

Supplemental Material

S1 Table. The community information collected by the external derivation cohort

Number	Time	Community name (Location)	Total number	Missing number n (%)	Qualified number n (%)
01	April 2nd to 4th	He Mei Community, Yu Zui Town	62	2 (3.2)	60 (96.8)
01	April 11th to 12th	Ziyu Village Community, Shuangbei Sub-district	58	2 (3.4)	56 (96.6)
01	April 17th to 19th	Dongzheng Community, Yuejin Village Sub-district	63	4 (6.3)	59 (93.7)
01	May 8th to 10th	Yanjiang Village, Guangyang Town	76	4 (5.3)	72 (94.7)
01	May 15th to 18th	Shiniugang Community, Huxi Sub-district	71	5 (7.0)	66 (93.0)
01	May 22th to 25th	Shuikou Community, Jinshan Town	58	1 (1.7)	57 (98.3)
02	June 2nd to 4th	Longjiang Village, Longmen Sub-district	51	3 (5.9)	48 (94.1)
02	June 11nd to 13th	Lianhua Village, Pingshan Town	65	2 (3.1)	63 (96.9)
03	June 20nd to 22th	Guiluo Village, Zhongping Township	55	3 (5.5)	52 (94.5)
Toatl			559	26 (4.7)	533 (95.3)

Note: Reflecting Chongqing's composition and counts of districts and counties, we selected 6, 2, and 1 communities or villages from districts, counties, and autonomous counties, respectively, ensuring no overlap with the communities or villages in the derivation cohort. The data collection period for external validation is from April to June 2025. External validation quality control was carried out in accordance with the method of derivation cohort, Details in manuscript “2.1 Study design” . Number 01: Districts; Number02: Counties; Number03:Autonomous counties.

S2 Table. The MoCA score, demographic characteristics and health-related factors with the percentage of missing data

Characteristics	Derivation cohort	Percentage of missing data, %	External validation cohort	Percentage of missing data, %
Age,years	2131	0.00	533	0.00
Gender	2131	0.00	533	0.00
Marital Status	2021	5.16	521	2.25
Hypertension	2101	1.41	528	0.94
Hyperglycemia	2114	0.80	525	1.50
Hyperlipidemia	2099	1.50	514	3.56
Cerebral Hemorrhage	2112	0.89	509	4.50
Cerebral Infarction	2009	5.73	515	3.38
Cerebral Atrophy	2120	0.52	533	0.00
Brain Trauma	2131	0.00	532	0.19
Parkinson	2108	1.08	533	0.00
Epilepsy	2105	1.22	533	0.00
CHD	2100	1.45	529	0.75
Smoking	2118	0.61	533	0.00
Drinking	2001	6.10	511	4.10
Depression	2001	6.10	511	4.10
Education	2001	6.10	511	4.10
MoCA score	2115	9.92	519	2.63

Note: The 114 cases in which participants or their families requested withdrawal were not included in the missing-data handling. Variables with more than 20% missing data were excluded. Abbreviation: CHD, coronary heart disease.

S3 Table. The inclusion and exclusion criteria

A. Inclusion criteria
1.Age $\geq$ 60 years old.
2.Clear consciousness with the ability to communicate.
3.Comprehension of Chongqing dialect or Mandarin.
4.Permanent residents of Chongqing Municipality (and have resided for at least 6 months).
B. Exclusion criteria
1.Seizure phase of epilepsy.
2.Recent cerebral hemorrhage or infarction (Within the last three months).
3.Congenital or acquired intellectual disability.
4.Severe visual or hearing impairment.
5.Severe anxiety or depression.
6.Inability to perform activities of daily living (caused by delirium or psychiatric disorders).

S4 Table. The rules and Settings of LASSO feature selection

Parameters	Content
CV configuration	5-fold StratifiedKfold cross-validation, AUC score, 50 C candidate values (corresponding $\lambda=1/C$)
Lambda	$\lambda=1/C$, where Lambda.min is the λ corresponding to the optimal C (such as maximizing AUC), and Lambda.1se can be derived from the standard deviation of CV
Core parameters	L1 regularization, saga solver, balanced category weights, 2000 iterations, and 42 random seeds ensure model convergence and reproducibility

S5 Table. The basic principles, architectures and training details of the three deep learning

Deep learning	principles, architectures and details
SNN	Network architecture: Input layer (configurable, default 10) $\rightarrow$ hidden layer (configurable, default 12) $\rightarrow$ output layer (1 unit, binary classification). Activations and regularization: Hidden layer uses ReLU with Dropout set to 25% ($p = 0.25$); output layer uses Sigmoid to produce probabilities. Weight and bias initialization: Fully connected weights are initialized uniformly using Xavier initialization (gain = 0.8); biases are initialized to 0.001. Training monitoring: The implementation maintains built-in lists of losses (losses) and accuracy rates (accs) for training and validation, enabling real-time recording of the training process. Initialization log: Core configuration (input and hidden dimensions, Dropout ratio, etc.) is automatically printed at initialization.
DNN	Network architecture: The model uses a progressively reduced hidden-layer dimensionality by default (input $\rightarrow 32 \rightarrow 16 \rightarrow 12 \rightarrow$ output), and the number of hidden layers can be adjusted via the hidden_dims parameter. Layer composition: Each hidden layer consists of a fully connected layer followed by a ReLU activation and Dropout; the Dropout rate is configurable (default 0.25). The output layer is a single-unit fully connected layer with a Sigmoid activation. Modular implementation: All layers are assembled with nn.Sequential to simplify forward propagation logic. Weight initialization: As in the shallow network, weights are initialized with Xavier uniform (gain=0.8) and biases are fixed to 0.001. Training monitoring: The implementation preserves real-time logging of training and validation loss and accuracy, and prints key configuration details at initialization, including the hidden layer dimension list and the Dropout ratio.
ResNet	(1) Residual block (core component) design: Each residual block comprises a main branch and a residual branch. Main branch: input $\rightarrow$ fully connected layer 1 $\rightarrow$ ReLU $\rightarrow$ Dropout (default $p = 0.25$) $\rightarrow$ fully connected layer 2. Residual branch: if input and output dimensions differ, apply a 1×1 fully connected layer to match dimensions; if they match, use an identity mapping (nn.Identity). The two branch outputs are combined by elementwise addition and then passed through ReLU to perform residual feature learning. (2) Overall residual network architecture (SmallDataResNet): The network uses a lightweight design with two residual blocks in series by default (input $\rightarrow$

Deep learning	principles, architectures and details
	32-dimensional residual block → 16-dimensional residual block → output). The number and widths of residual blocks can be configured via the hidden_dims parameter. Regularization and output: a Dropout layer (default $p = 0.25$) follows the final residual block; a 1-dimensional fully connected layer plus a Sigmoid activation produces the binary classification probability. Weight initialization: all fully connected layers use Xavier uniform initialization (gain = 0.8); all biases are initialized to 0.001. Training and logging: training and validation metrics are tracked, and during initialization the model reports core information such as input dimensions, residual block dimension list, and Dropout ratio.

Abbreviations: SNN: Shallow Neural Network, DNN: Deep Neural Network, ResNet: Residual Neural Network.

S6 Table. Details of SHAP implementation

Methods	Content
Interpreter selection principle	For linear models, use 'LinearExplainer', and for tree models, use 'TreeExplainer'. This ensures that the model types are adapted and the results are accurate
Feature processing logic	The heat map and scatter plot identify and present the top n core features. In contrast, the interaction graph utilizes all features, and its calculations are based on a tree model. Notably, neither visualization displays the total sum of feature aggregations.
Stability guarantee	The entire process is fixed with random seeds and consistent background data sampling to ensure that the results are reproducible and suitable for academic publication

S7 Table. The comparison of MoCA score, sociodemographic and health-related factors among the training and external validation cohorts

Characteristics	Total(n=2664)	Training test set (n=2131)	External validation set (n=533)	Statistics	p
Gender				0.00	0.969
Femal[n,%]	1985 (74.5)	1587 (74.5)	398 (74.7)		
male[n,%]	679 (25.5)	544 (25.5)	135 (25.3)		
Age [Median, IQR]	73.0 (68.0–80.0)	73.0 (68.0–80.0)	73.0 (68.0–80.0)	-0.65	0.516
Education				1.73	0.421
Illiterate [n,%]	434 (16.3)	356 (16.7)	78 (14.6)		
Primary school [n,%]	945 (35.5)	758 (35.6)	187 (35.1)		
Junior high school and above [n,%]	1285 (48.2)	1017 (47.7)	268 (50.3)		
Marital Status				0.80	0.372
No[n,%]	1032 (38.7)	835 (39.2)	197 (37.0)		
Yes[n,%]	1632 (61.3)	1296 (60.8)	336 (63.0)		
Hypertension				0.00	0.968
No [n,%]	1434 (53.8)	1148 (53.9)	286 (53.7)		
Yes [n,%]	1230 (46.2)	983 (46.1)	247 (46.3)		
Hyperglycemia				3.52	0.061
No[n,%]	2108 (79.1)	1670 (78.4)	438 (82.2)		
Yes[n,%]	556 (20.9)	461 (21.6)	95 (17.8)		
Hyperlipidemia				0.50	0.480
NO[n,%]	2018 (75.8)	1621 (76.1)	397 (74.5)		
Yes[n,%]	646 (24.2)	510 (23.9)	136 (25.5)		
Cerebral Hemorrhage				0.55	0.457
No[n,%]	2603 (97.7)	2085 (97.8)	518 (97.2)		
Yes[n,%]	61 (2.3)	46 (2.2)	15 (2.8)		
Cerebral Infarction				0.31	0.580
No[n,%]	2164 (81.2)	1736 (81.5)	428 (80.3)		
Yes[n,%]	500 (18.8)	395 (18.5)	105 (19.7)		
Cerebral Atrophy				0.00	1.000
NO[n,%]	1414 (53.1)	1131 (53.1)	283 (53.1)		
Yes[n,%]	1250 (46.9)	1000 (46.9)	250 (46.9)		
Brain Trauma				0.84	0.358
No[n,%]	2419 (90.8)	1941 (91.1)	478 (89.7)		
Yes[n,%]	245 (9.2)	190 (8.9)	55 (10.3)		
Parkinson's Disease				1.29	0.256

Characteristics	Total(n=2664)	Training test set (n=2131)	External validation set (n=533)	Statistics	p
No[n,%]	2482 (93.2)	1979 (92.9)	503 (94.4)	0.00	1.000
Yes[n,%]	182 (6.8)	152 (7.1)	30 (5.6)		
Epilepsy [n,%]					
No[n,%]	2637 (99.0)	2109 (99.0)	528 (99.1)	0.00	0.960
Yes[n,%]	27 (1.0)	22 (1.0)	5 (0.9)		
CHD					
No[n,%]	1984 (74.5)	1588 (74.5)	396 (74.3)	6.39	0.011
Yes[n,%]	680 (25.5)	543 (25.5)	137 (25.7)		
Smoking					
No[n,%]	2439 (91.6)	1936 (90.8)	503 (94.4)	0.14	0.712
Yes[n,%]	225 (8.4)	195 (9.2)	30 (5.6)		
Drinking					
No[n,%]	2263 (84.9)	1807 (84.8)	456 (85.6)	0.53	0.466
Yes[n,%]	401 (15.1)	324 (15.2)	77 (14.4)		
Depression [n,%]					
No[n,%]	2611 (98.0)	2086 (97.9)	525 (98.5)	0.56	0.573
Yes[n,%]	53 (2.0)	45 (2.1)	8 (1.5)		
MoCA score[Median, IQR]	21.0 (16.0–24.0)	21.0 (16.0–25.0)	20.0 (16.0–24.0)		
GCD				0.00	1.000
No [n,%]	1790 (67.2)	1432 (67.2)	358 (67.2)		
Yes [n,%]	874 (32.8)	699 (32.8)	175 (32.8)		

Abbreviations: SD: Standard deviation, CHD: Coronary heart disease, GCD: Geriatric cognitive disorders.

S8 Table. Independent Variable Assignment Table

Independent variable	Assignment situation
Age	Z-score standardization
MoCA score	Z-score standardization
Education	Illiterate=0, Primary school=1, Junior high school and above=2
Marital Status	No=0, Yes=1
Hypertension	No=0, Yes=1
Hyperglycemia	No=0, Yes=1
Cerebral Hemorrhage	No=0, Yes=1
Cerebral Infarction	No=0, Yes=1
Cerebral Atrophy	No=0, Yes=1

S9 Table. The results of multivariate logistic regression analysis in the derivation cohort

Variable	β	SE	Wald χ^2	P	OR	95% CI for OR
Intercept	0.9391	0.2385	15.5049	0.0001	2.5578	1.6027 - 4.0821
Age	-0.482	0.0618	60.8896	0.0000*	0.6175	0.5471 - 0.6970
MoCA score	-0.1768	0.0109	263.3488	0.0000*	0.8379	0.8202 - 0.8560
Marital status ("No" as reference)	0.1305	0.1165	1.2547	0.2627	1.1394	0.9068 - 1.4316
Hypertension ("No" as reference)	-0.1429	0.1096	1.7003	0.1923	0.8669	0.6993 - 1.0745
Hyperglycemia ("No" as reference)	-0.3524	0.137	6.6123	0.0101*	0.703	0.5374 - 0.9196
Cerebral hemorrhage ("No" as reference)	-1.8011	0.5338	11.3842	0.0007*	0.1651	0.0580 - 0.4701
Cerebral infarction ("No" as reference)	-1.1555	0.1643	49.4548	0.0000*	0.3149	0.2282 - 0.4345
Cerebral Atrophy ("No" as reference)	0.1208	0.183	0.4358	0.5092	1.1284	0.7883 - 1.6153
Education_1 (Reference Group=0)	1.4561	0.1969	54.7097	0.0000*	4.2893	2.9162 - 6.3090

Variable	β	SE	Wald χ^2	P	OR	95% CI for OR
Education_2 (Reference Group=0)	2.6356	0.2039	167.0776	0.0000*	13.9523	9.3558 - 20.8072

Note: 0 = illiterate, 1 = primary school, and 2 = junior high school and above. The logistic multivariate analysis identified variables associated with GCD: Age, MoCA score, Hyperglycemia, Cerebral Hemorrhage, Cerebral Infarction, Education_1, and Education_2 (*: $p < 0.05$).

S10 Table. Detailed results of LAASO regression selection

Rank	Feature	Coef	Abs(Coef)	Status
1	Education	0.930059	0.930059	Selected
2	MoCA score	-1.09043	1.09043	Selected
3	Cerebral infarction	-0.3563	0.3563	Selected
4	Cerebral atrophy	0.048226	0.048226	Selected
5	Age	-0.38966	0.389664	Selected
6	Hyperglycemia	-0.10841	0.108412	Selected
7	Cerebral hemorrhage	-0.13107	0.131072	Selected
8	Hypertension	-0.03943	0.039427	Selected
9	Smoking	0.064277	0.064277	Selected
10	Brain trauma	0.028677	0.028677	Selected
11	Marital status	0.011691	0.011691	Selected
12	Hyperlipidemia	0.000001	0.000001	Eliminated
13	Parkinson's	0	0	Eliminated
14	Depression	0	0	Eliminated
15	CHD	0	0	Eliminated
16	Epilepsy	0	0	Eliminated
17	Gender	0	0	Eliminated
18	Drinking	0	0	Eliminated

The coefficient of the characteristic parameter is not zero. Abbreviations: CHD: Coronary heart disease, MoCA: The Montreal Cognitive Assessment (Beijing Edition).

S11 Table. The multicollinearity of the independent variable in the derivation cohort by the variance inflation factor (VIF) method.

Rank	Feature	VIF
1	MoCA score	1.3031
2	Education	1.1139
3	Age	1.3297
4	Cerebral Infarction	1.1106
5	Cerebral Hemorrhage	1.0387
6	Hyperglycemia	1.0489
7	Smoking	1.0115
8	Cerebral Atrophy	1.0569
9	Hypertension	1.0779

Rank	Feature	VIF
10	Brain Trauma	1.0133
11	Marital Status	1.1459

Abbreviations : CHD : Coronary heart disease, MoCA: The Montreal Cognitive Assessment (Beijing Edition).

S12 Table. Comparison between the optimal XGB model based on MoCA score, sociodemographic and health-related factors and the existing cognitive risk scoring tools

Models	Optimal model	Kappa	Sensitivity	Specificity	Accuracy	PPV	NPV	F1-score	Youden index	AUC(95%CI)
A	XGB	0.2029	1.0000	0.3465	0.5033	0.3258	1	0.4915	0.3465	0.6859(0.6333~0.7384)
B	GBC	0.3645	0.9821	0.4638	0.6458	0.4977	0.9796	0.6607	0.4459	0.7911(0.7464~0.8357)
C	XGB	0.6870	0.9657	0.7905	0.8480	0.6926	0.9792	0.8067	0.7562	0.9349(0.9139~0.9558)

Note: A:Modeled with ADL. B:Model with MoCA (Beijing Edition). C:The model XGB established based on MoCA score, sociodemographic and health-related factors. Abbreviations: ADL:activity of daily living, MoCA:Montreal Cognitive Assessment (Beijing Edition), XGB:eXtreme Gradient Boosting, GBC:Gradient Boosting Classifier.

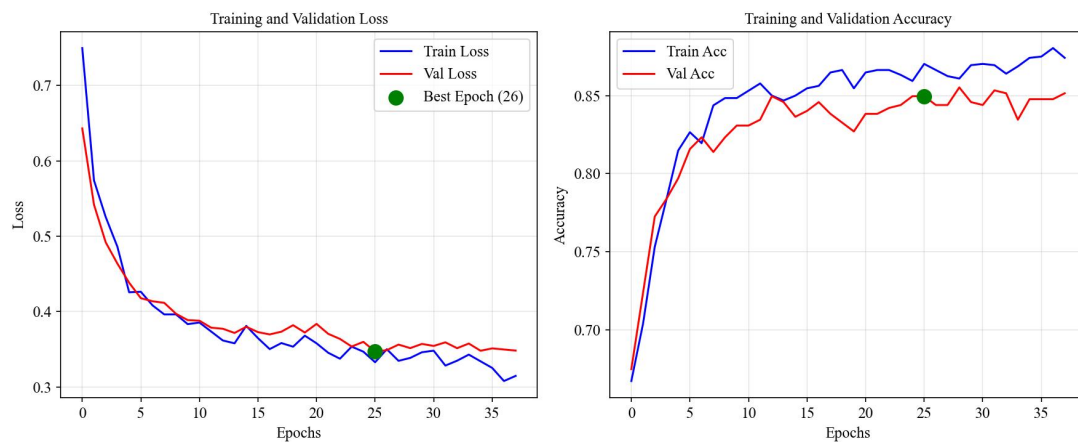
S13 Table. The result of the Delong test

Model	P value
XGB(Main research) vs XGB (ADL)	<0.01
XGB(Main research) vs GBC(MoCA)	<0.01

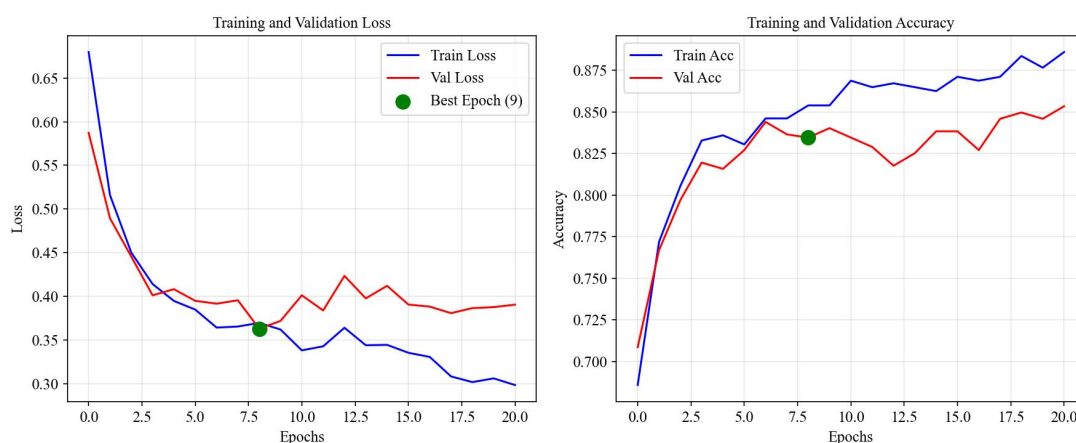
Abbreviations: XGB:eXtreme Gradient Boosting, GBC:Gradient Boosting Classifier, ADL:activity of daily living, MoCA: Montreal Cognitive Assessment (Beijing Edition).

S1 Fig. The performance curve changes of the training sets and validation sets of the three types of deep learning models

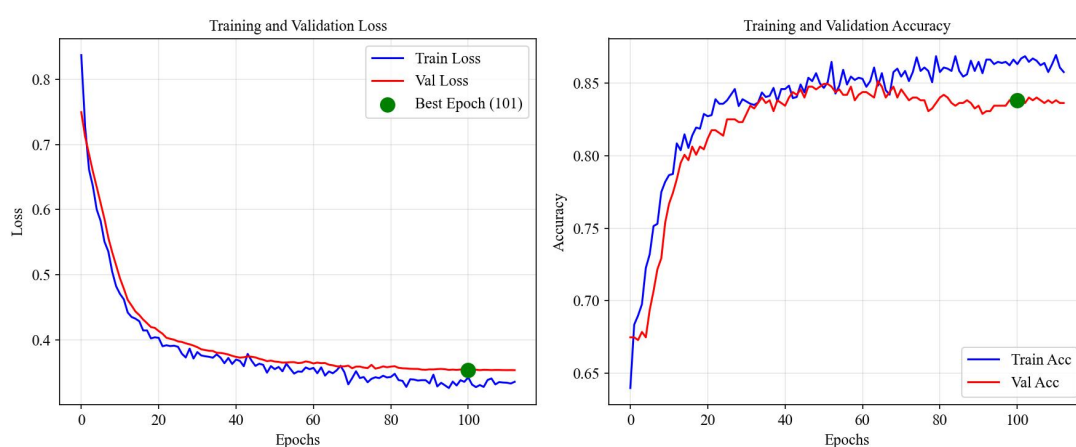
A



B

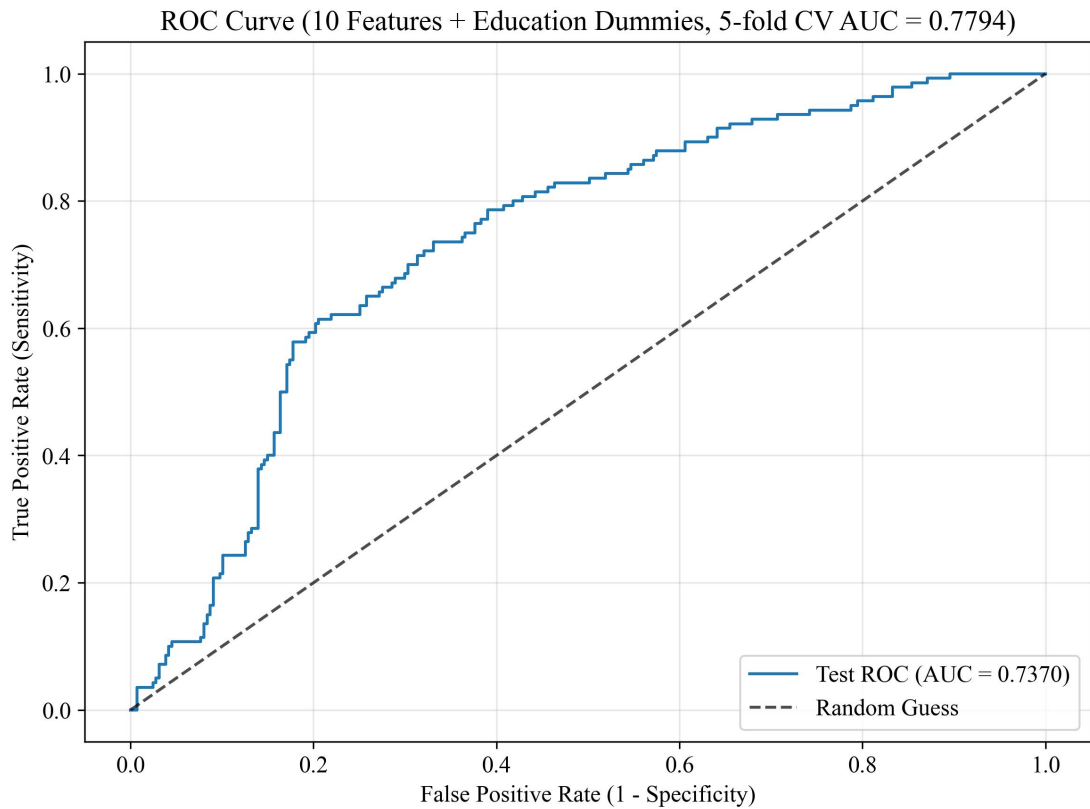


C

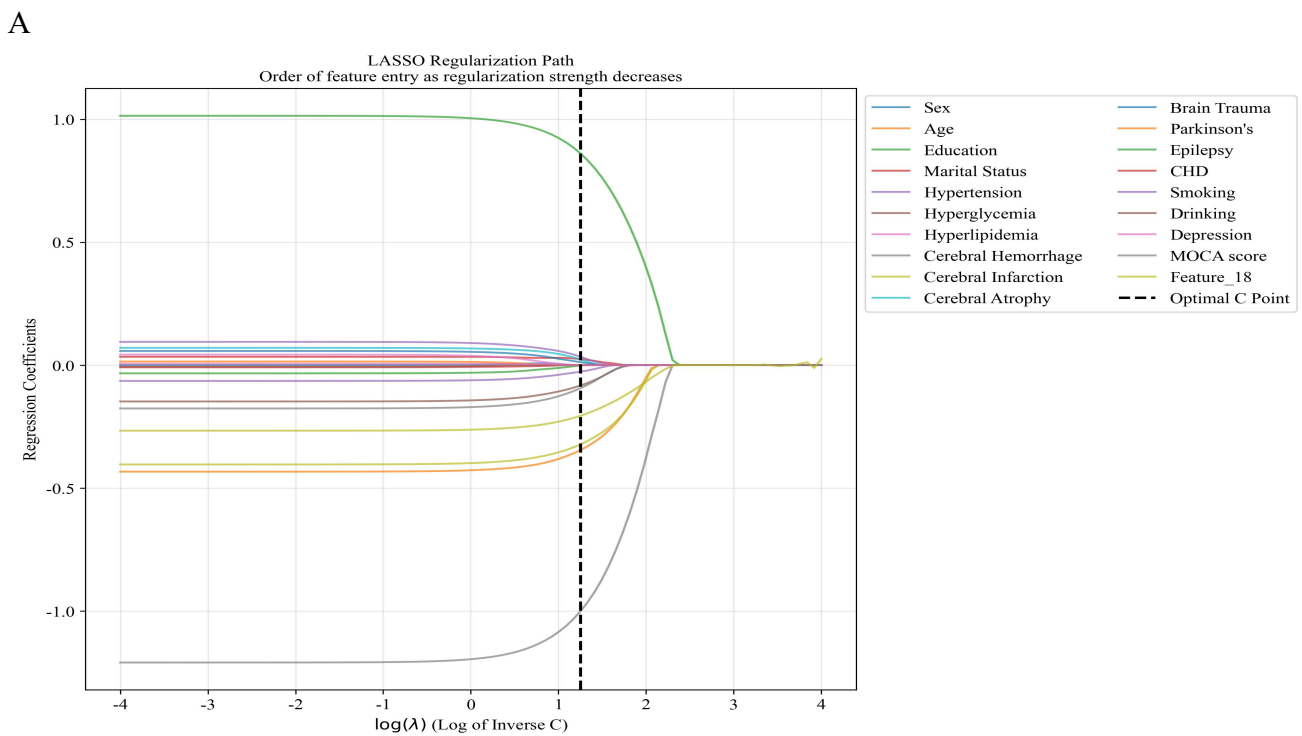


Note: The X-axis shows the number of training sessions. The left panel illustrates the variations in the loss functions for both the training and validation sets. Conversely, the right panel displays the changes in accuracy rates for these same sets. The optimal number of training sessions was identified based on the insights gained from these two graphs. A: Deep-Learning Neural Network (DNN). B: Residual Network (ResNet). C: Shallow Neural Network (SNN).

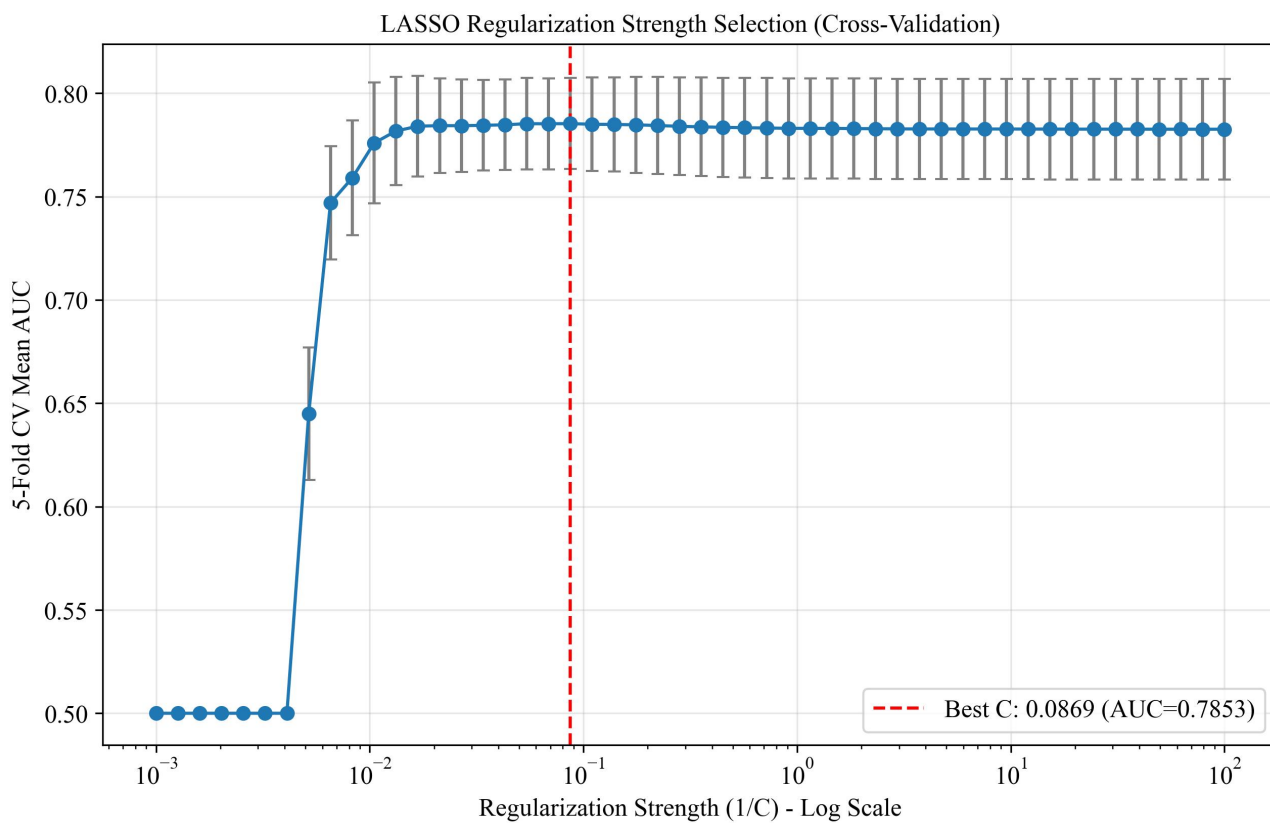
S2 Fig. The area under the ROC curve (AUC) of the derived cohort multivariate logistic regression analysis



S3 Fig. Feature selection



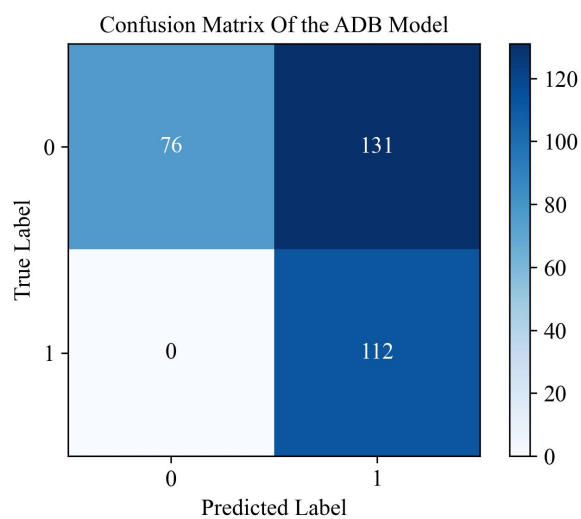
B



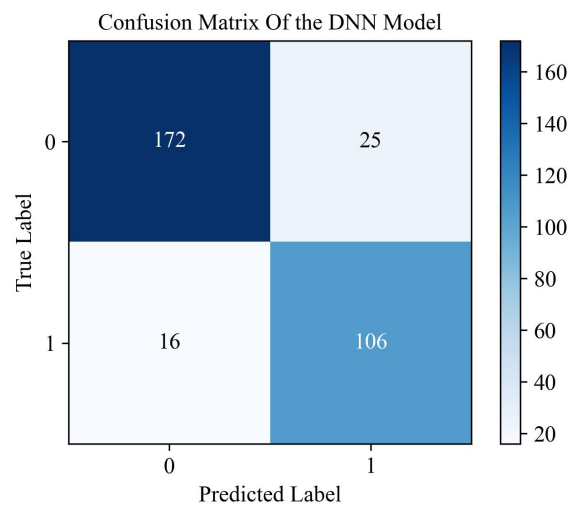
Note: A: LASSO Regularization Path order of feature entry as Regularization strength decreases, B: LASSO Regularization Strength selection (Cross-Validation).

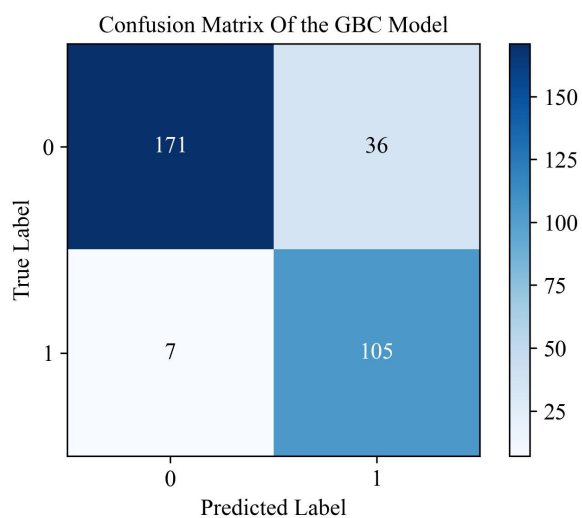
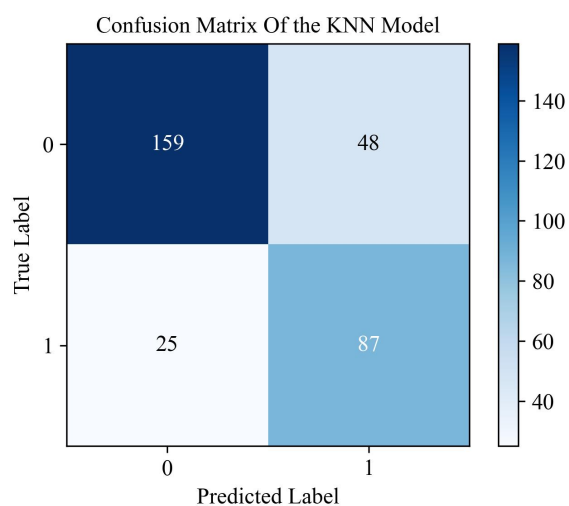
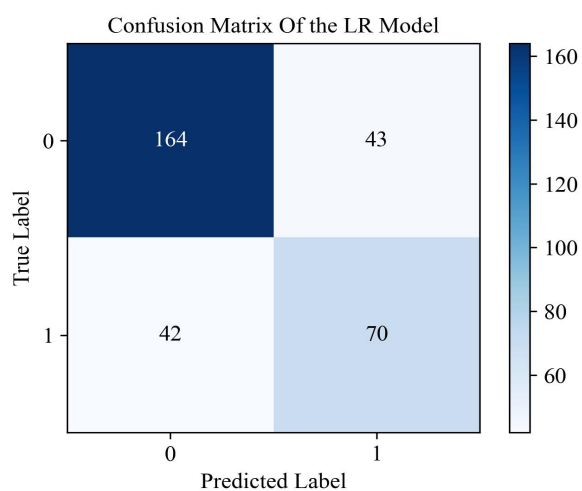
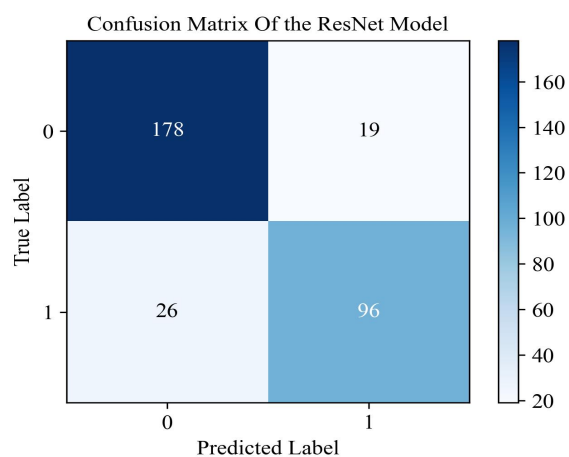
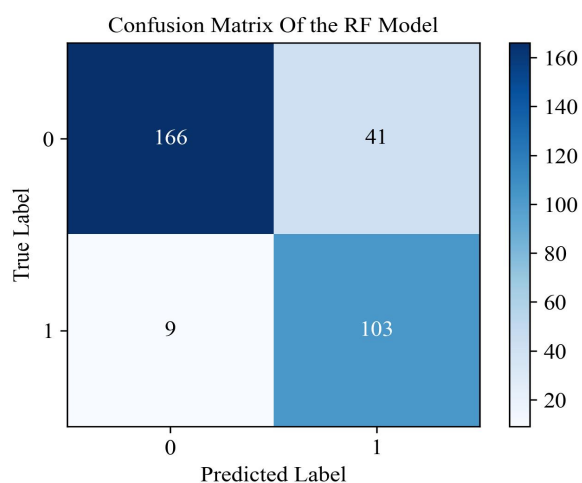
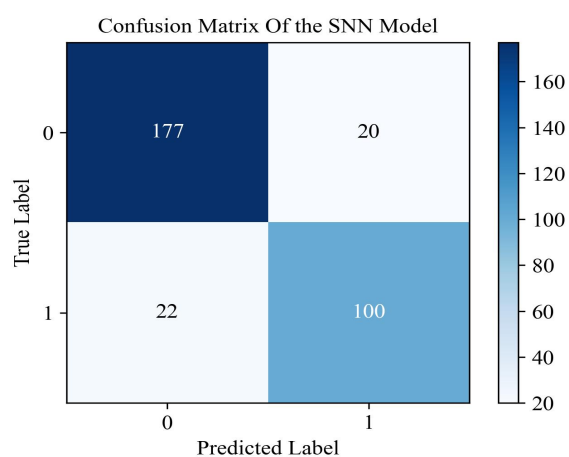
S4 Fig. Confusion matrix heatmap of the nine ML models for GCD prediction in the internal validation cohort.

A

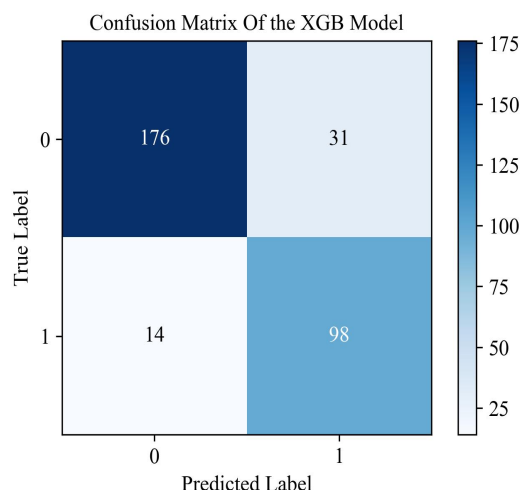


B



C**D****E****F****G****H**

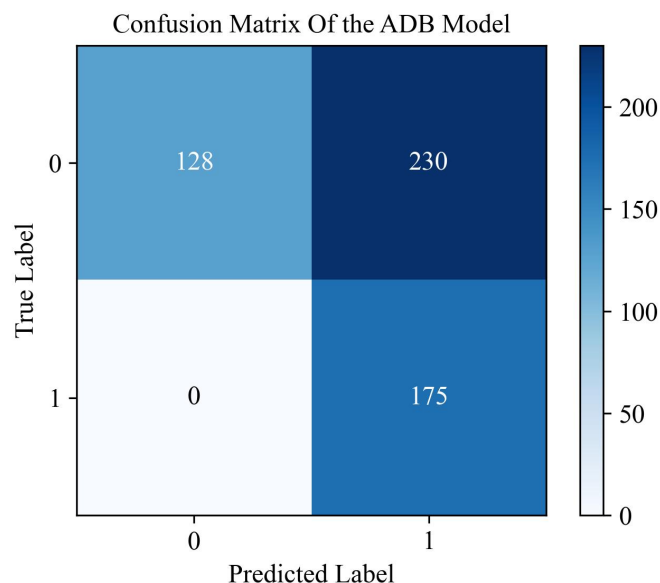
I



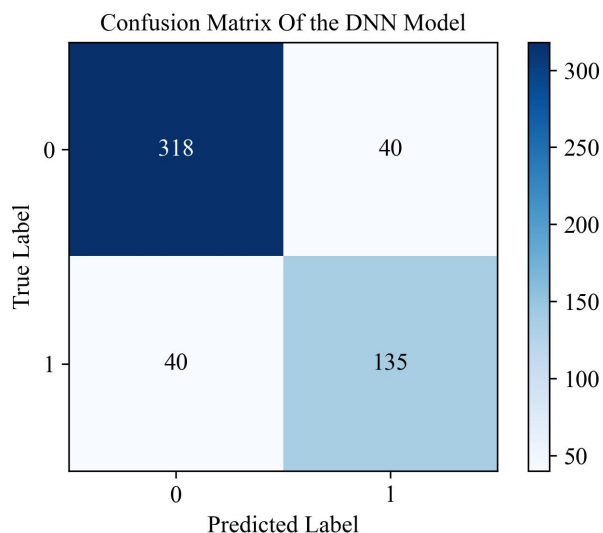
Abbreviations: GCD: Geriatric Cognitive Disorders. XGB: eXtreme Gradient Boosting, LR: Logistic regression classifier, RF: Random forest classifier, ADB: Adaptive Boosting, GBC: Gradient Boosting Classifier, KNN: K-Nearest Neighbor, SNN: Shallow Neural Network, DNN: Deep Neural Network, ResNet: Residual Neural Network.

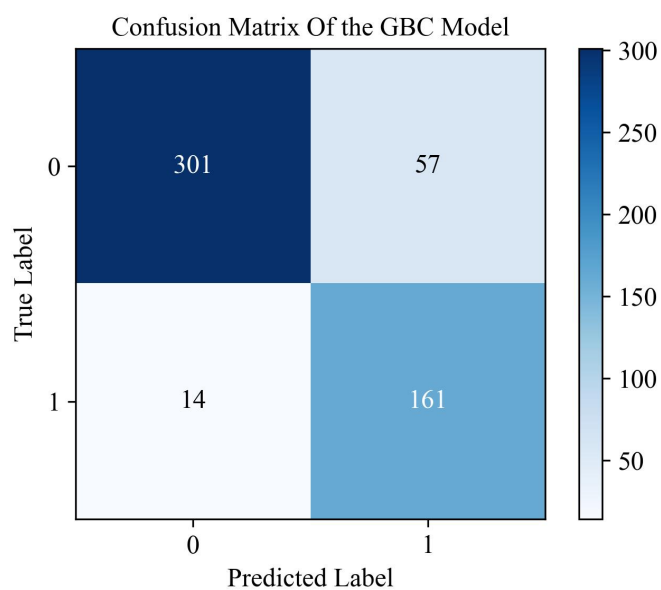
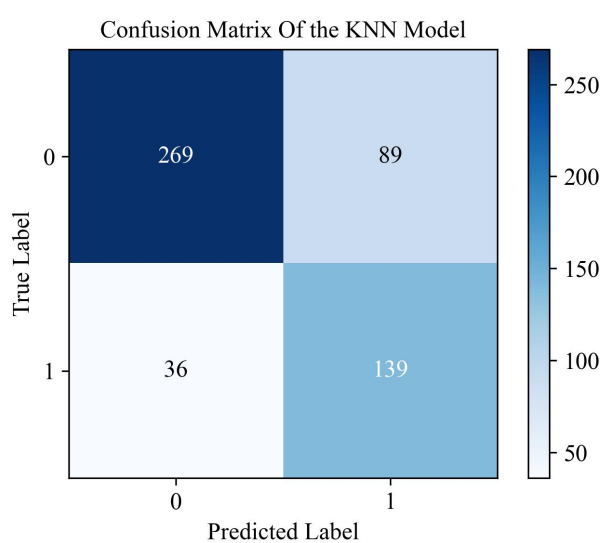
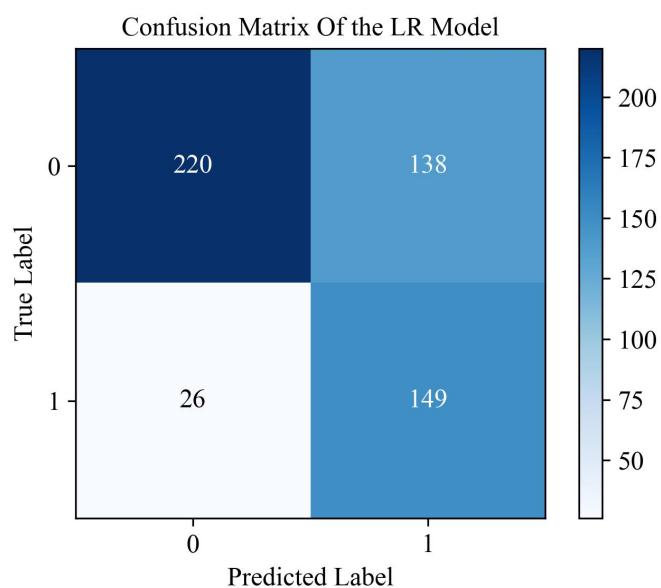
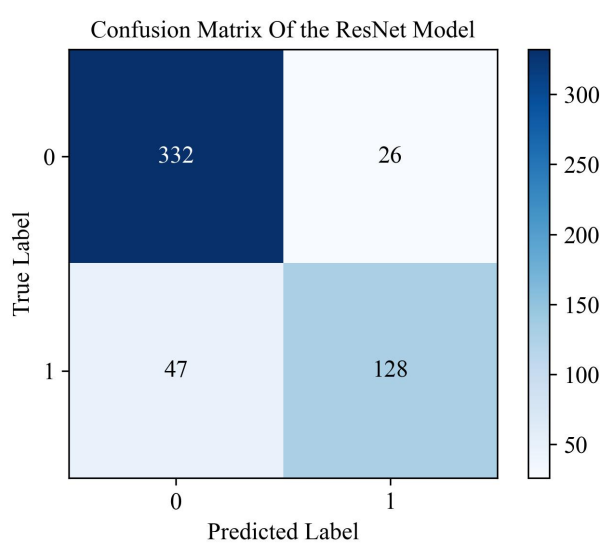
S5 Fig. Confusion matrix heatmap of the nine ML models for GCD prediction in the temporal external validation cohort.

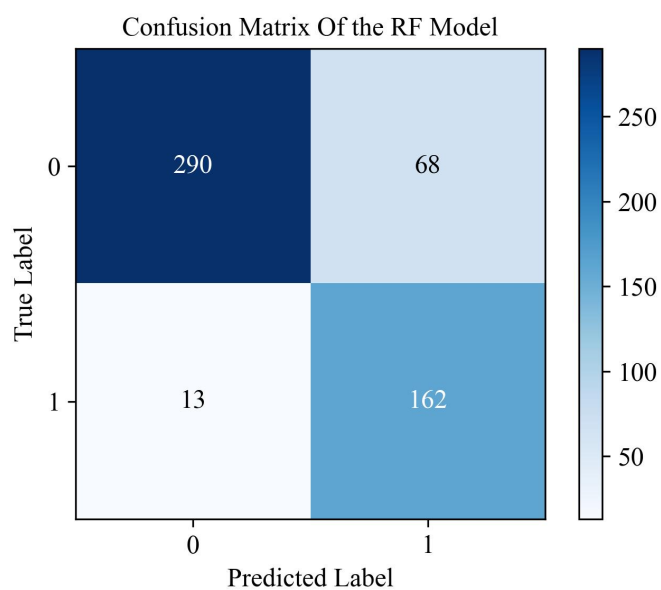
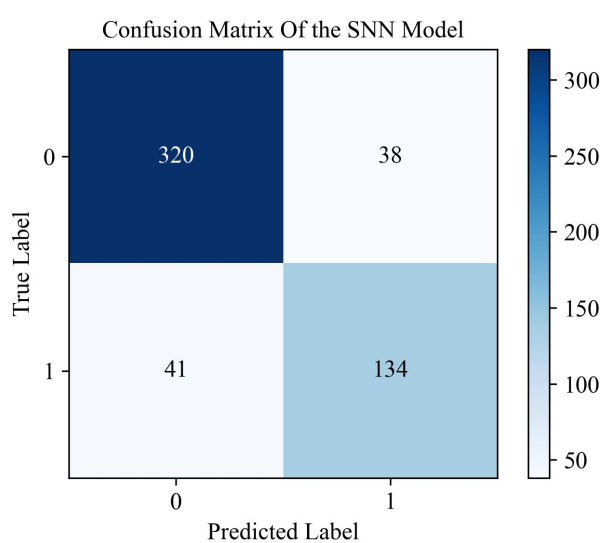
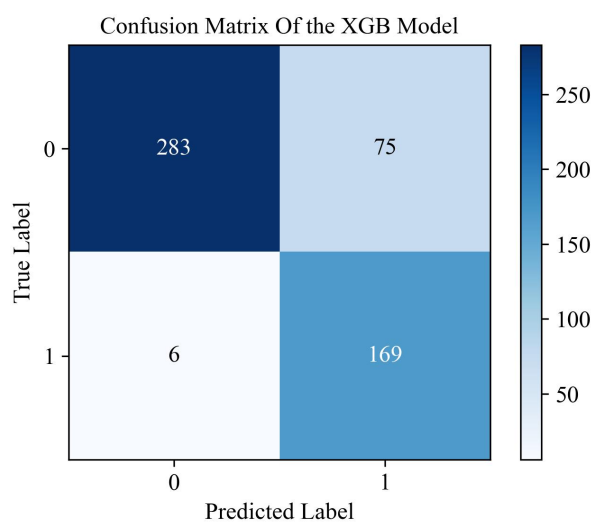
A



B



C**D****E****F**

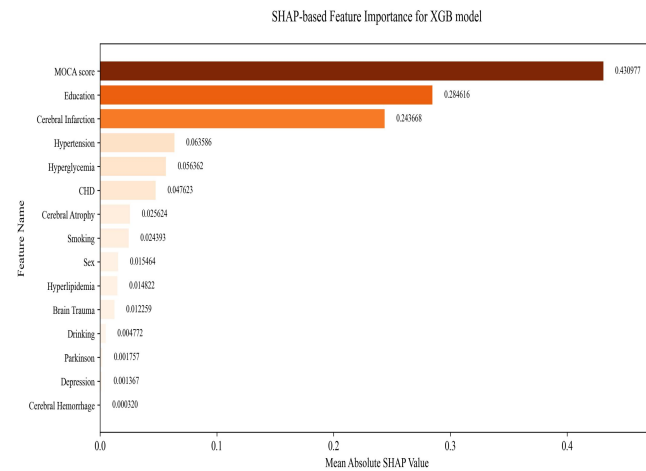
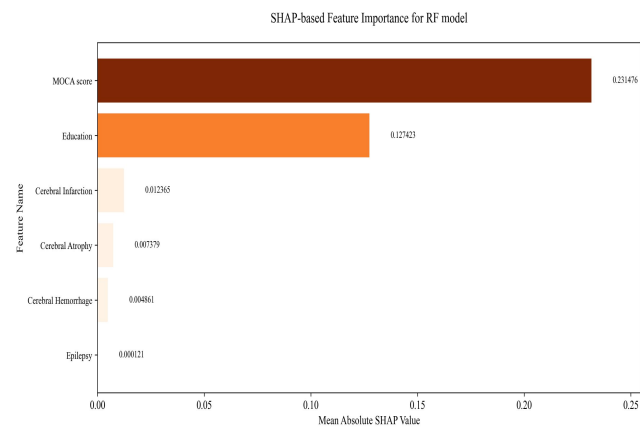
G**H****I**

Abbreviations: GCD: Geriatric Cognitive Disorders. XGB: eXtreme Gradient Boosting, LR: Logistic regression classifier, RF: Random forest classifier, ADB: Adaptive Boosting, GBC: Gradient Boosting Classifier, KNN: K-Nearest Neighbor, SNN: Shallow Neural Network, DNN: Deep Neural Network, ResNet: Residual Neural Network.

S6 Fig. The SHAP feature importance plot of the five subgroups.

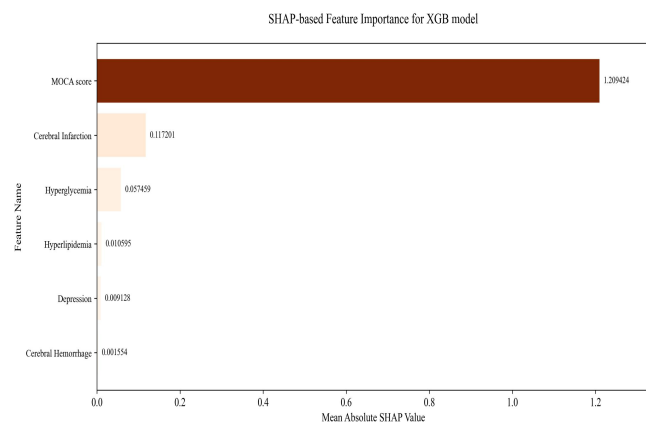
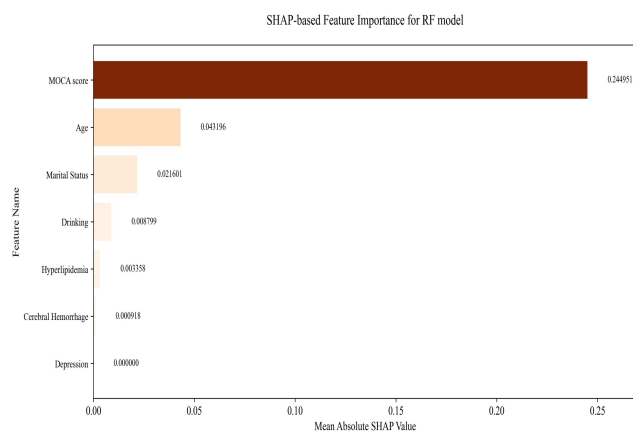
(B) ≥ 75 years

(A) 60–74 years

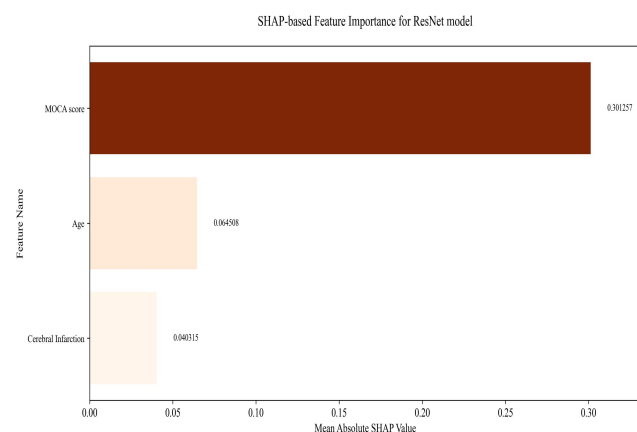


(C) Illiterate

(D) Primary school



(E) Junior high school or above

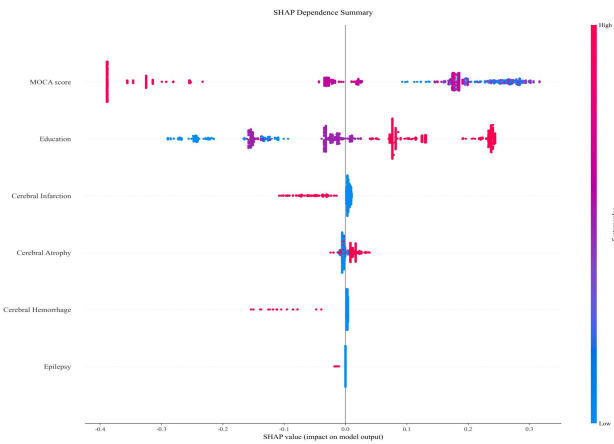


Abbreviations: XGB: eXtreme Gradient Boosting, RF: Random forest classifier, ResNet: Residual Neural Network. MoCA: Montreal Cognitive Assessment (Beijing

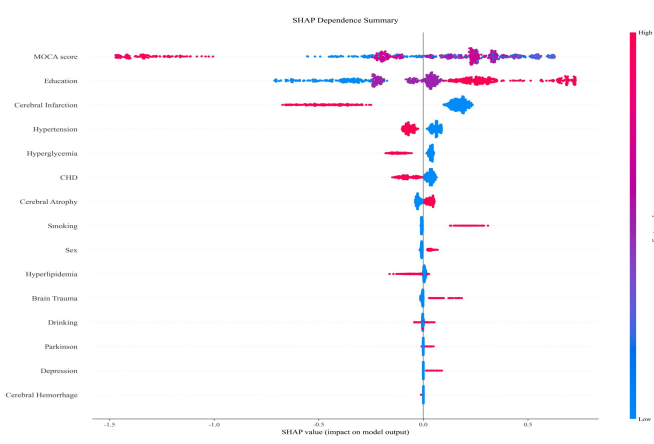
Edition), CHD: Coronary Heart Disease.

S7 Fig. The SHAP summary dot plot of the five subgroups.

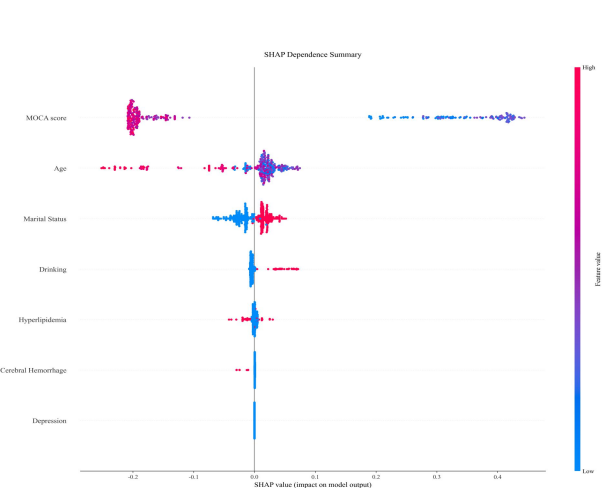
(A) 60–74 years



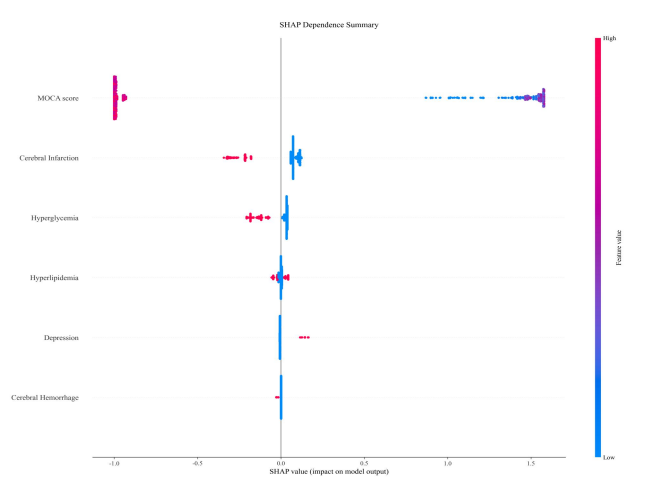
(B) ≥ 75 years



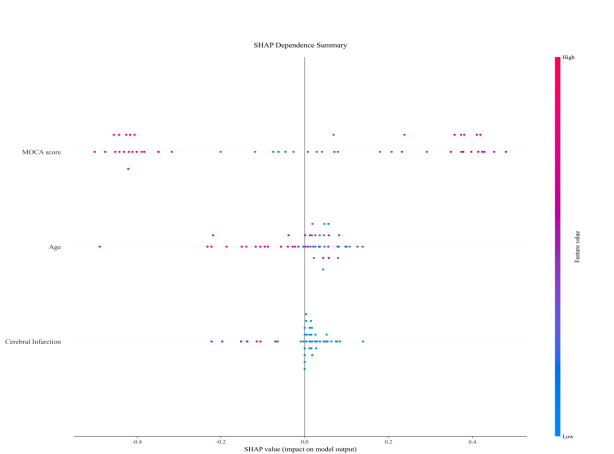
(C) Illiterate



(D) Primary school



(E) Junior high school or above



Abbreviations: MoCA: Montreal Cognitive Assessment (Beijing Edition), CHD: Coronary Heart Disease.