Sepsis Hour 1: logistic

Carter Sevick

November 6, 2019

In this report the glmnet model is used only as a variable reduction algorithm and the model under study is estimated by ML.



Based on the plot, 0.011691 was selected as the lambda resulting from keeping 9 variables.

##### Logistic regression parameter estimates

## Estimate Std. Error z value Pr(>|z|)
## (Intercept) 5.033602449 2.4940816495 2.0182188 4.356848e-02
## Temp, initial -0.134435997 0.0637533586 -2.1086889 3.497145e-02
## SBP, initial -0.028297218 0.0069854109 -4.0509025 5.102047e-05
## Oncological comorbidity -0.996658769 0.3241940782 -3.0742658 2.110213e-03
## ArrivalMode\_Cat 0.328028614 0.2185781862 1.5007381 1.334233e-01
## Hospitalized Last Year -0.307763485 0.1801339213 -1.7085260 8.753879e-02
## Central line -0.217669543 0.3425222096 -0.6354903 5.251087e-01
## DBP less than 69 0.027018102 0.0106757934 2.5307817 1.138087e-02
## Age by respiration rate 0.001659762 0.0008845469 1.8763979 6.060067e-02
## Shock index by Age 0.073323982 0.0202752356 3.6164306 2.986934e-04

Rate of septic shock in each data set:

|  |  |  |  |
| --- | --- | --- | --- |
| Dataset | N | Cases | Percent |
| Training | 1697 | 192 | 11.31 |
| Time | 698 | 80 | 11.46 |
| Geographic | 69 | 10 | 14.49 |
| Comparison of expected VS observed values, within deciles of risk. |  |  |  |
|  |  |  |  |
|  |  |  |  |

Training set 

Calibration line intercept and slope

|  |  |
| --- | --- |
| Intercept | Slope |
| -0.004 | 1.032 |

Time test



Calibration line intercept and slope

|  |  |
| --- | --- |
| Intercept | Slope |
| 0.01 | 0.961 |

Geography



Calibration line intercept and slope

|  |  |
| --- | --- |
| Intercept | Slope |
| -0.043 | 1.08 |

####

#### Area under the curve:

Training Set

## Area under the curve: 0.794

## 95% CI: 0.7614-0.8265 (DeLong)

Time holdout

## Area under the curve: 0.7505

## 95% CI: 0.6891-0.812 (DeLong)

Geography holdout

## Area under the curve: 0.8525

## 95% CI: 0.7126-0.9925 (DeLong)

ROC plots for the three sets: 

ROC tests:

Training VS time holdout

Septic shock

##
## DeLong's test for two ROC curves
##
## data: rocobj and ttrocobj
## D = 1.2253, df = 1107.4, p-value = 0.2207
## alternative hypothesis: true difference in AUC is not equal to 0
## sample estimates:
## AUC of roc1 AUC of roc2
## 0.7939819 0.7505057

Training VS geographic holdout

Septic shock

##
## DeLong's test for two ROC curves
##
## data: rocobj and grrocobj
## D = -0.79899, df = 75.555, p-value = 0.4268
## alternative hypothesis: true difference in AUC is not equal to 0
## sample estimates:
## AUC of roc1 AUC of roc2
## 0.7939819 0.8525424

Time VS geographic holdout

Septic shock

##
## DeLong's test for two ROC curves
##
## data: ttrocobj and grrocobj
## D = -1.3087, df = 96.414, p-value = 0.1937
## alternative hypothesis: true difference in AUC is not equal to 0
## sample estimates:
## AUC of roc1 AUC of roc2
## 0.7505057 0.8525424

Table 2. Test characteristics of the model in training and test sets, evaluated with 2 thresholds: one designed to optimize both sensitivity and specificity (Youden’s J), and with a threshold designed for a 90% sensitivity.

The estimate and 95% CI on Youden’s threshold is: 0.16 ( 0.08, 0.16)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| group | Threshold | Sensitivity | Specificity | PPV | NPV |
| Training - Youden’s | 0.16 | 0.63 (0.56, 0.7) | 0.83 (0.81, 0.85) | 0.32 (0.27, 0.37) | 0.95 (0.93, 0.96) |
| Training - 90% sens | 0.06 | 0.9 (0.85, 0.94) | 0.44 (0.42, 0.47) | 0.17 (0.15, 0.2) | 0.97 (0.96, 0.98) |
| Temporal Test - Youden’s | 0.16 | 0.55 (0.43, 0.66) | 0.84 (0.81, 0.86) | 0.3 (0.23, 0.39) | 0.93 (0.91, 0.95) |
| Temporal Test - 90% sens | 0.06 | 0.81 (0.71, 0.89) | 0.46 (0.42, 0.5) | 0.16 (0.13, 0.2) | 0.95 (0.92, 0.97) |
| Geograpic Test - Youden’s | 0.16 | 0.8 (0.44, 0.97) | 0.69 (0.56, 0.81) | 0.31 (0.14, 0.52) | 0.95 (0.84, 0.99) |
| Geograpic Test - 90% sens | 0.06 | 0.9 (0.55, 1) | 0.31 (0.19, 0.44) | 0.18 (0.09, 0.31) | 0.95 (0.74, 1) |

###

### Precision - recall curves

Training 

|  |  |  |  |
| --- | --- | --- | --- |
| modnames | dsids | curvetypes | aucs |
| m1 | 1 | ROC | 0.7939819 |
| m1 | 1 | PRC | 0.3216627 |

Time test 

|  |  |  |  |
| --- | --- | --- | --- |
| modnames | dsids | curvetypes | aucs |
| m1 | 1 | ROC | 0.7505057 |
| m1 | 1 | PRC | 0.3548257 |

Geography test 

|  |  |  |  |
| --- | --- | --- | --- |
| modnames | dsids | curvetypes | aucs |
| m1 | 1 | ROC | 0.8525424 |
| m1 | 1 | PRC | 0.5304875 |

####

#### Shock with Vasoactives

Rate of Shock with Vasoactives in each data set:

|  |  |  |  |
| --- | --- | --- | --- |
| Dataset | N | Cases | Percent |
| Training | 1697 | 76 | 4.48 |
| Time | 698 | 29 | 4.15 |
| Geographic | 69 | 4 | 5.80 |

Training set:

## Area under the curve: 0.7471

## 95% CI: 0.6885-0.8057 (DeLong)

Time test set:

## Area under the curve: 0.7106

## 95% CI: 0.6059-0.8153 (DeLong)

Geographic test set:

## Area under the curve: 0.8192

## 95% CI: 0.6742-0.9643 (DeLong)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| group | Threshold | Sensitivity | Specificity | PPV | NPV |
| Training - Youden’s | 0.16 | 0.64 (0.53, 0.75) | 0.8 (0.77, 0.81) | 0.13 (0.1, 0.17) | 0.98 (0.97, 0.99) |
| Temporal Test - Youden’s | 0.16 | 0.48 (0.29, 0.67) | 0.8 (0.77, 0.83) | 0.1 (0.05, 0.16) | 0.97 (0.96, 0.98) |
| Geograpic Test - Youden’s | 0.16 | 0.75 (0.19, 0.99) | 0.65 (0.52, 0.76) | 0.12 (0.02, 0.3) | 0.98 (0.88, 1) |
| Training - 90% sens | 0.06 | 0.83 (0.73, 0.91) | 0.42 (0.39, 0.44) | 0.06 (0.05, 0.08) | 0.98 (0.97, 0.99) |
| Temporal Test - 90% sens | 0.06 | 0.79 (0.6, 0.92) | 0.44 (0.4, 0.48) | 0.06 (0.04, 0.09) | 0.98 (0.96, 0.99) |
| Geograpic Test - 90% sens | 0.06 | 1 (0.4, 1) | 0.29 (0.19, 0.42) | 0.08 (0.02, 0.19) | 1 (0.82, 1) |