Continuous vs. Binary over varying training size

Ivana Malenica

Summer, 2019

Comparison of continuous and binary outcome prediction based on the subset of data

The below results summarize a test run comparing the Combined Online SL and Regular SL at various time points, for both the continuous and binary outcome.

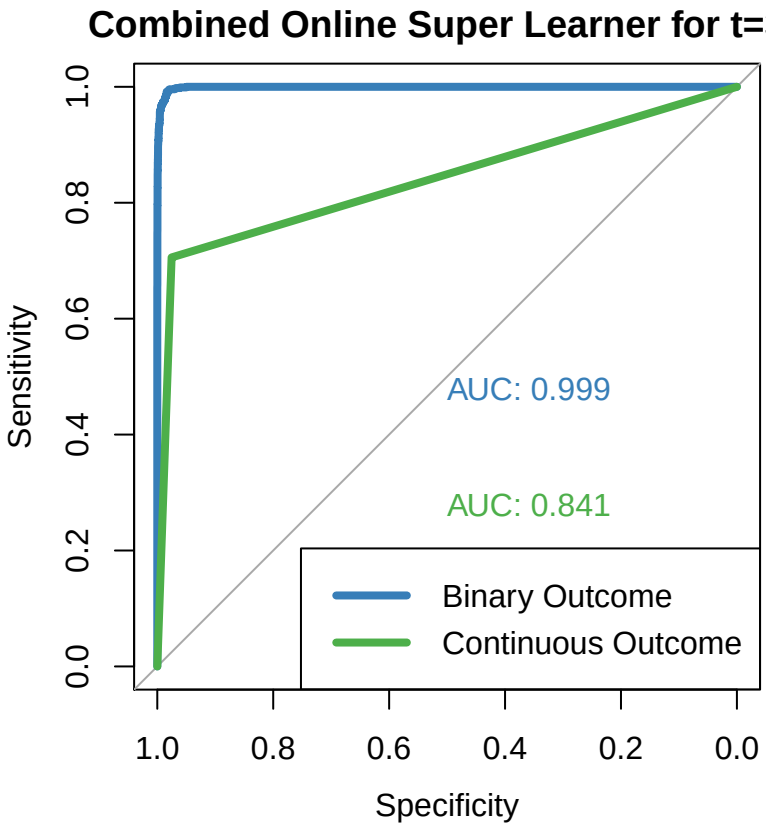
1. Combined Online SL: consists of both global and individualized learners, as well as various variations of the global learners that differentially adjust for baseline covariates.
2. Regular SL: just global SL.

The prediction is based on the following covariates:

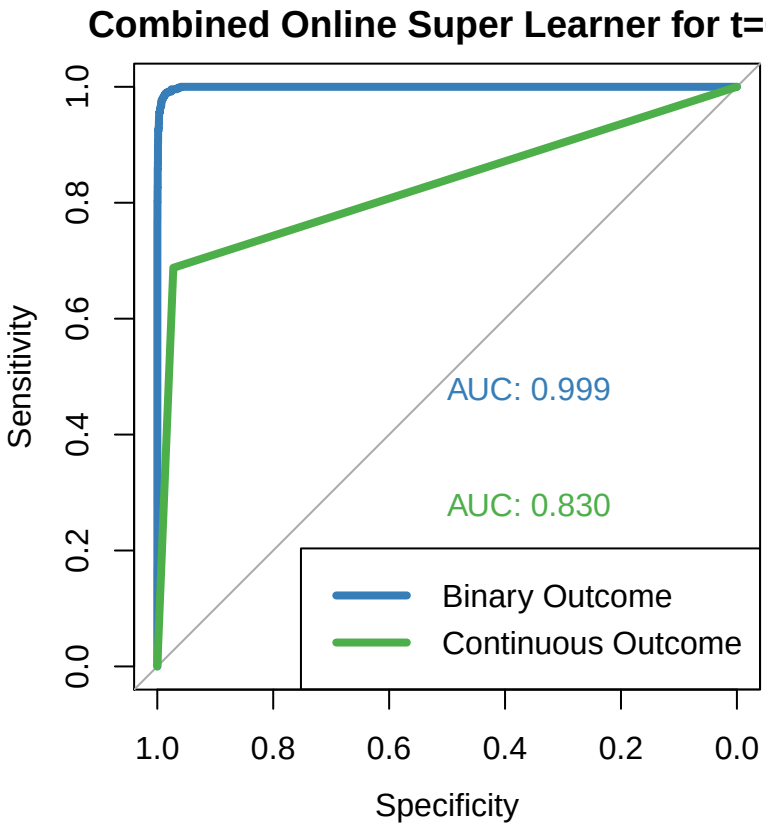
1. Baseline and Time-varying: "gender", "age", "care_unit", "admission_type_descr", "sapsi_first", "sofa_first", "amine","sedation","ventilation","spo2","abpsys", "abpdias"
2. Baseline: "amine","sedation","ventilation","spo2","abpsys","abpdias"

Initial training size included $n = 400$ samples, trained at $t = \{30, 60, 90, 120, 150, 180, 240\}$ minutes. For each t, we follow a gap of 30 minutes and predict the event for the following 20 minutes after the gap.

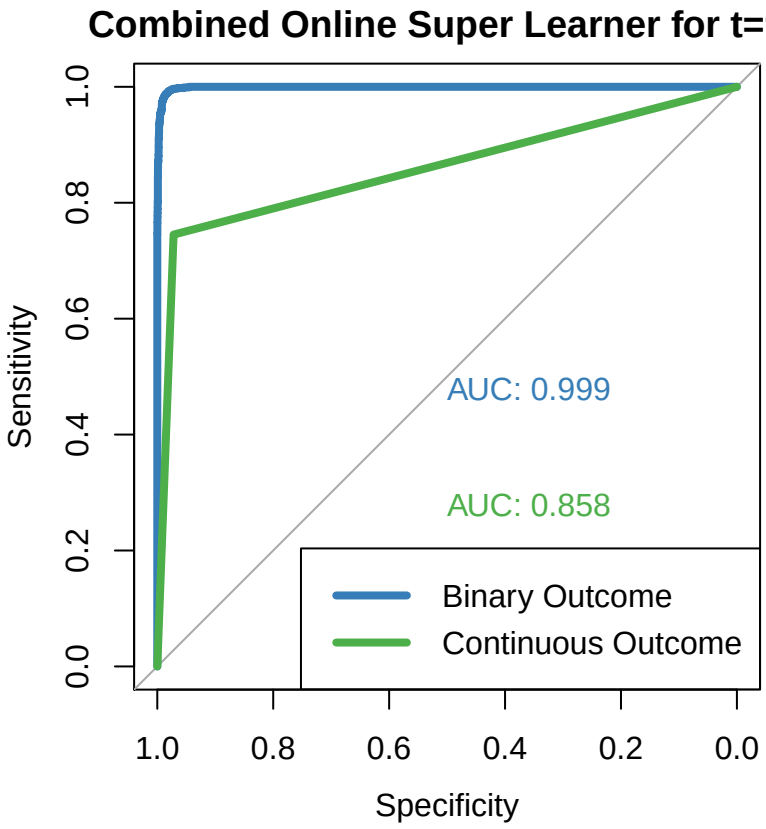
```
##
## Call:
## roc.default(response = calc_bin_t30$truth$truth, predictor = calc_bin_t30$pred_fin$pred,     plot = T
##
## Data: calc_bin_t30$pred_fin$pred in 7052 controls (calc_bin_t30$truth$truth 0) < 948 cases (calc_bin_
## Area under the curve: 0.9991
```



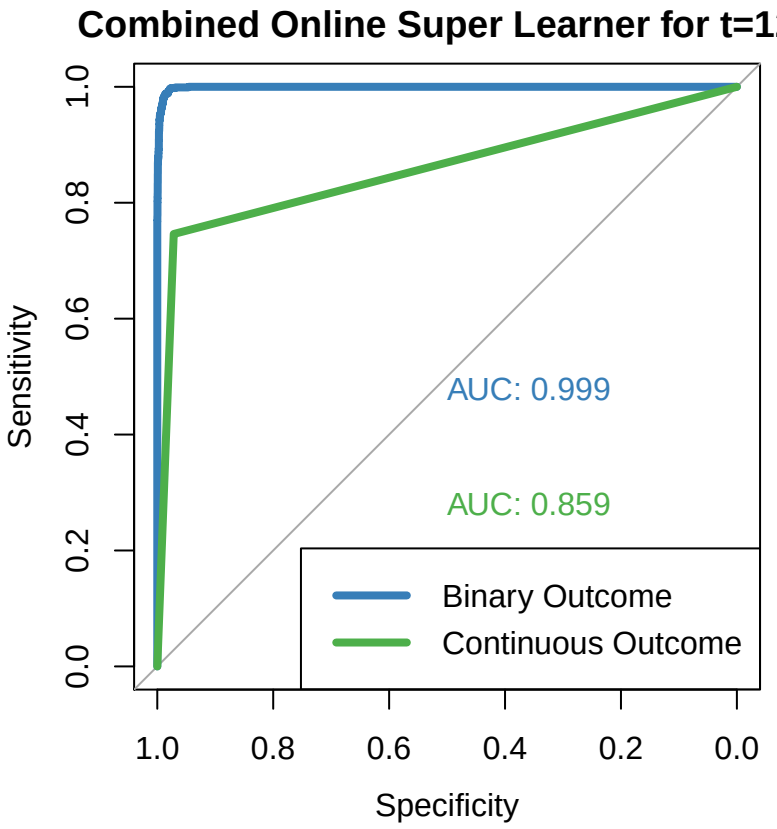

```
##
## Call:
## roc.default(response = calc_bin_t60$truth$truth, predictor = calc_bin_t60$pred_fin$pred,     plot = T
##
## Data: calc_bin_t60$pred_fin$pred in 7052 controls (calc_bin_t60$truth$truth 0) < 948 cases (calc_bin_
## Area under the curve: 0.9991
```



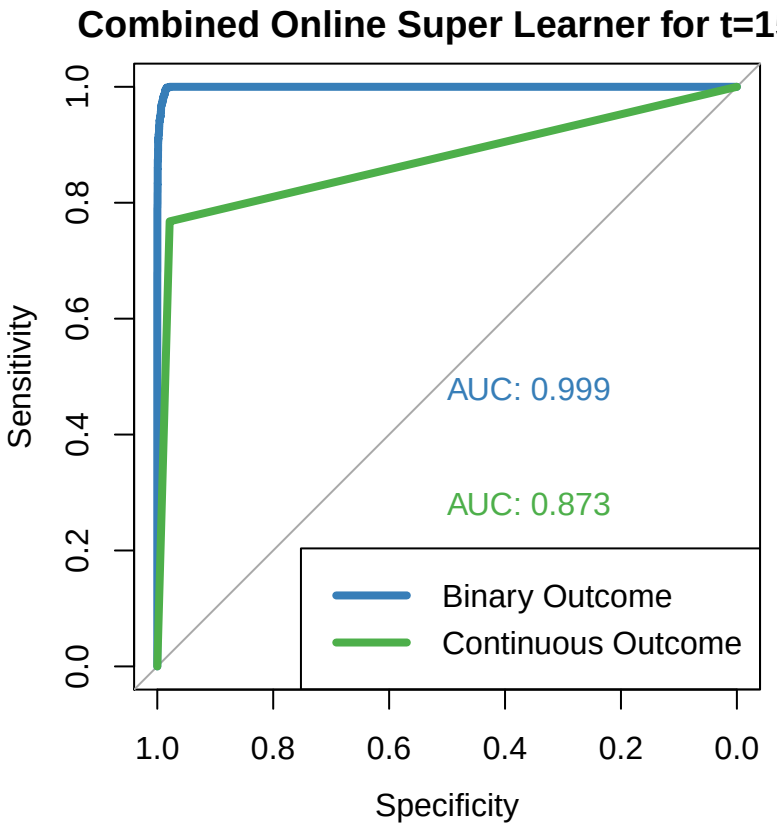

```
##
## Call:
## roc.default(response = calc_bin_t90$truth$truth, predictor = calc_bin_t90$pred_fin$pred,     plot = T
##
## Data: calc_bin_t90$pred_fin$pred in 7047 controls (calc_bin_t90$truth$truth 0) < 953 cases (calc_bin_
## Area under the curve: 0.999
```



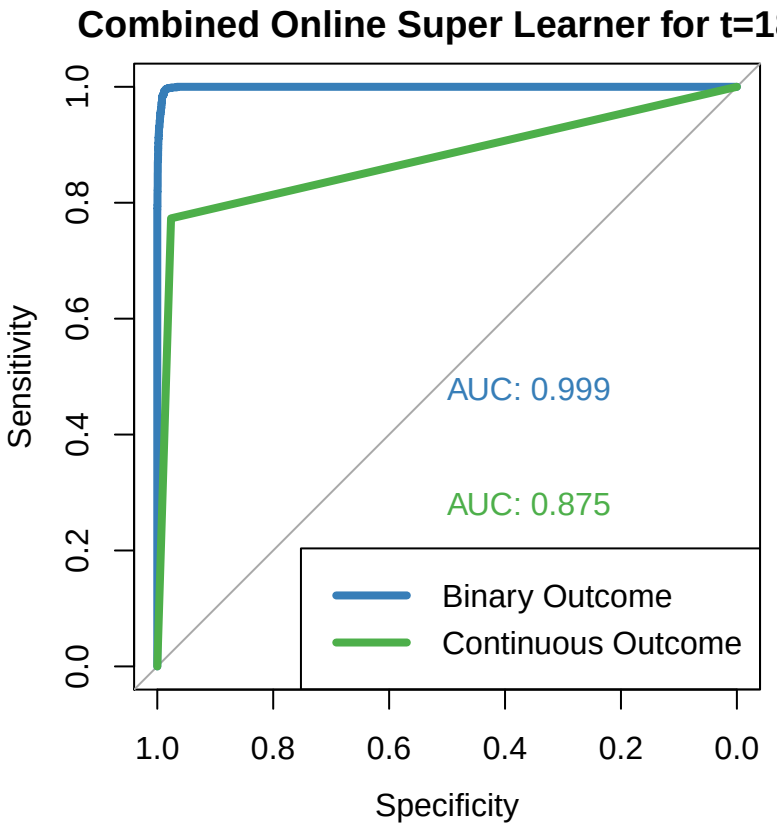

```
##
## Call:
## roc.default(response = calc_bin_t120$truth$truth, predictor = calc_bin_t120$pred_fin$pred,     plot =
##
## Data: calc_bin_t120$pred_fin$pred in 7047 controls (calc_bin_t120$truth$truth 0) < 953 cases (calc_bi
## Area under the curve: 0.9991
```



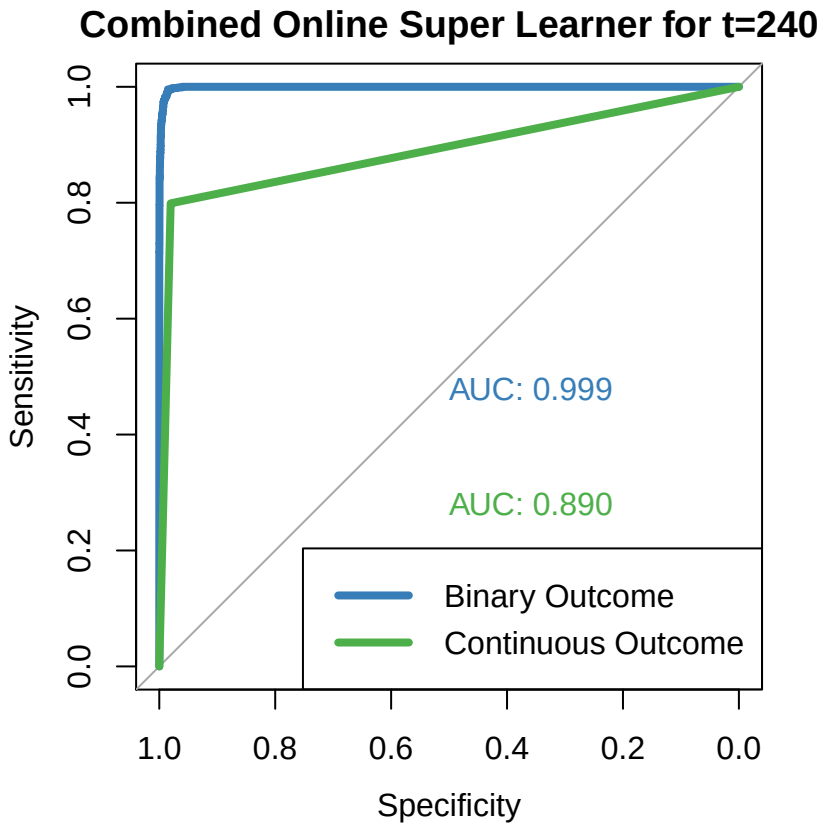

```
##
## Call:
## roc.default(response = calc_bin_t150$truth$truth, predictor = calc_bin_t150$pred_fin$pred,     plot =
##
## Data: calc_bin_t150$pred_fin$pred in 7035 controls (calc_bin_t150$truth$truth 0) < 965 cases (calc_bi
## Area under the curve: 0.9993
```




```
##
## Call:
## roc.default(response = calc_bin_t180$truth$truth, predictor = calc_bin_t180$pred_fin$pred,     plot =
##
## Data: calc_bin_t180$pred_fin$pred in 7035 controls (calc_bin_t180$truth$truth 0) < 965 cases (calc_bi
## Area under the curve: 0.9993
```




```
##
## Call:
## roc.default(response = calc_bin_t240$truth$truth, predictor = calc_bin_t240$pred_fin$pred,     plot =
##
## Data: calc_bin_t240$pred_fin$pred in 7075 controls (calc_bin_t240$truth$truth 0) < 925 cases (calc_bi
## Area under the curve: 0.9992
```


Combined Online Super Learner for t=240
Sensitivity
1.0
0.8
0.6
0.4
0.2
0.0
AUC: 0.999
AUC: 0.890
Binary Outcome
Continuous Outcome
1.0
0.8
0.6
0.4
0.2
0.0
Specificity