

| | training dataset | | | | | | | validation cohort 1 | | | | | | | validation cohort 2 | | | | | | |
|-------------|------------------|-------|------|----|----|---------|------|---------------------|-------|------|------|------|---------|------|---------------------|-------|------|------|------|---------|------|
| specificity | 0.83 | 0.84 | 0.99 | 1 | 1 | 0.9 | 0.85 | 0.8 | 0.81 | 0.67 | 0.88 | 0.78 | 0.69 | 0.86 | 0.8 | 0.81 | 0.67 | 0.88 | 0.78 | 0.69 | 0.86 |
| sensitivity | 0.7 | 0.7 | 0.89 | 1 | 1 | 0.7 | 0.68 | 0.58 | 0.58 | 0.75 | 0.19 | 0.53 | 0.61 | 0.58 | 0.58 | 0.58 | 0.75 | 0.19 | 0.53 | 0.61 | 0.58 |
| accuracy | 0.82 | 0.82 | 0.98 | 1 | 1 | 0.88 | 0.83 | 0.79 | 0.8 | 0.68 | 0.84 | 0.76 | 0.69 | 0.84 | 0.79 | 0.8 | 0.68 | 0.84 | 0.76 | 0.69 | 0.84 |
| ppv | 0.3 | 0.31 | 0.9 | 1 | 1 | 0.44 | 0.32 | 0.17 | 0.17 | 0.13 | 0.1 | 0.14 | 0.12 | 0.22 | 0.17 | 0.17 | 0.13 | 0.1 | 0.14 | 0.12 | 0.22 |
| npv | 0.96 | 0.96 | 0.99 | 1 | 1 | 0.97 | 0.96 | 0.97 | 0.97 | 0.98 | 0.94 | 0.96 | 0.96 | 0.97 | 0.97 | 0.97 | 0.98 | 0.94 | 0.96 | 0.96 | 0.97 |
| precision | 0.3 | 0.31 | 0.9 | 1 | 1 | 0.44 | 0.32 | 0.17 | 0.17 | 0.13 | 0.1 | 0.14 | 0.12 | 0.22 | 0.17 | 0.17 | 0.13 | 0.1 | 0.14 | 0.12 | 0.22 |
| recall | 0.7 | 0.7 | 0.89 | 1 | 1 | 0.7 | 0.68 | 0.58 | 0.58 | 0.75 | 0.19 | 0.53 | 0.61 | 0.58 | 0.58 | 0.58 | 0.75 | 0.19 | 0.53 | 0.61 | 0.58 |
| F1 | 0.42 | 0.43 | 0.9 | 1 | 1 | 0.54 | 0.43 | 0.26 | 0.27 | 0.23 | 0.13 | 0.22 | 0.2 | 0.32 | 0.26 | 0.27 | 0.23 | 0.13 | 0.22 | 0.2 | 0.32 |
| AUC | 0.77 | 0.77 | 0.94 | 1 | 1 | 0.83 | 0.81 | 0.66 | 0.66 | 0.74 | 0.54 | 0.68 | 0.66 | 0.66 | 0.66 | 0.66 | 0.74 | 0.54 | 0.68 | 0.66 | 0.66 |
| | GLM | LASSO | SVM | DT | RF | XGBoost | GBM | GLM | LASSO | SVM | DT | RF | XGBoost | GBM | GLM | LASSO | SVM | DT | RF | XGBoost | GBM |