.776050  0.233575
cluster_tree1.717117  0.908418  2.480395  0.374473
                          z       p
age                   1.659 0.09708
gender               -1.676 0.09371
hr                    2.660 0.00781
sysbp                 1.090 0.27592
diasbp               -2.315 0.02059
bmi                  -2.695 0.00704
cvd                  -0.458 0.64677
afb                   0.015 0.98837
sho                   2.750 0.00596
chf                   2.987 0.00281
av3                   0.801 0.42321
miord                -0.157 0.87518
mitype               -0.183 0.85483
los                  -1.306 0.19142
y1997                -0.997 0.31896
y1999                -1.085 0.27771
cluster_tree1.717117  2.426 0.01527

Likelihood ratio test=141.5  on 17 df, p=< 2.2e-16
n= 333, number of events= 148

$model_list[[1]]$clusters
[1] 1.717117 0.334251


$model_list[[2]]
$model_list[[2]]$treemodel
n= 334

node), split, n, deviance, yval
      * denotes terminal node

1) root 334 468.1744 1.0000000
  2) age< 71.5 160 151.2419 0.3840971 *
  3) age>=71.5 174 241.5789 1.7932300 *

$model_list[[2]]$modcoxmodel
Call:
coxph(formula = as.formula(paste("Surv(df_train$time, df_train$event) ~",
    paste(predict.factors, collapse = "+"))), data = df_train,
    x = TRUE)
```

```
                        coef exp(coef)  se(coef)      z
age                 0.025588  1.025918  0.012198  2.098
gender             -0.502935  0.604753  0.185469 -2.712
hr                  0.009955  1.010005  0.004005  2.486
sysbp              -0.001318  0.998683  0.003639 -0.362
diasbp             -0.012320  0.987755  0.006039 -2.040
bmi                -0.059015  0.942692  0.022278 -2.649
cvd                 0.007037  1.007062  0.219557  0.032
afb                -0.099761  0.905053  0.219254 -0.455
sho                 1.514605  4.547625  0.346503  4.371
chf                 0.935761  2.549154  0.204715  4.571
av3                 0.125707  1.133950  0.617303  0.204
miord              -0.013661  0.986432  0.187640 -0.073
mitype             -0.220695  0.801961  0.240199 -0.919
los                -0.002285  0.997717  0.018180 -0.126
y1997              -0.686179  0.503496  0.249625 -2.749
y1999              -0.393274  0.674844  0.227574 -1.728
cluster_tree1.79323  0.658206  1.931325  0.346798  1.898
                           p
age                  0.03592
gender               0.00669
hr                   0.01293
sysbp                0.71721
diasbp               0.04133
bmi                  0.00807
cvd                  0.97443
afb                  0.64911
sho                 1.24e-05
chf                 4.85e-06
av3                  0.83864
miord                0.94196
mitype               0.35820
los                  0.89997
y1997                0.00598
y1999                0.08397
cluster_tree1.79323  0.05770

Likelihood ratio test=176.3  on 17 df, p=< 2.2e-16
n= 334, number of events= 146

$model_list[[2]]$clusters
[1] 1.793230 0.384097


$model_list[[3]]
$model_list[[3]]$treemodel
n= 333

node), split, n, deviance, yval
      * denotes terminal node

1) root 333 454.9301 1.0000000
  2) age< 73.5 187 174.1279 0.4548238 *
  3) age>=73.5 146 207.2122 2.0269480 *

$model_list[[3]]$modcoxmodel
```

```
Call:
coxph(formula = as.formula(paste("Surv(df_train$time, df_train$event) ~",
    paste(predict.factors, collapse = "+"))), data = df_train,
    x = TRUE)

                          coef exp(coef)  se(coef)
age                   0.035726  1.036372  0.015282
gender               -0.217038  0.804899  0.183138
hr                    0.011583  1.011650  0.003892
sysbp                 0.001939  1.001941  0.003625
diasbp               -0.014187  0.985913  0.005981
bmi                  -0.038478  0.962253  0.020715
cvd                  -0.159733  0.852371  0.238623
afb                   0.274991  1.316518  0.235283
sho                   1.490594  4.439730  0.397227
chf                   0.834878  2.304533  0.188010
av3                   0.031270  1.031764  0.558029
miord                 0.193020  1.212908  0.185828
mitype               -0.412966  0.661685  0.265493
los                  -0.001056  0.998944  0.020345
y1997                -0.493121  0.610718  0.251002
y1999                -0.430172  0.650398  0.233070
cluster_tree2.026948  0.497970  1.645378  0.340480
                          z        p
age                   2.338 0.019396
gender               -1.185 0.235975
hr                    2.976 0.002917
sysbp                 0.535 0.592793
diasbp               -2.372 0.017702
bmi                  -1.858 0.063239
cvd                  -0.669 0.503244
afb                   1.169 0.242498
sho                   3.752 0.000175
chf                   4.441 8.97e-06
av3                   0.056 0.955312
miord                 1.039 0.298943
mitype               -1.555 0.119835
los                  -0.052 0.958603
y1997                -1.965 0.049459
y1999                -1.846 0.064940
cluster_tree2.026948  1.463 0.143589

Likelihood ratio test=169  on 17 df, p=< 2.2e-16
n= 333, number of events= 136

$model_list[[3]]$clusters
[1] 0.454824 2.026948



$time
Time difference of 1.336202 secs
```