Supplementary Table 3. Performance of each model for prediction (training set)

Model	AUC	Accuracy	Sensitivity	Specificity	PPV	NPV	F1 score	Kappa
XGBoost	0.765 (0.606–0.919)	0.807 (0.768–0.845)	0.674 (0.570–0.779)	0.848 (0.812–0.885)	0.311 (0.257–0.366)	0.949 (0.928–0.971)	0.413 (0.358–0.468)	0.304 (0.248–0.360)
logistic	0.737 (0.583–0.892)	0.763 (0.696–0.830)	0.666 (0.578–0.754)	0.814 (0.745–0.882)	0.31 (0.237–0.384)	0.932 (0.920–0.944)	0.408 (0.339–0.476)	0.262 (0.183–0.341)
LightGBM	0.686 (0.530–0.841)	0.86 (0.827–0.893)	0.468 (0.372–0.564)	0.884 (0.833–0.935)	-	0.917 (0.894–0.940)	-	0.141 (0.063–0.219)
MLP	0.482 (0.310–0.654)	0.653 (0.526–0.780)	0.487 (0.289–0.685)	0.686 (0.482–0.889)	-	0.878 (0.838–0.918)	-	0.031 (–0.027–0.089)
SVM	0.577 (0.411–0.744)	0.742 (0.660–0.825)	0.652 (0.461–0.843)	0.625 (0.463–0.786)	0.221 (0.137–0.306)	0.909 (0.884–0.934)	0.31 (0.209–0.412)	0.097 (0.046–0.147)

Values are presented as 95% confidence interval.
AUC, area under the curve; PPV, positive predictive value; NPV, negative predictive value; XGBoost, extreme gradient boosting; LightGBM, light gradient boosting machine; MLP, multiple layers perception; SVM, support vector machine.