



OPEN ACCESS

EDITED BY

Eliza Russu,
George Emil Palade University of Medicine,
Pharmacy, Sciences and Technology of Târgu
Mureș, Romania

REVIEWED BY

Niranjana Sampathila,
Manipal Academy of Higher Education, India
Stoian Adina,
George Emil Palade University of Medicine,
Pharmacy, Sciences and Technology of Târgu
Mureș, Romania

*CORRESPONDENCE

Fang Tang

✉ tangfangsdu@gmail.com

RECEIVED 23 February 2024

ACCEPTED 09 July 2024

PUBLISHED 23 July 2024

CITATION

Zhang X, Tang C, Wang S, Liu W, Yang W,
Wang D, Wang Q and Tang F (2024) A
stacking ensemble model for predicting the
occurrence of carotid atherosclerosis.
Front. Endocrinol. 15:1390352.
doi: 10.3389/fendo.2024.1390352

COPYRIGHT

© 2024 Zhang, Tang, Wang, Liu, Yang, Wang,
Wang and Tang. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

A stacking ensemble model for predicting the occurrence of carotid atherosclerosis

Xiaoshuai Zhang¹, Chuanping Tang¹, Shuohuan Wang²,
Wei Liu³, Wangxuan Yang⁴, Di Wang¹, Qinghuan Wang¹
and Fang Tang^{3,5,6*}

¹Department of Data Science, School of Statistics and Mathematics, Shandong University of Finance and Economics, Jinan, China, ²Information Technology Division, Shandong International Trust Co., Ltd., Jinan, China, ³Department of Medical Ultrasound, The First Affiliated Hospital of Shandong First Medical University & Shandong Provincial Qianfoshan Hospital, Shandong Engineering Research Center of Diagnosis and Treatment Technology for Bariatric and Metabolism-Associated Surgery, Jinan, China, ⁴School of Public Health, Harbin Medical University, Harbin, China, ⁵Center for Big Data Research in Health and Medicine, The First Affiliated Hospital of Shandong First Medical University & Shandong Provincial Qianfoshan Hospital Qianfoshan Data Open Innovative Application Laboratory, Jinan, China, ⁶Shandong Provincial Qianfoshan Hospital, Cheeloo College of Medicine, Shandong University, Jinan, China

Background: Carotid atherosclerosis (CAS) is a significant risk factor for cardio-cerebrovascular events. The objective of this study is to employ stacking ensemble machine learning techniques to enhance the prediction of CAS occurrence, incorporating a wide range of predictors, including endocrine-related markers.

Methods: Based on data from a routine health check-up cohort, five individual prediction models for CAS were established based on logistic regression (LR), random forest (RF), support vector machine (SVM), extreme gradient boosting (XGBoost) and gradient boosting decision tree (GBDT) methods. Then, a stacking ensemble algorithm was used to integrate the base models to improve the prediction ability and address overfitting problems. Finally, the SHAP value method was applied for an in-depth analysis of variable importance at both the overall and individual levels, with a focus on elucidating the impact of endocrine-related variables.

Results: A total of 441 of the 1669 subjects in the cohort were finally diagnosed with CAS. Seventeen variables were selected as predictors. The ensemble model outperformed the individual models, with AUCs of 0.893 in the testing set and 0.861 in the validation set. The ensemble model has the optimal accuracy, precision, recall and F1 score in the validation set, with considerable performance in the testing set. Carotid stenosis and age emerged as the most significant predictors, alongside notable contributions from endocrine-related factors.

Conclusion: The ensemble model shows enhanced accuracy and generalizability in predicting CAS risk, underscoring its utility in identifying individuals at high risk.

This approach integrates a comprehensive analysis of predictors, including endocrine markers, affirming the critical role of endocrine dysfunctions in CAS development. It represents a promising tool in identifying high-risk individuals for the prevention of CAS and cardio-cerebrovascular diseases.

KEYWORDS

carotid atherosclerosis, endocrine-related markers, prediction, stacking, machine learning

1 Introduction

Carotid atherosclerosis (CAS) is a multifaceted disease characterized by the progressive accumulation of atherosclerotic plaques within the carotid arteries (1). As a manifestation of atherosclerosis in local blood vessels, the continuous development of CAS is a major and potentially preventable cause of ischaemic stroke (2). Early manifestations of CAS such as intermittent dizziness or mild headaches are subtle, often leading to missed diagnoses. As CAS progresses, it severely impacts the physical and psychological well-being of individuals, imposing substantial financial strains on their families. Therefore, early prediction and prevention of CAS are crucial to mitigate the risk of subsequent cardio-cerebrovascular events.

Current research on CAS has mainly focused on the analysis of risk factors, the most common of which include age, smoking status, physical inactivity, abnormal blood glucose levels, hypertension and others (3–7). These factors, particularly hyperglycemia and hypertension, indicative of underlying metabolic and hormonal imbalances, contribute to the endothelial dysfunction, inflammation, and subsequent plaque formation characteristic of atherosclerosis. Despite numerous studies on CAS risk factors, there is a scarcity of research dedicated to developing predictive models for CAS, with existing models primarily using cross-sectional data for disease diagnosis rather than predicting its onset.

Machine learning methods offer the potential to achieve precise predictive ability to assess diagnostic and prognostic outcomes (8–11). Among various machine learning approaches, ensemble learning, which includes techniques like bagging (e.g., random forests), boosting (e.g., XGBoost, GBDT), and stacking, stands out by integrating multiple weak classifiers to form a robust classifier, thereby improving prediction accuracy and model generalizability (12–15). Stacking ensemble models, which train different weak learners in parallel, have shown superior performance across various domains, from healthcare to financial forecasting (16–20). However, they also present significant challenges such as computational complexity and a lack of interpretability, often referred to as the “black box” phenomenon, which can obscure understanding of decision-making processes (21).

In this study, we employ a stacking ensemble learning algorithm to construct a risk prediction model for the occurrence of CAS based on a routine health checkup cohort. The predictive performance of the ensemble model was compared with that of the individual models. We utilize the SHapley Additive exPlanation (SHAP) method (22) to elucidate the predictive relationships between CAS and various risk factors, with a particular focus on endocrine-related markers.

2 Materials and methods

2.1 Study design and data collection

The study cohort was derived from the routine health check-up system of the First Affiliated Hospital of Shandong First Medical University in Jinan, China. All the participants were free of CAS at the first check-up and underwent three health checks during the follow-up. Individuals who had been diagnosed with coronary heart disease, previous coronary heart disease, cerebral ischaemia, cerebral infarction, cerebral artery stenosis, cerebral artery spasm, coronary artery stenosis, coronary atherosclerotic heart disease, and those with missing information were excluded. CAS was diagnosed by carotid B-mode ultrasonography as a carotid intima-media thickness of 1.0 mm or greater or plaque formation. The study was approved by the Ethics Committee of the First Affiliated Hospital of Shandong First Medical University, and informed consent was obtained from all eligible participants.

2.2 Study variables

The study variables consisted of three sets of data: demographic data, laboratory indicators, and clinical history. All the individuals in this study cohort underwent anthropometric and laboratory tests. The height and weight of the participants were measured while they were wearing light clothing and no shoes. Peripheral blood samples were collected from the subjects after an overnight fast, and the variables included blood urea nitrogen (BUN), lymphocyte percentage (LYM), aspartate aminotransferase (AST), red cell volume distribution width standard deviation (RDW-SD), red

blood cell count (RBC), mean corpuscular haemoglobin concentration (MCHC), mean platelets (MPV), fasting blood glucose (GLU), platelet count (PLT), eosinophil percent (PEOS), white blood cell count (WBC), and carcinoembryonic antigen (CEA). Disease history was also collected, such as carotid stenosis (CS), diabetes mellitus (DM), and hypertension. All the measurements were collected following the same standard procedures.

2.3 SMOTE sampling

To address the issue of data imbalance, we applied the synthetic minority oversampling technique (SMOTE) in our study. Since the number of individuals without CAS was larger than those with CAS, SMOTE was employed to generate synthetic samples of the minority class (22). The new synthetic records were generated using the existing samples of the minority class via linear interpolation. After we obtain new minority sample data, a balanced dataset can be obtained by merging with majority samples.

2.4 Variable selection

Variable selection is an important step in the application of machine learning to ensure that the most relevant predictors are used. Both filtering and embedded feature selection methods were used to select the predictors. The variables were first selected using univariate logical regression with a threshold P value of 0.1. Second, we applied three tree-based machine learning methods-random forest (RF), eXtreme Gradient Boost (XGBoost), and gradient boosting decision tree (GBDT))-to assess the importance of each variable. These methods are well-suited for identifying important variables due to their ability to capture complex interactions and non-linear relationships. Variables were ranked based on their importance scores from these models. Finally, the correlation coefficients of the

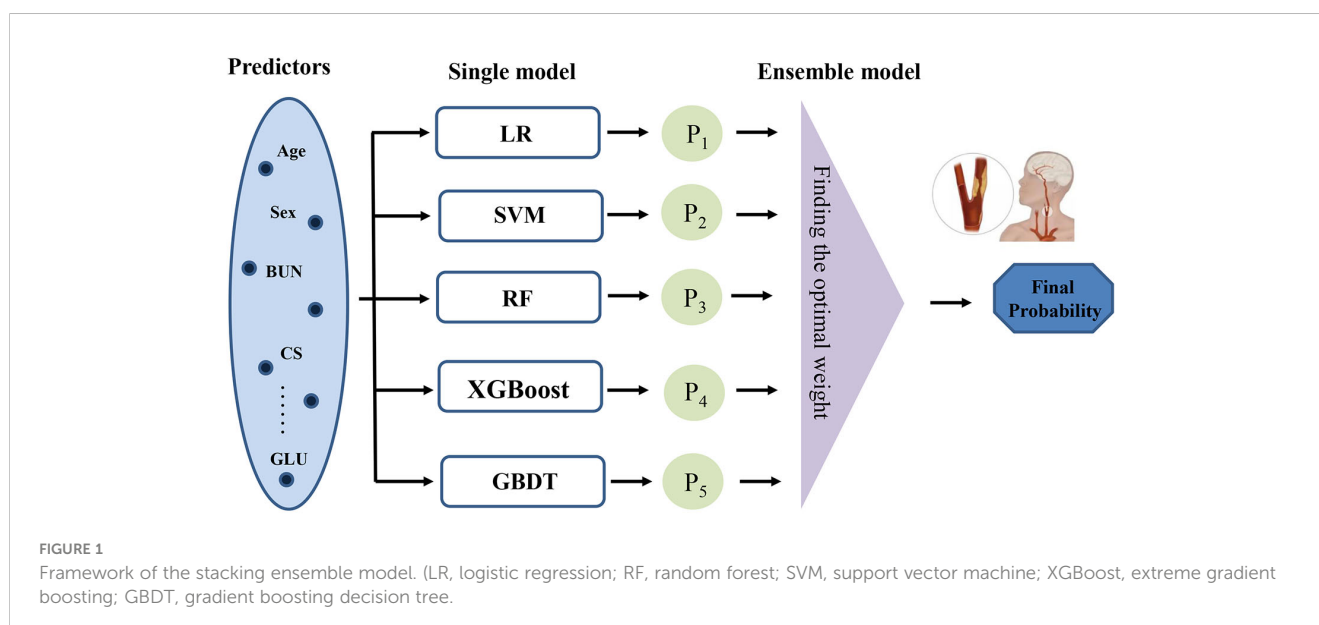
continuous variables were calculated to address multicollinearity, which can distort the model's estimates and reduce interpretability. Features with low importance among the highly related variables were eliminated for determining the predictors.

2.5 Statistical analysis

The baseline characteristics were assessed for CAS and non-CAS patients during the follow-up. Continuous variables were described by the mean and standard deviation (SD), and categorical features were described as proportions; we compared the baseline features using the t test and the chi-square test. To predict the probability of CAS, we employed five machine learning models: logistic regression (LR), support vector machine (SVM), RF, XGBoost, and GBDT (11, 23–26). We then used a stacking ensemble model, specifically the super learner, which combines these individual models by assigning them different weights to optimize predictive performance (27). The final predicted value is a weighted sum of the individual model predictions, where the weights are determined to minimize the cross-validation risk (28). Figure 1 shows the framework of the super learner ensemble model. We compared the predictive performance of the super learner against the individual models (LR, RF, SVM, XGBoost, and GBDT).

2.6 Hyperparameter tuning

We optimized the hyperparameters using a random search method, which is a traditional and efficient technique for tuning in classification methods (29). This process was conducted within a 5-fold cross-validation framework. Specifically, each model configuration was trained on four folds and validated on the remaining fold. This cycle was repeated five times, with each fold serving as the validation set once, to ensure a comprehensive evaluation across the entire dataset.



We performed the random search 1,000 times, selecting the hyperparameter combination with the highest average areas under the receiver operating characteristic curve (AUC).

2.7 Predictive performance assessment

The performance of the prediction model was validated using both the testing and hold-out validation set. Several metrics were used to evaluate the performance of the prediction models: accuracy, precision, recall, F1 score, and AUC. Compared to commonly used performance metrics, the AUC better reflects model performance in unbalanced datasets. Hence, the AUC was the main metric, while the others were considered of secondary priority.

2.8 Model interpretation

To solve the “black box” problem in machine learning, we report the feature importance ranking of each predictor based on SHAP values (22). SHAP values are useful for explaining the prediction of a machine learning model by computing the contribution of each feature to the prediction. Kernel-based SHAP values were used to rank the variables in terms of their ability to predict the CAS, which is an additive feature attribution method using kernel functions enabling consistent explanation of feature importance (30).

3 Results

3.1 Data description

A total of 1669 participants were included in this study, including 1426 (85.4%) males and 243 (14.6%) females. A total of 441 participants were diagnosed with CAS during the follow-up, including 395 men and 46 women. A total of 1228 participants were not diagnosed with CAS. The SMOTE method was used to address the sample imbalance problem. Figure 2 shows the roadmap

of the data processing. All the samples were first divided into a training set and a hold-out validation set. Specifically, 5% of the subjects were randomly selected in advance as the validation set, and the remaining 95% were used for model construction; 419 patients with CAS and 1167 non-CAS patients were included. Of the remaining data, 70% of the data were used as the training set where SMOTE resampling was applied to address class imbalance. The remaining 30% was used as the testing set. In the original dataset, the ratio of non-CAS to CAS was approximately 2.78 to 1, which was adjusted in the training set to approximately 1 to 1.

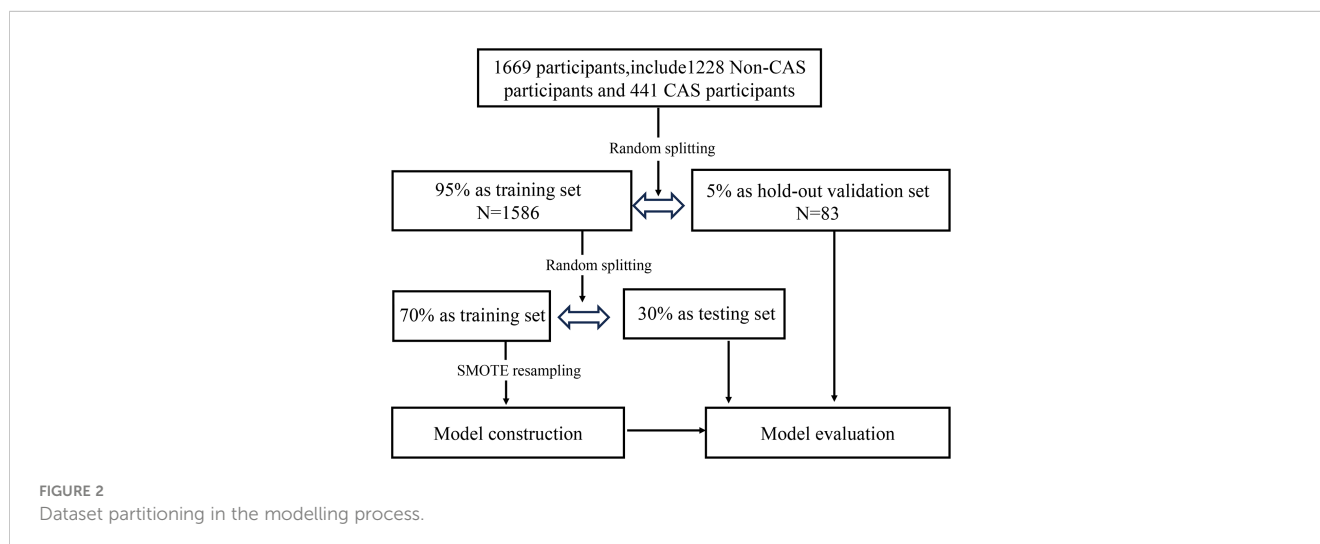
3.2 Variable selection results

A total of 28 variables with a P value < 0.1 were retained in the univariate logistic regression and subsequently included in the three machine learning models. Supplementary Figure S1 shows the variable importance rankings for RF, GBDT, and XGBoost. A total of 22 of these variables were common among all three machine learning algorithms (please see the Supplementary File). The correlations between the selected continuous variables are shown in Supplementary Figure S2 in the Supplementary File. Variables with high correlations and the lowest importance were removed; thus, five variables were excluded. Finally, 17 variables were selected, including age, sex, BUN, LYM, AST, RDW-SD, RBC, MCHC, MPV, GLU, PLT, PEOS, WBC, CEA, CS, hypertension and DM.

Table 1 summarizes the baseline characteristics of the incident CAS status. Overall, individuals who developed CAS were more likely to be male CS, DM, hypertension, older age, MCHC, BUN, RDW-SD, MPV, GLU, PEOS, WBC, and CEA and lower PLT, LYM, AST, and RBC at baseline; these variables were significantly different at a level of 0.1.

3.3 Model comparison

The super learner algorithm creates an optimal weighted average of the five models (LR, RF, SVM, XGBoost, and GBDT).



Supplementary Table S1 in the Supplementary File depicts the weight coefficients of the super learner model. Supplementary Table S2 shows the hyperparameters used in the models. The weight coefficients of LR and SVM are 0, indicating that they were not used in the super learner model, while the coefficient of RF is 0.820, which is much greater than that of the other four models, indicating that RF contributes most in the prediction model.

Table 2 shows the predictive performance of the six models on the testing set and the hold-out validation set. It can be seen that the predictive performance varies across the five models. SVM has the highest precision, while its performance in the validation set is inferior. Logistic regression had the lowest performance in the testing set. These results also indicate the reason that the two methods are not selected in the super learner model. Combined with the advantages of RF, XGBoost and GBDT, the performance of the super learner was improved. The super learner model had the

optimal performance measures in the validation set, with considerable performance in the testing set. The ROC curves of the different machine learning models are shown in Supplementary Figure S3. The super learner has the largest ROC curve area in the validation set, and its AUC is 0.861. Overall, the predictive performance of the super learner model is superior to that of the other five models, especially regarding the overfitting problem.

3.4 Model interpretation

In this paper, the SHAP value was used to quantify the impact of each variable on the prediction of CAS, and the results are shown in Figure 3. Figure 3A shows the contribution of all the features to the prediction, which was sorted according to the average SHAP values. CS and age are the two most important predictors with the largest SHAP values, followed by RDW-SD, GLU and hypertension.

To further explain how each variable affects the occurrence of CAS, we illustrate two sample cases. Figures 3B, C depict the SHAP value of each variable for individuals I and II. The blue bars on the left (SHAP value less than 0) indicate variables that reduce the probability of the individual being predicted as CAS; the orange bars on the right (SHAP value greater than 0) indicate variables that increase the probability of the individual being predicted as CAS. Larger areas indicate greater impacts of that factor. For individual I, diabetes and an increase in glucose are the main reasons for the increased risk of CAS. Due to the absence of CS, relatively young age and other variables with negative impacts, the predicted probability of CAS for individuals is low. In contrast, for individual II, CS is the main reason for the increased risk of CAS, and most of the variables have a positive impact in predicting CAS. The probability of CAS for individual II is only slightly reduced by the absence of hypertension, diabetes and young age; thus, this individual is more likely to develop CAS in the future. Therefore, through the SHAP framework, we can directly determine the main causes for the increased individual probability of CAS; thus, corresponding interventions could be taken to reduce the risk.

TABLE 1 Baseline characteristics by incident CAS status.

Variables	N=1669	Non-CAS N=1228 (73.6%)	CAS N=441 (26.4%)	t/ χ^2	P
Age	55.1 ± 7.6	53.7 ± 6.2	59.1 ± 9.4	-11.28	<0.001
MCHC	339.4 ± 10.1	339.0 ± 10.4	340.3 ± 9.3	-2.29	0.022
PLT	221.1 ± 50.9	222.6 ± 50.2	216.9 ± 52.7	1.98	0.048
BUN	5.2 ± 1.3	5.2 ± 1.3	5.4 ± 1.3	-2.39	0.017
LYM	0.4 ± 0.1	0.4 ± 0.1	0.3 ± 0.1	3.29	0.001
AST	20.0 ± 8.0	20.4 ± 8.5	19.0 ± 6.0	3.72	<0.001
RDW-SD	42.4 ± 2.7	42.2 ± 2.7	42.8 ± 2.6	-3.63	<0.001
RBC	4.9 ± 0.4	4.9 ± 0.4	4.8 ± 0.4	2.46	0.014
MPV	10.3 ± 0.9	10.3 ± 0.9	10.4 ± 0.8	-2.52	0.012
GLU	5.7 ± 1.3	5.6 ± 1.1	5.9 ± 1.7	-4.43	<0.001
EOS	0.026 ± 0.022	0.025 ± 0.020	0.028 ± 0.026	-1.87	0.061
WBC	6.3 ± 1.6	6.2 ± 1.5	6.4 ± 1.7	-2.29	0.023
CEA	2.0 ± 1.4	1.9 ± 1.3	2.1 ± 1.6	-2.93	0.004
Sex: n (%)	1426				
0	(85.4%)	1031(84.0%)	395(89.6%)	8.21	0.004
1	243(14.6%)	197(16.0%)	46 (10.4%)		
CS: n (%)	1088			121.24	<0.001
0	(65.2%)	895(72.9%)	193(43.8%)		
1	581 (34.8%)	333(27.1%)	248(56.2%)		
DM: n (%)	1475			26.53	<0.001
0	(88.4%)	1115(90.8%)	360(81.6%)		
1	194(11.6%)	113(9.2%)	81(18.4%)		
Hypertension: n (%)	1150			34.34	<0.001
0	(68.9%)	895(72.9%)	255(57.8%)		
1	519(31.1%)	333(27.1%)	186(42.2%)		

TABLE 2 Predictive performance of the six machine learning models.

Models	Performance metrics ^a				
	accuracy	precision	recall	F1	AUC
LR	0.734(0.783)	0.732(0.563)	0.701 (0.818)	0.716 (0.667)	0.797 (0.857)
SVM	0.825(0.602)	0.793(0.359)	0.859 (0.636)	0.825 (0.459)	0.909 (0.639)
RF	0.805(0.759)	0.783(0.528)	0.822 (0.864)	0.802 (0.655)	0.891 (0.848)
XGBoost	0.777(0.723)	0.761(0.487)	0.780 (0.864)	0.770 (0.623)	0.872 (0.817)
GBDT	0.789(0.759)	0.771(0.529)	0.797 (0.818)	0.784 (0.642)	0.866 (0.828)
Super learner	0.795(0.795)	0.785(0.571)	0.788 (0.909)	0.786 (0.701)	0.893 (0.861)

^aThe predictive performance values for the hold-out validation set are shown in parentheses.

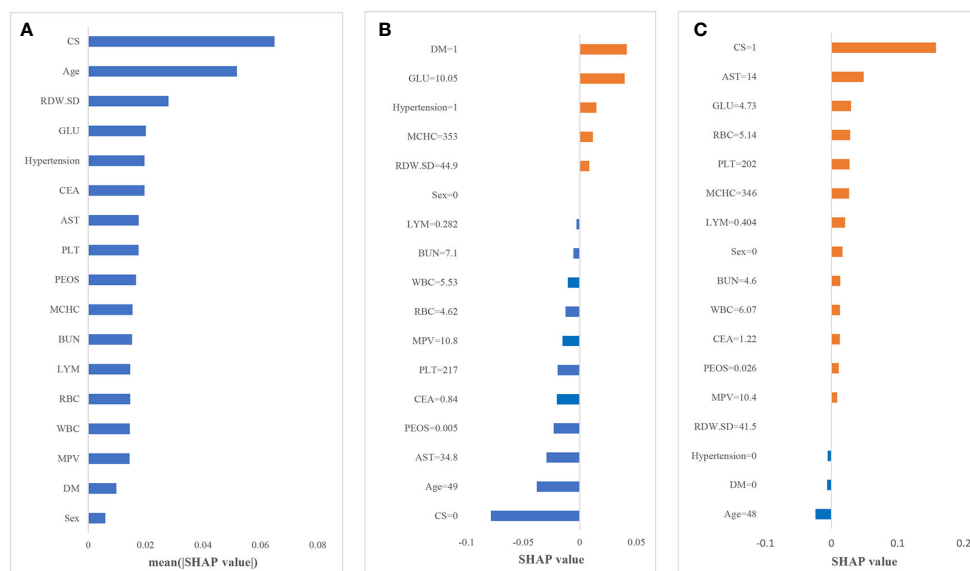


FIGURE 3

Results of the SHAP analysis. (A) Mean (|SHAP value|) of each variable; (B) the contribution of each variable in non-CAS individual I; (C) the contribution of each variable in CAS individual II.

4 Discussion

In this study, based on a routine health check-up cohort, we constructed a stacking ensemble prediction model for quantifying the risks of incident CAS. Demographic information, such as age and sex, and clinical factors, including BUN, LYM, AST, RDW-SD, RBC, MCHC, MPV, GLU, PLT, PEOS, WBC, CEA, CS, hypertension and DM, were important predictors of CAS.

We established five machine learning models to predict CAS and found that the performance of the individual models varied in the testing and validation sets. Most of the models performed better on the testing set and inferiorly on the hold-out validation set, indicating the overfitting problem. Therefore, we used the super learner algorithm to integrate the models, which significantly improved their performance. The super learner model not only demonstrated superior discrimination but also effectively managed overfitting, with AUC scores of 0.893 and 0.861 in the testing and validation sets, respectively. Our findings align with recent research that demonstrates the superior performance of stacking models in various biomedical applications (31–33). Studies like those conducted by Zhou have shown that stacking models provide enhanced accuracy in predicting diabetes which is consistent with our results (34).

In accordance with several studies (4, 5, 35), age was identified as a risk factor for CAS. According to the results of the feature importance analysis for the three machine learning models and the SHAP explanatory framework, we found that age and CS were the two most important factors affecting the occurrence of CAS. The demographic shift towards an aging population warrants increased societal attention, given the anticipated rise in CAS incidence (36). Additionally, since CS is a symptom of CAS, its presence is an important signal of CAS, and these two groups of people in particular need to take corresponding measures to prevent CAS. Moreover, our

analysis extends beyond conventional risk factors by incorporating endocrine-related markers within the predictive framework. The integration of these markers, including but not limited to abnormal blood glucose levels and hypertension, underscores the intricate relationship between endocrine dysfunctions and atherosclerosis. Through the SHAP framework, the contribution of each feature to the CAS risk is quantified, offering a personalized risk assessment. It underscores the necessity of a multifaceted risk assessment strategy that not only considers traditional factors like age and CS but also gives weight to the underlying endocrine dysfunctions contributing to the disease's pathogenesis.

The SHAP score becomes an invaluable tool for clinicians, enhancing the interpretability of machine learning predictions and enabling personalized preventive measures. For instance, older individuals with CS may benefit from increased screening and early intervention strategies, facilitating early detection and management of CAS. Similarly, for individuals with abnormal blood glucose or hypertension, personalized medical interventions including adjustments in medication and lifestyle changes such as diet and exercise could be advised based on their specific risk profiles. Regular monitoring of blood pressure and glucose levels can further aid early intervention and management, demonstrating the dynamic utility of predictive models in clinical settings.

One of the limitations of our study was that information about important risk factors for CAS, such as lifestyle, was not available. However, the model in our study still achieved acceptable performance without these predictors. Moreover, study subjects in the routine check-up cohort were limited to a single source area, and the prediction model was only internally validated. External validation with an independent population is needed to evaluate the generalizability of the model. Future studies will aim to collaborate with various institutions across different geographic regions to ensure that our models are robust and applicable to a broader

population. This approach will not only help in validating our current model externally but also in assessing its effectiveness across different demographic settings.

5 Conclusion

In this study, we developed a stacking model to predict the risk of incident CAS, enhancing the application of machine learning in the disease prediction. This approach not only provides a new method for the risk calculation of CAS but also highlights the critical role of endocrine dysfunctions in CAS development. By integrating a comprehensive analysis of predictors, and utilizing SHAP for model interpretation, our model effectively identifies high-risk individuals. This allows for targeted interventions that could substantially reduce the health and economic burdens associated with CAS. The study demonstrates the potential of advanced machine learning techniques to enhance preventive healthcare strategies.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors. Further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by the Ethics Committee of the First Affiliated Hospital of Shandong First Medical University, and informed consent was obtained from all the participants. The studies were conducted in accordance with the local legislation and institutional requirements.

Author contributions

XZ: Conceptualization, Formal analysis, Methodology, Writing – review & editing. CT: Formal analysis, Writing – original draft. SW:

Writing – review & editing. WL: Writing – review & editing. WY: Writing – original draft. DW: Writing – original draft. QW: Writing – original draft. FT: Conceptualization, Data curation, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The work was supported by National Natural Science Foundation of China (81903410&71804093).

Conflict of interest

Author SW were employed by Shandong International Trust Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fendo.2024.1390352/full#supplementary-material>

References

- Sirimarco G, Amarenco P, Labreuche J, Touboul P-J, Alberts M, Goto S, et al. Carotid atherosclerosis and risk of subsequent coronary event in outpatients with atherosclerosis. *Stroke*. (2013) 44:373–9. doi: 10.1161/STROKEAHA.112.673129
- Martinez E, Martorell J, Rimbau V. Review of serum biomarkers in carotid atherosclerosis. *J Vasc Surg*. (2020) 71:329–41. doi: 10.1016/j.jvs.2019.04.488
- Hollander M, Bots ML, Del Sol AI, Koudstaal PJ, Witteman JMC, Grobbee DE, et al. Carotid plaques increase the risk of stroke and subtypes of cerebral infarction in asymptomatic elderly: the Rotterdam study. *Circulation*. (2002) 105:2872–7. doi: 10.1161/01.CIR.0000018650.58984.75
- Taylor BA, Zaleski AL, Capizzi JA, Ballard KD, Troyanos C, Baggish AL, et al. Influence of chronic exercise on carotid atherosclerosis in marathon runners. *BMJ Open*. (2014) 4:e004498. doi: 10.1136/bmjopen-2013-004498
- van den Munckhof ICL, Jones H, Hopman MTE, de Graaf J, Nyakayiru J, van Dijk B, et al. Relation between age and carotid artery intima-medial thickness: a systematic review. *Clin Cardiol*. (2018) 41:698–704. doi: 10.1002/clc.22934
- Wu D, Li C, Chen Y, Xiong H, Tian X, Wu W, et al. Influence of blood pressure variability on early carotid atherosclerosis in hypertension with and without diabetes. *Med (Baltimore)*. (2016) 95:e3864. doi: 10.1097/MD.0000000000003864
- Yuan T, Yang T, Chen H, Fu D, Hu Y, Wang J, et al. New insights into oxidative stress and inflammation during diabetes mellitus-accelerated atherosclerosis. *Redox Biol*. (2019) 20:247–60. doi: 10.1016/j.redox.2018.09.025
- Jiang H, Mao H, Lu H, Lin P, Garry W, Lu H, et al. Machine learning-based models to support decision-making in emergency department triage for patients with suspected cardiovascular disease. *Int J Med Inform*. (2021) 145:104326. doi: 10.1016/j.ijmedinf.2020.104326
- Byra M, Galperin M, Ojeda-Fournier H, Olson L, O'Boyle M, Comstock C, et al. Breast mass classification in sonography with transfer learning using a deep convolutional neural network and color conversion. *Med Phys*. (2019) 46:746–55. doi: 10.1002/mp.13361

10. Danielsen AA, Fenger MHJ, Østergaard SD, Nielbo KL, Mors O. Predicting mechanical restraint of psychiatric inpatients by applying machine learning on electronic health data. *Acta Psychiatr Scand.* (2019) 140:147–57. doi: 10.1111/acps.13061
11. Yu D, Liu Z, Su C, Han Y, Duan X, Zhang R, et al. Copy number variation in plasma as a tool for lung cancer prediction using Extreme Gradient Boosting (XGBoost) classifier. *Thorac Cancer.* (2020) 11:95–102. doi: 10.1111/1759-7714.13204
12. Schultebrucks K, Shalev AY, Michopoulos V, Grudzen CR, Shin S-M, Stevens JS, et al. A validated predictive algorithm of post-traumatic stress course following emergency department admission after a traumatic stressor. *Nat Med.* (2020) 26:1084–8. doi: 10.1038/s41591-020-0951-z
13. Shim J-G, Kim DW, Ryu K-H, Cho E-A, Ahn J-H, Kim J-I, et al. Application of machine learning approaches for osteoporosis risk prediction in postmenopausal women. *Arch Osteoporos.* (2020) 15:169. doi: 10.1007/s11657-020-00802-8
14. van Os HJA, Ramos LA, Hilbert A, van Leeuwen M, van Walderveen MAA, Kruyt ND, et al. Predicting outcome of endovascular treatment for acute ischemic stroke: potential value of machine learning algorithms. *Front Neurol.* (2018) 9:784. doi: 10.3389/fneur.2018.00784
15. Jiang F, Jiang Y, Zhi H, Dong Y, Hi H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. *Stroke Vasc Neurol.* (2017) 2:230–43. doi: 10.1136/svn-2017-000101
16. Liang N, Wang C, Duan J, Xie X, Wang Y. Efficacy prediction of noninvasive ventilation failure based on the stacking ensemble algorithm and autoencoder. *BMC Med Inform Decis Mak.* (2022) 22:27. doi: 10.1186/s12911-022-01767-z
17. Liang M, Chang T, An B, Duan X, Du L, Wang X, et al. A stacking ensemble learning framework for genomic prediction. *Front Genet.* (2021) 12:600040. doi: 10.3389/fgene.2021.600040
18. Verma AK, Pal S. Prediction of skin disease with three different feature selection techniques using stacking ensemble method. *Appl Biochem Biotechnol.* (2020) 191:637–56. doi: 10.1007/s12010-019-03222-8
19. Gantenberg JR, McConeghy KW, Howe CJ, Steingrimsson J, van Aalst R, Chit A, et al. Predicting seasonal influenza hospitalizations using an ensemble super learner: A simulation study. *Am J Epidemiol.* (2023) 192:1688–700. doi: 10.1093/aje/kwad113
20. Zhu X, Zhang P, Jiang H, Kuang J, Wu L. Using the Super Learner algorithm to predict risk of major adverse cardiovascular events after percutaneous coronary intervention in patients with myocardial infarction. *BMC Med Res Methodol.* (2024) 24:59. doi: 10.1186/s12874-024-02179-5
21. Li F, Chen J, Ge Z, Wen Y, Yue Y, Hayashida M, et al. Computational prediction and interpretation of both general and specific types of promoters in *Escherichia coli* by exploiting a stacked ensemble-learning framework. *Brief Bioinform.* (2021) 22:2126–40. doi: 10.1093/bib/bbaa049
22. Lundberg SM, Lee S-I. (2017). A unified approach to interpreting model predictions, in: *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Curran Associates Inc., Red Hook, NY, USA. pp. 4768–77.
23. Liu W, Wang S, Ye Z, Xu P, Xia X, Guo M, et al. Prediction of lung metastases in thyroid cancer using machine learning based on SEER database. *Cancer Med.* (2022) 11:2503–15. doi: 10.1002/cam4.4617
24. Kop R, Hoogendoorn M, Teije AT, Büchner FL, Slottje P, Moons LMG, et al. Predictive modeling of colorectal cancer using a dedicated pre-processing pipeline on routine electronic medical records. *Comput Biol Med.* (2016) 76:30–8. doi: 10.1016/j.combiomed.2016.06.019
25. Singal AG, Mukherjee A, Elmunzer BJ, Higgins PDR, Lok AS, Zhu J, et al. Machine learning algorithms outperform conventional regression models in predicting development of hepatocellular carcinoma. *Am J Gastroenterol.* (2013) 108:1723–30. doi: 10.1038/ajg.2013.332
26. Xu H, Wang H, Yuan C, Zhai Q, Tian X, Wu L, et al. Identifying diseases that cause psychological trauma and social avoidance by GCN-Xgboost. *BMC Bioinf.* (2020) 21:504. doi: 10.1186/s12859-020-03847-1
27. van der Laan MJ, Polley EC, Hubbard AE. Super learner. *Stat Appl Genet Mol Biol.* (2007) 6:Article25. doi: 10.2202/1544-6115.1309
28. Naimi AI, Balzer LB. Stacked generalization: an introduction to super learning. *Eur J Epidemiol.* (2018) 33:459–64. doi: 10.1007/s10654-018-0390-z
29. Dalal S, Onyema EM, Malik A. Hybrid XGBoost model with hyperparameter tuning for prediction of liver disease with better accuracy. *World J Gastroenterol.* (2022) 28:6551–63. doi: 10.3748/wjg.v28.i46.6551
30. Štrumbelj E, Kononenko I. Explaining prediction models and individual predictions with feature contributions. *Knowl Inf Syst.* (2014) 41:647–65. doi: 10.1007/s10115-013-0679-x
31. Ahmed SH, Bose DB, Khandoker R, Rahman MS. StackDPP: a stacking ensemble based DNA-binding protein prediction model. *BMC Bioinf.* (2024) 25:111. doi: 10.1186/s12859-024-05714-9
32. Biswas SK, Nath Boruah A, Saha R, Raj RS, Chakraborty M, Bordoloi M, et al. Early detection of Parkinson disease using stacking ensemble method. *Comput Methods Biomech Biomed Engin.* (2023) 26:527–39. doi: 10.1080/10255842.2022.2072683
33. Kapila R, Saleti S. Optimizing fetal health prediction: Ensemble modeling with fusion of feature selection and extraction techniques for cardiocography data. *Comput Biol Chem.* (2023) 107:107973. doi: 10.1016/j.compbiolchem.2023.107973
34. Zhou H, Xin Y, Li S. A diabetes prediction model based on Boruta feature selection and ensemble learning. *BMC Bioinf.* (2023) 24:224. doi: 10.1186/s12859-023-05300-5
35. Fine-Edelstein JS, Wolf PA, O'Leary DH, Poehlman H, Belange AJ, Kase CS, et al. Precursors of extracranial carotid atherosclerosis in the Framingham Study. *Neurology.* (1994) 44:1046–50. doi: 10.1212/WNL.44.6.1046
36. Fan J, Chen M, Luo J, Luo J, Yang S, Shi J, Yao Q, et al. The prediction of asymptomatic carotid atherosclerosis with electronic health records: a comparative study of six machine learning models. *BMC Med Inform Decis Mak.* (2021) 21:115. doi: 10.1186/s12911-021-01480-3