

Appendix A. Hyperparameter Configurations

Table A1 Key Hyperparameter Configurations

Algorithm	Key hyperparameter settings
Logistic Regression	solver = “lbfgs”; max_iter = 5000; class_weight = “balanced”
Random Forest	n_estimators = 1000; min_samples_leaf = 3; class_weight = “balanced_subsample”
HistGradientBoosting	learning_rate = 0.06; max_leaf_nodes = 31; l2_regularization = 0.10

Appendix B. Variable Categorization

Table B1 Categorization of Variables in the Modeling Pipeline

Variable category	Variables
Outcome-definition variables, excluded from predictors	FBS; current antidiabetic medication
Administrative variables, excluded from predictors	Participant ID; visit code; phase indicator; date-related variables
Candidate predictors before selection	Age; sex; BMI; waist circumference; blood-pressure measures; total cholesterol; triglycerides; HDL; LDL; physical activity; BAI anxiety; BDI-II depression; derived TG/HDL ratio, pulse pressure, mean arterial pressure, and BMI thresholds
Final selected predictors	Age; triglycerides; waist circumference; TG/HDL ratio; LDL; BMI; pulse pressure; total cholesterol; anxiety; depression; systolic blood pressure; HDL

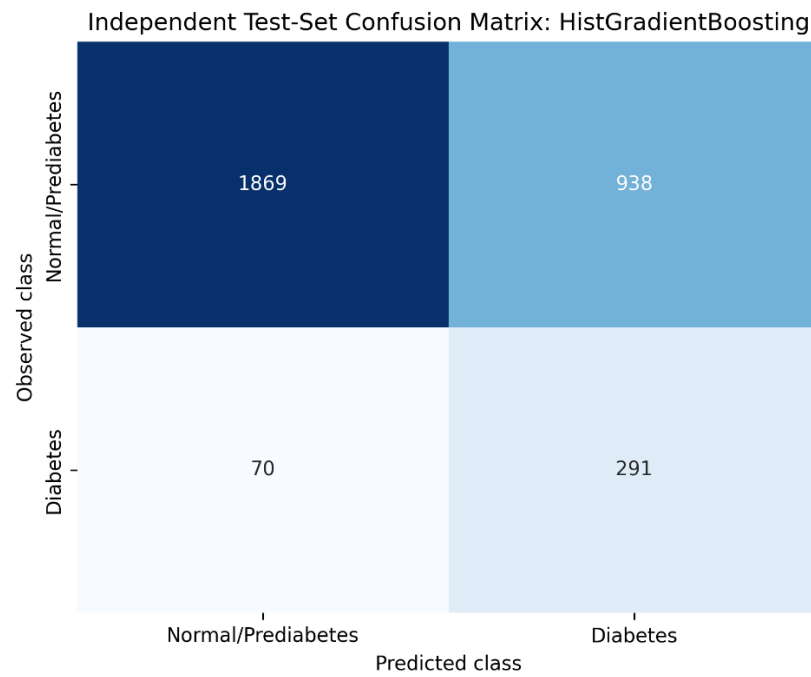


Figure S1. Supplementary Figure S1. Visual confusion-matrix representation of the classification counts reported in Table 4 for calibrated HistGradientBoosting in the hold-out record-level test set, using the selected threshold of 0.119.

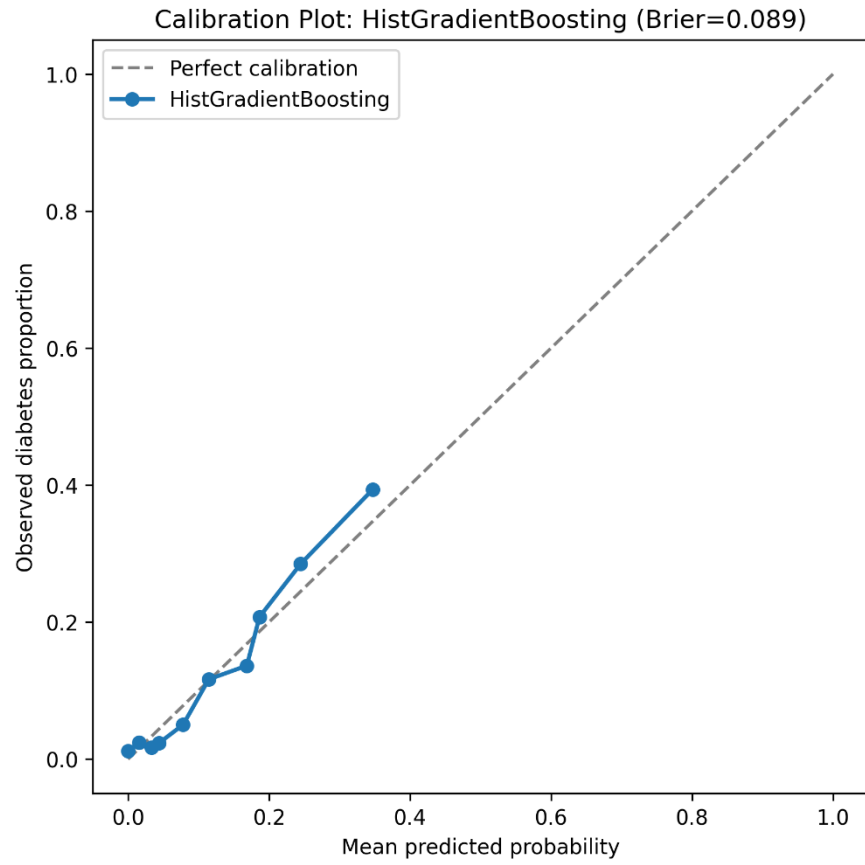


Figure S2. Calibration plot for calibrated HistGradientBoosting in the hold-out record-level test set. The Brier score was 0.089.

Supplementary Table S1. Class distribution

Dataset	Glycemic_class	N	Percent
Overall	Normal	10900	68.82
Overall	Prediabetes	3136	19.8
Overall	Diabetes	1803	11.38
Development	Normal	8703	68.68
Development	Prediabetes	2526	19.94
Development	Diabetes	1442	11.38
Hold-out test	Normal	2197	69.35
Hold-out test	Prediabetes	610	19.26
Hold-out test	Diabetes	361	11.4

Supplementary Table S2. feature-selection ranking

Predictor	Mean permutation importance	SD	Rank	Retained for modeling
age_exam	0.089651	0.003421	1	TRUE
tg	0.019	0.000723	2	TRUE
a5_wc	0.018	0.000768	3	TRUE
tg_hdl_ratio	0.005683	0.000317	4	TRUE
ldl	0.005655	0.00027	5	TRUE
bmi	0.00486	0.00022	6	TRUE
pulse_pressure	0.003926	0.000173	7	TRUE
chol	0.003153	0.000143	8	TRUE
anxiety	0.002035	0.000121	9	TRUE
depression	0.001989	0.000121	10	TRUE
e4_sys	0.001766	0.000127	11	TRUE
hdl	0.001745	0.000109	12	TRUE
e3_sys	0.001279	9.88e-05	13	FALSE

d2_sex	0.001155	8.41e-05	14	FALSE
mean_arterial_pressur e	0.000972	5.35e-05	15	FALSE
phys_level	0.000482	3.45e-05	16	FALSE
bmi_ge_25	0.000233	2.38e-05	17	FALSE
e3_dias	0.000215	1.99e-05	18	FALSE
e4_dias	0.000207	1.93e-05	19	FALSE
bmi_ge_30	5.39e-05	5.82e-06	20	FALSE

Supplementary Table S3. Extended hold-out record-level test performance metrics for the evaluated machine-learning models.

Model	Threshold	Threshold-subset F1	Accuracy (95% CI)	Macro-F1 (95% CI)	Precision (95% CI)	NPV (95% CI)
Random Forest	0.055	0.269	0.389 (0.373–0.406)	0.371 (0.355–0.386)	0.153 (0.138–0.167)	0.983 (0.974–0.991)
HistGradientBoosting	0.119	0.392	0.682 (0.664–0.697)	0.577 (0.558–0.595)	0.237 (0.214–0.261)	0.964 (0.956–0.972)
Logistic Regression	0.084	0.362	0.691 (0.674–0.706)	0.580 (0.560–0.600)	0.238 (0.214–0.263)	0.960 (0.951–0.968)

Thresholds were selected within the development data by maximizing F1 score subject to a minimum sensitivity target of 0.75 and were applied unchanged to the hold-out record-level test set. Values are proportions. Ninety-five percent confidence intervals were estimated using 1,000 bootstrap resamples of the hold-out record-level test set.