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	0.8139	0.2351	0.1682	0.7591	0.5093	0.0974	5
train	0.878	0.1651	0.4211	0.8084	0.9769	0.0848	5

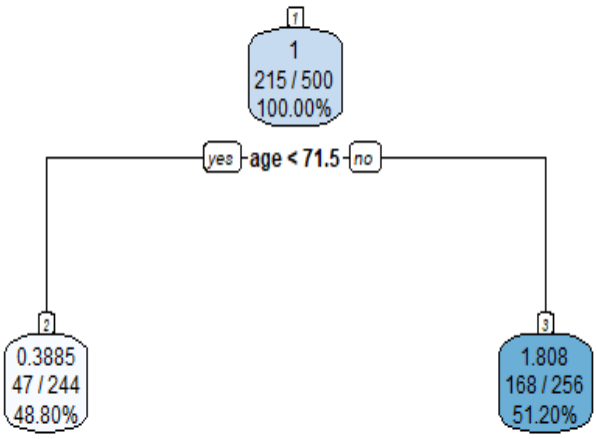
Showing 1 to 2 of 2 entries Previous 1 Next

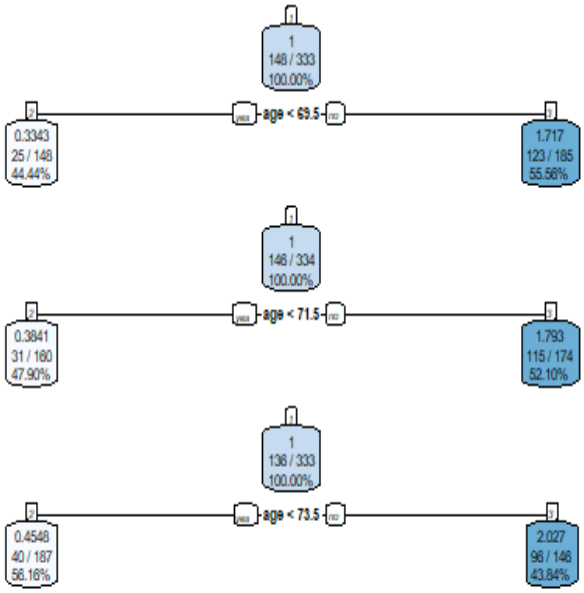
Internally cross-validated Test results for each CV fold:

Show 10 entries Search:

	AUCROC	BS	BS_scaled	C_score	Calib_slope	Calib_alpha	T
test.1	0.8375	0.2913	0.0091	0.7673	0.3728	-0.0296	5
test.2	0.7834	0.224	0.1961	0.7693	0.389	0.1464	5
test.3	0.8209	0.1899	0.2995	0.7407	0.766	0.1755	5

Showing 1 to 3 of 3 entries Previous 1 Next




1
148 / 333
100.00%
age < 69.5
0.3343
25 / 148
44.44%
1.717
123 / 185
55.56%
1
146 / 334
100.00%
age < 71.5
0.3841
31 / 160
47.90%
1.793
115 / 174
52.10%
1
136 / 333
100.00%
age < 73.5
0.4548
40 / 187
56.16%
2.027
96 / 146
43.84%

```
[[1]]
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        p
age     0.028928  1.029350  0.012415  2.330 0.019806
gender -0.457675  0.632753  0.163197 -2.804 0.005040
hr      0.005741  1.005758  0.003591  1.599 0.109874
sysbp  -0.001188  0.998813  0.003308 -0.359 0.719496
diasbp -0.006670  0.993352  0.005810 -1.148 0.250952
bmi    -0.068455  0.933835  0.020126 -3.401 0.000670
cvd    -0.216775  0.805111  0.208808 -1.038 0.299196
afb    -0.045571  0.955451  0.203171 -0.224 0.822523
sho     1.563340  4.774741  0.332497  4.702 2.58e-06
chf     0.640928  1.898242  0.174975  3.663 0.000249
av3     0.362507  1.436928  0.496151  0.731 0.464999
miord  -0.177376  0.837465  0.171818 -1.032 0.301907
mitype -0.381762  0.682658  0.241241 -1.582 0.113538
los    -0.014750  0.985358  0.020548 -0.718 0.472850
y1997  -0.546643  0.578890  0.231503 -2.361 0.018212
y1999  -0.343942  0.708970  0.210792 -1.632 0.102749

Likelihood ratio test=88.62  on 16 df, p=4.495e-12
n= 256, number of events= 168

[[2]]
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        p
age     0.023175  1.023445  0.020769  1.116 0.264506
gender  0.408663  1.504804  0.365636  1.118 0.263706
hr      0.027240  1.027614  0.008104  3.361 0.000776
sysbp   0.001163  1.001164  0.006431  0.181 0.856441
diasbp -0.028494  0.971908  0.009184 -3.103 0.001918
bmi    -0.013318  0.986770  0.031095 -0.428 0.668434
cvd     0.241430  1.273069  0.434382  0.556 0.578347
afb    -0.202830  0.816417  0.454511 -0.446 0.655409
sho     1.580730  4.858501  1.021891  1.547 0.121895
chf     1.044437  2.841797  0.389864  2.679 0.007385
av3    -0.487595  0.614102  1.319368 -0.370 0.711705
miord   0.860084  2.363360  0.376648  2.284 0.022400
mitype  0.437793  1.549284  0.385540  1.136 0.256152
los    -0.022224  0.978022  0.034116 -0.651 0.514784
y1997  -0.323355  0.723717  0.446941 -0.723 0.469381
y1999  -0.298567  0.741880  0.444337 -0.672 0.501623

Likelihood ratio test=67.38  on 16 df, p=2.854e-08
n= 244, number of events= 47
```

```
$test
  T    AUCROC        BS   BS_scaled   C_score
1 5 0.8374792 0.2913315 0.009146407 0.7672572
2 5 0.7834346 0.2239705 0.196061484 0.7693175
3 5 0.8209124 0.1898875 0.299458805 0.7407323
  Calib_slope Calib_alpha test cv_n
1   0.3728322 -0.02961608    1    1
2   0.3889984  0.14640679    1    2
3   0.7660380  0.17551161    1    3

$train
  T    AUCROC        BS BS_scaled   C_score Calib_slope
1 5 0.8587692 0.1825040 0.3712892 0.8028312   0.9789669
2 5 0.9020335 0.1481189 0.4759745 0.8109569   0.9527392
3 5 0.8730626 0.1646268 0.4160248 0.8115070   0.9990690
  Calib_alpha test cv_n
1  0.07614262    0    1
2  0.08831713    0    2
3  0.08999548    0    3

$testaverage
          T      AUCROC          BS   BS_scaled
 5.00000000  0.81394206  0.23506316  0.16822223
    C_score Calib_slope Calib_alpha        test
 0.75910232  0.50928955  0.09743411  1.00000000

$trainaverage
          T      AUCROC          BS   BS_scaled
 5.00000000  0.87795509  0.16508326  0.42109616
    C_score Calib_slope Calib_alpha        test
 0.80843168  0.97692506  0.08481841  0.00000000

$model_list
$model_list[[1]]
$model_list[[1]]$vimp10
          age           chf           los            hr
 6.724164e-02  1.020074e-02  5.888476e-03  4.639405e-03
          bmi        gender         y1999        mitype
 2.810409e-03  5.650558e-04  2.825279e-04  0.000000e+00
        miord           cvd        diasbp           av3
-8.921933e-05 -3.122677e-04 -3.271375e-04 -4.312268e-04
        y1997           afb         sysbp
-5.650558e-04 -7.286245e-04 -1.784387e-03

$model_list[[1]]$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 *

$model_list[[1]]$coxmodels
```

```
$model_list[[1]]$coxmodels[[1]]
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       p
age     0.019100  1.019283  0.015232  1.254 0.20988
gender -0.428408  0.651545  0.189461 -2.261 0.02375
hr      0.003987  1.003995  0.004338  0.919 0.35810
sysbp  -0.001136  0.998864  0.004258 -0.267 0.78959
diasbp -0.006559  0.993463  0.007795 -0.841 0.40014
bmi    -0.066325  0.935827  0.022436 -2.956 0.00311
cvd    -0.111167  0.894789  0.250479 -0.444 0.65717
afb    -0.108625  0.897067  0.241406 -0.450 0.65273
sho     1.382773  3.985938  0.416114  3.323 0.00089
chf     0.454712  1.575720  0.209216  2.173 0.02975
av3     0.231414  1.260382  0.518270  0.447 0.65523
miord  -0.266290  0.766217  0.213427 -1.248 0.21214
mitype -0.350075  0.704635  0.288441 -1.214 0.22487
los    -0.038491  0.962240  0.030170 -1.276 0.20202
y1997  -0.508341  0.601493  0.266103 -1.910 0.05609
y1999  -0.504708  0.603682  0.252673 -1.997 0.04577

Likelihood ratio test=53.55  on 16 df, p=6.142e-06
n= 185, number of events= 123

$model_list[[1]]$coxmodels[[2]]
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      p
age     3.073e-03  1.003e+00  2.705e-02  0.114 0.9096
gender  5.453e-01  1.725e+00  5.155e-01  1.058 0.2901
hr      5.806e-02  1.060e+00  1.561e-02  3.719 0.0002
sysbp   1.799e-02  1.018e+00  9.702e-03  1.855 0.0637
diasbp -3.005e-02  9.704e-01  1.301e-02 -2.310 0.0209
bmi    -6.798e-02  9.343e-01  4.784e-02 -1.421 0.1554
cvd     5.115e-02  1.052e+00  5.377e-01  0.095 0.9242
afb    -7.029e-01  4.952e-01  7.776e-01 -0.904 0.3661
sho    -1.497e+01  3.153e-07  7.537e+03 -0.002 0.9984
chf     9.690e-01  2.635e+00  5.656e-01  1.713 0.0867
av3    -1.540e+01  2.060e-07  1.724e+04 -0.001 0.9993
miord   1.152e+00  3.163e+00  5.508e-01  2.091 0.0366
mitype  6.295e-01  1.877e+00  5.723e-01  1.100 0.2714
los    -9.467e-02  9.097e-01  6.870e-02 -1.378 0.1682
y1997   7.556e-01  2.129e+00  8.875e-01  0.851 0.3946
y1999   7.383e-01  2.092e+00  8.831e-01  0.836 0.4032

Likelihood ratio test=45.69  on 16 df, p=0.0001087
n= 148, number of events= 25


$model_list[[1]]$clusters
```

```
[1] 1.717117 0.334251


$model_list[[2]]
$model_list[[2]]$vimp10
          age           chf           sho           bmi
 3.591597e-02  1.862364e-02  9.553208e-03  7.227083e-03
           hr         y1999           afb         sysbp
 2.823536e-03  1.609269e-03  1.448342e-03  1.170378e-03
        y1997         miord        diasbp           cvd
 9.509319e-04  6.290780e-04  5.266700e-04  2.925944e-05
       gender           los           av3
 0.000000e+00 -4.388917e-05 -2.925944e-04

$model_list[[2]]$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 *

$model_list[[2]]$coxmodels
$model_list[[2]]$coxmodels[[1]]
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        p
age     0.019451  1.019641  0.014655  1.327 0.184428
gender -0.551852  0.575882  0.205310 -2.688 0.007190
hr      0.006641  1.006663  0.004513  1.472 0.141150
sysbp  -0.001663  0.998338  0.004125 -0.403 0.686794
diasbp -0.007878  0.992153  0.007261 -1.085 0.277995
bmi    -0.074172  0.928512  0.026634 -2.785 0.005354
cvd    -0.089418  0.914463  0.251176 -0.356 0.721843
afb    -0.262738  0.768943  0.251763 -1.044 0.296675
sho     1.670552  5.315102  0.393463  4.246 2.18e-05
chf     0.865101  2.375245  0.233317  3.708 0.000209
av3     1.279578  3.595123  0.683815  1.871 0.061312
miord  -0.243116  0.784180  0.211536 -1.149 0.250436
mitype -0.638307  0.528186  0.314379 -2.030 0.042318
los    -0.006952  0.993072  0.022613 -0.307 0.758515
y1997  -0.580034  0.559880  0.287382 -2.018 0.043556
y1999  -0.229472  0.794953  0.258250 -0.889 0.374237

Likelihood ratio test=73.45  on 16 df, p=2.463e-09
n= 174, number of events= 115

$model_list[[2]]$coxmodels[[2]]
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        p
age      0.037911   1.038639   0.030876  1.228 0.219506
gender   0.460461   1.584804   0.519774  0.886 0.375679
hr       0.023285   1.023558   0.010932  2.130 0.033176
sysbp   -0.001590   0.998411   0.008387 -0.190 0.849606
diasbp  -0.036256   0.964394   0.012897 -2.811 0.004938
bmi     -0.008697   0.991341   0.043307 -0.201 0.840845
cvd      0.198426   1.219481   0.620040  0.320 0.748953
afb     -0.371857   0.689453   0.573676 -0.648 0.516855
sho      6.082124 437.958599   1.224272  4.968 6.77e-07
chf      1.592411   4.915584   0.536183  2.970 0.002979
av3     -6.046317   0.002367   1.820690 -3.321 0.000897
miord    0.716404   2.047059   0.523015  1.370 0.170762
mitype   0.510005   1.665299   0.516432  0.988 0.323371
los      0.004319   1.004328   0.036596  0.118 0.906051
y1997   -0.466781   0.627017   0.578149 -0.807 0.419452
y1999   -0.374509   0.687627   0.582743 -0.643 0.520441

Likelihood ratio test=61.27  on 16 df, p=3.199e-07
n= 160, number of events= 31


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


$model_list[[3]]
$model_list[[3]]$vimp10
          age           chf            hr           bmi
 0.0432633070  0.0195601281  0.0043894776  0.0042046575
          sho         miord           los        mitype
 0.0011551257  0.0009703056  0.0006314687 -0.0002156235
          av3           afb         y1997        gender
-0.0004312469 -0.0005852637 -0.0005852637 -0.0006006654
        y1999           cvd         sysbp
-0.0006160670 -0.0006776737 -0.0007084771

$model_list[[3]]$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 *

$model_list[[3]]$coxmodels
$model_list[[3]]$coxmodels[[1]]
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       p
age     0.0250934  1.0254109  0.0211165  1.188 0.23470
gender  0.2901066  1.3365700  0.3999416  0.725 0.46822
hr      0.0145886  1.0146955  0.0068322  2.135 0.03274
sysbp  -0.0028485  0.9971555  0.0088030 -0.324 0.74625
diasbp -0.0238292  0.9764525  0.0134098 -1.777 0.07557
bmi     0.0125542  1.0126333  0.0380716  0.330 0.74159
cvd    -0.1113814  0.8945974  0.4723621 -0.236 0.81359
afb     0.1131474  1.1197969  0.5083636  0.223 0.82387
sho     0.3993737  1.4908907  1.0494594  0.381 0.70354
chf     1.2016030  3.3254433  0.4147898  2.897 0.00377
av3     0.6739841  1.9620387  1.0447573  0.645 0.51886
miord   0.7631286  2.1449765  0.4097126  1.863 0.06252
mitype -0.0007882  0.9992121  0.4667957 -0.002 0.99865
los     0.0072370  1.0072632  0.0357690  0.202 0.83966
y1997  -1.1391495  0.3200911  0.4944073 -2.304 0.02122
y1999  -1.1028080  0.3319377  0.4842273 -2.277 0.02276

Likelihood ratio test=56.96  on 16 df, p=1.686e-06
n= 187, number of events= 40

$model_list[[3]]$coxmodels[[2]]
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        p
age     0.039913  1.040720  0.023030  1.733  0.08308
gender -0.455704  0.634002  0.228382 -1.995  0.04600
hr      0.008159  1.008192  0.004865  1.677  0.09354
sysbp   0.001582  1.001584  0.004132  0.383  0.70172
diasbp -0.010274  0.989779  0.006907 -1.487  0.13692
bmi    -0.067032  0.935165  0.026399 -2.539  0.01111
cvd    -0.462528  0.629690  0.286320 -1.615  0.10622
afb     0.292751  1.340109  0.278809  1.050  0.29372
sho     2.353602 10.523412  0.497257  4.733 2.21e-06
chf     0.692808  1.999322  0.222847  3.109  0.00188
av3    -0.160939  0.851344  0.784570 -0.205  0.83747
miord  -0.092528  0.911624  0.224916 -0.411  0.68079
mitype -0.558160  0.572261  0.340284 -1.640  0.10095
los    -0.017612  0.982542  0.025792 -0.683  0.49470
y1997  -0.366634  0.693064  0.302920 -1.210  0.22615
y1999  -0.464261  0.628599  0.289418 -1.604  0.10869

Likelihood ratio test=58.93  on 16 df, p=7.909e-07
n= 146, number of events= 96


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

```
$time
Time difference of 15.38528 secs
```