Default

Overall Statistics

Accuracy : 0.3323

95% CI : (0.3022, 0.3634)

No Information Rate : 0.9766

P-Value [Acc > NIR] : 1

Kappa : -0.0106

Mcnemar's Test P-Value : <2e-16

Statistics by Class:

Class: 0 Class: 1 Class: 2

Sensitivity 0.33479 0.166667 0.500000

Specificity 0.45455 0.682953 0.654545

Pos Pred Value 0.96238 0.010169 0.006154

Neg Pred Value 0.01613 0.976708 0.996743

Prevalence 0.97657 0.019169 0.004260

Detection Rate 0.32694 0.003195 0.002130

Detection Prevalence 0.33972 0.314164 0.346113

Balanced Accuracy 0.39467 0.424810 0.577273

With weights

Overall Statistics

Accuracy : 0.3365

95% CI : (0.3063, 0.3678)

No Information Rate : 0.9766

P-Value [Acc > NIR] : 1

Kappa : 0.005

Mcnemar's Test P-Value : <2e-16

Statistics by Class:

Class: 0 Class: 1 Class: 2

Sensitivity 0.33588 0.33333 0.50000

Specificity 0.77273 0.67427 0.65989

Pos Pred Value 0.98403 0.01961 0.00625

Neg Pred Value 0.02716 0.98104 0.99677

Prevalence 0.97657 0.01917 0.00426

Detection Rate 0.32801 0.00639 0.00213

Detection Prevalence 0.33333 0.32588 0.34079

Balanced Accuracy 0.55430 0.50380 0.57995

Hyper parameter tuned with argmax classification:

Overall Statistics

Accuracy : 0.9723

95% CI : (0.9597, 0.9818)

No Information Rate : 0.9723

P-Value [Acc > NIR] : 0.5519

Kappa : 0

Mcnemar's Test P-Value : NA

Statistics by Class:

Class: 0 Class: 1 Class: 2

Sensitivity 1.0000 0.00000 0.00000

Specificity 0.0000 1.00000 1.00000

Pos Pred Value 0.9723 NaN NaN

Neg Pred Value NaN 0.97444 0.99787

Prevalence 0.9723 0.02556 0.00213

Detection Rate 0.9723 0.00000 0.00000

Detection Prevalence 1.0000 0.00000 0.00000

Balanced Accuracy 0.5000 0.50000 0.50000

Tuned and calibrated:  
Overall Statistics

Accuracy : 0.886

95% CI : (0.864, 0.9057)

No Information Rate : 0.9766

P-Value [Acc > NIR] : 1

Kappa : -0.0328

Mcnemar's Test P-Value : NA

Statistics by Class:

Class: 0 Class: 1 Class: 2

Sensitivity 0.9073 0.00000 0.00000

Specificity 0.0000 0.92400 0.98396

Pos Pred Value 0.9742 0.00000 0.00000

Neg Pred Value 0.0000 0.97929 0.99567

Prevalence 0.9766 0.01917 0.00426

Detection Rate 0.8860 0.00000 0.00000

Detection Prevalence 0.9095 0.07455 0.01597

Balanced Accuracy 0.4537 0.46200 0.49198

✅ Best multiclass log loss: 1.08137

✅ Best Parameters:

$booster

[1] "gbtree"

$objective

[1] "multi:softprob"

$eval\_metric

[1] "mlogloss"

$max\_depth

[1] 3

$min\_child\_weight

[1] 1

$eta

[1] 0.05

$gamma

[1] 0.5

$subsample

[1] 0.9

$colsample\_bytree

[1] 0.9

$nthread

[1] 1

$num\_class

[1] 3

✅ Best Number of Rounds: 50

Externalizing

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | CBCL | | | | BPM | | | |
| Model Configuration | Default | With Weights | Tuned + Argmax | Tuned + Threshold-Cal. | Default | With Weights | Tuned + Argmax | Tuned + Threshold-Cal. |
| Overall Performance | | | | | | | | |
| Accuracy | 0.3478 | 0.3498 | 0.9835 | 0.9546 | 0.3323 | 0.3365 | 0.9723 | 0.886 |
| 95% CI | 0.3178–0.3787 | 0.3198–0.3808 | 0.9733–0.9905 | 0.9395–0.9668 | 0.3022–0.3634 | 0.3063–0.3678 | 0.9597–0.9818 | 0.864–0.9057 |
| Cohen’s Kappa | 0.0113 | 0.0084 | 0 | -0.0173 | -0.0106 | 0.005 | 0 | -0.0328 |
| Class Balance | | | | | | | | |
| Balanced Accuracy | 0.5524 | 0.5545 | 0.5 | 0.4879 | 0.4248 | 0.5038 | 0.5 | 0.462 |
| Sensitivity (Borderline) | 0.4 | 0.4 | 0.0 | 0.0 | 0.1667 | 0.3333 | 0.0 | 0.0 |
| Sensitivity (Clinical) | 0.7273 | 0.5455 | 0.0 | 0.0 | 0.5 | 0.5 | 0.0 | 0.0 |
| Specificity (Borderline) | 0.6705 | 0.6621 | 1.0 | 0.9875 | 0.683 | 0.6743 | 1.0 | 0.924 |
| Specificity (Clinical) | 0.6754 | 0.6858 | 1.0 | 0.9885 | 0.6545 | 0.6599 | 1.0 | 0.984 |
| Class-Level Discrimination | | | | | | | | |
| Precision (Borderline) | 0.0125 | 0.0122 | NaN | 0.0 | 0.0102 | 0.0196 | NaN | 0.0 |
| Precision (Clinical) | 0.0251 | 0.0195 | NaN | 0.0 | 0.0062 | 0.0063 | NaN | 0.0 |
| F1 Score (Borderline) | 0.0236 | 0.0236 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| F1 Score (Clinical) | 0.0248 | 0.0248 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |

|  |  |  |
| --- | --- | --- |
| Parameter | CBCL | BPM |
| Booster | Gbtree | Gbtree |
| Objective | Multi:softprob | Multi:softprob |
| Evaluation Metric | mlogloss | mlogloss |
| Max Depth | 3 | 3 |
| Min Child Weight | 5 | 1 |
| Eta | 0.1 | 0.05 |
| Gamma | 0.5 | 0.5 |
| Subsample | 0.6 | 0.9 |
| Colsample by Tree | 0.9 | 0.9 |
| Number of Classes | 3 | 3 |
| Best number of rounds | 150 | 50 |
| Best multiclass Log Loss | 1.091149 | 1.08137 |

**Data Split for Externalizing Symptom**

BPM

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Split | Class 0 | Class 1 | Class 2 | Total | Minority % |
| Full | 6096 | 144 | 21 | 6261 | ~2.5% |
| Training | 4266 | 102 | 15 | 4383 | ~2.7% |
| Calibration | 913 | 24 | 2 | 939 | ~2.8% |
| Test | 917 | 18 | 4 | 939 | ~2.3% |

CBCL

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Split | Class 0 | Class 1 | Class 2 | Total | Minority % |
| Full | 6356 | 66 | 38 | 6460 | ~1.6% |
| Training | 4455 | 44 | 23 | 4522 | ~1.5% |
| Calibration | 953 | 12 | 4 | 969 | ~1.7% |
| Test | 948 | 10 | 11 | 969 | ~2.2% |

Best thersholds bpm: [1] 0.95 0.04 0.00

Cbcl: [1] 0.98 0.01 0.01

CBCL

With weights:  
Confusion Matrix and Statistics

Reference

Prediction 0 1 2

0 329 3 2

1 321 4 3

2 298 3 6

Overall Statistics

Accuracy : 0.3498

95% CI : (0.3198, 0.3808)

No Information Rate : 0.9783

P-Value [Acc > NIR] : 1

Kappa : 0.0084

Mcnemar's Test P-Value : <2e-16

Statistics by Class:

Class: 0 Class: 1 Class: 2

Sensitivity 0.3470 0.400000 0.545455

Specificity 0.7619 0.662148 0.685804

Pos Pred Value 0.9850 0.012195 0.019544

Neg Pred Value 0.0252 0.990640 0.992447

Prevalence 0.9783 0.010320 0.011352

Detection Rate 0.3395 0.004128 0.006192

Detection Prevalence 0.3447 0.338493 0.316821

Balanced Accuracy 0.5545 0.531074 0.615629

tuned argmax

Overall Statistics

Accuracy : 0.9835

95% CI : (0.9733, 0.9905)

No Information Rate : 0.9835

P-Value [Acc > NIR] : 0.566

Kappa : 0

Mcnemar's Test P-Value : NA

Statistics by Class:

Class: 0 Class: 1 Class: 2

Sensitivity 1.0000 0.00000 0.000000

Specificity 0.0000 1.00000 1.000000

Pos Pred Value 0.9835 NaN NaN

Neg Pred Value NaN 0.98762 0.995872

Prevalence 0.9835 0.01238 0.004128

Detection Rate 0.9835 0.00000 0.000000

Detection Prevalence 1.0000 0.00000 0.000000

Balanced Accuracy 0.5000 0.50000 0.500000

Tuned and calibrated on test set:  
Overall Statistics

Accuracy : 0.9546

95% CI : (0.9395, 0.9668)

No Information Rate : 0.9783

P-Value [Acc > NIR] : 1

Kappa : -0.0173

Mcnemar's Test P-Value : NA

Statistics by Class:

Class: 0 Class: 1 Class: 2

Sensitivity 0.9757 0.00000 0.00000

Specificity 0.0000 0.98749 0.98852

Pos Pred Value 0.9778 0.00000 0.00000

Neg Pred Value 0.0000 0.98955 0.98852

Prevalence 0.9783 0.01032 0.01135

Detection Rate 0.9546 0.00000 0.00000

Detection Prevalence 0.9763 0.01238 0.01135

Balanced Accuracy 0.4879 0.49374 0.49426