scientific reports

OPEN

The interpretable machine learning model for depression associated with heavy metals via EMR mining method

Site Xu & Mu Sun✉

Limited research exists on the association between depression and heavy metal exposure. This study aims to develop an interpretable and efficient machine learning (ML) model with robust performance to identify depression linked to heavy metal exposure. Data were derived from the US National Health and Nutrition Examination Survey (NHANES) spanning from 2013 to March 2020. We constructed 5 ML models to detect depression based on heavy metal exposure and assessed their performance using 10 discrimination metrics. The optimal model was selected after parameter tuning with a Genetic Algorithm (GA). To enhance the interpretability of the model's predictions, we applied SHapley Additive exPlanation (SHAP) and Local Interpretable Model-Agnostic Explanations (LIME) algorithms. The study included 19,368 participants. The highest-performing model, an eXtreme Gradient Boosting (XGB) algorithm optimized with GA, identified depression using 16 heavy metal indicators (AUC: 0.686; 95% CI: 0.68–0.69; accuracy: 97.1%). SHAP analysis revealed that elevated blood cadmium levels had a positive influence on the model's prediction of depression, while urine concentrations of barium, thallium, tin, manganese, antimony, lead, and tungsten, along with blood levels of lead, cadmium, mercury, selenium, and manganese, showed a negative influence. In conclusion, the study successfully utilized an efficient and robust GA-XGB model to identify depression linked to heavy metal exposure, supported by SHAP and LIME explanations. Blood cadmium was positively correlated with depression, whereas barium, thallium, tin, manganese, antimony, lead, and tungsten in urine, along with lead, cadmium, mercury, selenium, and manganese in blood, were negatively correlated with depression.

Abbreviations

ML	Machine learning
NHANES	National Health and Nutrition Examination Survey
GA	Genetic algorithm
SHAP	SHapley additive exPlanation
LIME	Local interpretable model-agnostic explanations
XGB	eXtreme gradient boosting
ICD-10	International Statistical Classification of Diseases and Related Health Problems, Tenth Revision
PCA	Principal component analysis
SKB	Select K best
DNN	Deep neural networks
SVM	Support vector machine
GNB	Gaussian Naive Bayes
DT	Decision tree
AAUC	Average area under the curve
BAUC	Best area under the curve
APS	Average precision score
ROC	Receiver operating characteristic

Ruijin Hospital, Shanghai Jiaotong University School of Medicine, Shanghai 200025, China. ✉email: schuster_ter@163.com

Depression is marked by persistent and profound emotional distress, and it has become an increasingly serious global mental health issue. It leads to a significant reduction in physical functioning and quality of life, while contributing to a rise in disease incidence and mortality rates. In 2017, approximately 17.3 million adults in the United States, aged 18 and older, experienced at least one major depressive episode, representing a prevalence of around 7.1%[1]. According to the World Health Organization's (WHO) 2017 report, "Depression and Other Common Mental Disorders: Global Health Estimates", there were 322 million individuals worldwide living with depression. Nearly half of these cases are concentrated in Southeast Asia and the Western Pacific regions, including countries like China and India. In China, depressive disorders have been identified as the second leading cause of years lived with disability (YLDs)[2]. The prevalence of depression varies by age, with rates peaking in the elderly. Among women aged 55–74, the estimated prevalence exceeds 7.5%.

Given the low cure rates and ineffective treatments, identifying risk factors for depression is crucial. Early prevention and intervention can effectively slow the progression of the disorder. Various factors have been linked to depression, including age, gender, occupation, and lifestyle[3]. As a leading cause of disability worldwide, depression also has been linked to various environmental factors, including exposure to heavy metals.

Human health, including the occurrence of depression, is often influenced by the combined impact of multiple metals. However, most studies have focused solely on specific metal exposures[4–6], utilizing traditional statistical or ML analyses[7–9]. To more effectively explore the relationship between depression and heavy metal exposure, a novel analytical approach is required that accounts for the combined effects of multiple metal exposures.

Traditional methodologies for disease identification involve numerous stringent standards for preparing datasets. However, with advancements in computer science and the growing volume of information, researchers are increasingly challenged to uncover hidden insights from big data[10]. Machine learning (ML), with its black-box nature, requires fewer preprocessing standards, thereby enhancing the ability to analyze large volumes of information. This capability supports hazard identification and other health-related decision-making processes[11].

Recent research leveraging machine learning techniques has provided novel insights into the relationship between heavy metals and depression, particularly among aging populations. This review synthesizes findings from the past 5 years, focusing on the role of heavy metals in the etiology of depression. Xia et al.[12] conducted a study using machine learning algorithms to analyze data from the National Health and Nutrition Examination Survey (NHANES) 2017–2018, revealing significant associations between depression and specific heavy metals. Their research indicated that cadmium (Cd), ethyl mercury (EtHg), and mercury (Hg) were particularly associated with depression, with Cd and EtHg showing positive correlations and Hg a negative one. This study contributes to a broader literature on the environmental determinants of mental health. For example, Berk et al.[13] found associations between persistent organic pollutants and depressive symptoms, while Scinicariello et al.[3] reported links between hearing loss and depression, which may be influenced by heavy metal exposure. The neurotoxic effects of Cd are well-established, with mechanisms including oxidative stress and interference with essential minerals like zinc and calcium[14]. Mercury's role in neuropsychiatric disorders is also recognized, with EtHg and inorganic mercury species being of particular concern due to their ability to cross the blood-brain barrier[15]. The inverse relationship between Hg and depression, as noted by Xia et al., is intriguing and suggests a complex interplay between environmental toxins and mental health. This finding aligns with other studies that suggest a protective effect of fish consumption against depression, likely due to the nutritional benefits of omega-3 fatty acids outweighing the risks associated with mercury exposure[1].

In our study, we analyzed datasets from the NHANES (2013–March 2020) to explore the relationship between depression and heavy metal exposure. We developed 5 machine learning (ML) models to identify depression based on heavy metal exposure and compared their performance metrics. The best-performing model was further optimized using a GA to enhance its efficiency. Additionally, we incorporated advanced electronic medical record (EMR) mining techniques, specifically SHAP[16] and LIME[17], to evaluate the contribution of heavy metals in the identification of depression. This approach highlights potential associations that could inform future research and guide epidemiological investigations, contributing to a better understanding of early risk factors for intervention.

Methods
Participants
The US NHANES investigates the US population using various survey strategies to collect demographics, dietary, examination, laboratory, and questionnaire data. All data are available on the American Centers for Disease Control and Prevention website (https://www.cdc.gov/nchs/nhanes). Our study sample comprised three contiguous cycles of NHANES data from 2013 to March 2020.

The study applied specific inclusion criteria: participants were required to be over 20 years old, have completed both blood and urine tests for heavy metals, and provided responses to the NHANES questionnaire, which included information on their depression status. The exclusion criteria included participants with more than 10% missing data or any contradictory information. As a result, 19,368 participants were included in the final analysis.

Data collection
Demographics characteristics of the study participants
Participants' demographic and relevant characteristics were gathered from NHANES, including gender, age (in years at screening), Race/Hispanic origin w/ NH Asian, education level (college or above, high school or equivalent, and less than high school), poverty-to-income ratio (PIR) (≤ 1, 1–4, and ≥ 4)[18], and body mass index (BMI, kg/m^2).

Heavy metals
Our analysis incorporated the urinary and blood concentrations of 16 heavy metals. The National Center for Environmental Health implemented strict quality control protocols to ensure accurate detection of all heavy metal levels[19].

Outcome ascertainment
Since the 2013–2014 data release cycle, professional physicians have diagnosed major depressive disorder in NHANES using the codes F32.9 and F33.9, in accordance with the International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10)[20].

Pre-processing of features

In our study, we selected 22 variables (referred to as features in the field of machine learning), including 19 continuous and 3 categorical variables. After splitting the data into training and test sets, we applied data preprocessing only to the training set, ensuring the independence of the test set and preventing data leakage. We excluded data with a loss rate of 10% or higher. Missing values in continuous variables were imputed with the median, unordered categorical variables with the mode, and ordinal categorical variables with the nearest neighbor values. Features were standardized using the Standard Scaler, and categorical variables were transformed using one-hot encoding[21]. We employed Principal Component Analysis (PCA) and the Select K Best (SKB) algorithm for feature extraction[22]. Variables contributing little to the model were removed during preprocessing to prevent overfitting.

The entire process of feature selection and preprocessing was strictly conducted within the training data and did not involve the test set at any stage. The test set remained completely independent and was not used for any feature selection or preprocessing steps.

Model establishment

Repeated K-Fold cross-validation was applied on the training set to construct and evaluate the model[23]. We employed 5 ML algorithms commonly used in the field of EMR mining[8,9,24], including Deep Neural Networks (DNN), Support Vector Machine (SVM), Gaussian Naive Bayes (GNB), Decision Tree (DT), and eXtreme Gradient Boosting (XGB), to establish models for identifying depression based on heavy metal exposure. Each of these 5 models has distinct characteristics. DNN: Typically offers higher accuracy with a simple structure for data training but possesses strong black-box characteristics, making it difficult to understand its decision-making principles[25]. SVM: Robust to data variations and capable of handling nonlinear, multidimensional datasets[26]. GNB: Performs well on small-scale data, supports multiple classification tasks, and is suitable for incremental training, though it may introduce noise and redundancy[27,28]. DT: Supports visual analytics, is easy to understand and interpret, but is prone to overfitting[29]. XGB: An optimized library designed to increase distributed gradient boosting, offering high efficiency, flexibility, and portability[30]. However, it has numerous parameters that need adjustment for optimal performance[31]. To mitigate the class imbalance issue, we applied the built-in class weighting function in the model, which assigns higher importance to the minority class, enhancing the model's ability to detect cases of depression.

Initially, each algorithm's mean performance was evaluated on the training set using K-fold cross-validation, where hyperparameters were tuned to achieve the most stable performance on the validation set (derived from the training set). The most effective machine learning algorithm was then selected based on its performance on an independent test set. We then used a Genetic Algorithm (GA) to fine-tune the parameters of the chosen model to overcome its limitations. SHAP and LIME were applied to interpret the model by highlighting relevant risk variables for identifying depression in participants from 2013 to March 2020[32]. SHAP provided an overall interpretation of the model, while LIME was used for more localized, partial interpretations.

Statistical analysis

Continuous variables were presented as medians with interquartile ranges, while categorical variables were described as counts with percentages. The chi-square test was used to compare group-specific characteristics. Heavy metal concentrations were expressed as geometric means with geometric standard deviations. Trends over the 8 + years (across 3 data release cycles) were analyzed using the Mann-Kendall test.

Model effectiveness was evaluated using several indicators, including average area under the curve (AAUC)[35] and 95% confidence intervals (95%CI), best AUC (BAUC), average precision score (APS), average recall, average f1 score, average accuracy, average Brier score loss, average cross-entropy loss, average Jaccard index, and average Cohen's kappa of each model by repeated K-Fold cross-validation. Focusing on these metrics is more appropriate for imbalanced datasets and provides a more comprehensive evaluation of how well the model identifies cases of depression.

All analyses were conducted using Python 3.9.7, with the majority of the modeling and evaluation processes implemented using the scikit-learn library. A significance level was set at $P < 0.05$. An overview of the methodology is presented in Fig. 1.

Results

Participants' demographics characteristics

The characteristics of the study participants are presented in Table 1. A total of 19,368 individuals were included in the analysis. Of these, 555 were diagnosed with major depressive disorder. The cohort consisted of 9,397 men (48.5%), and the median age of participants was 57 years (33, 69). Those with major depressive disorder were more likely to be women, younger, have a higher BMI, and be non-Hispanic white (all $P < 0.05$).
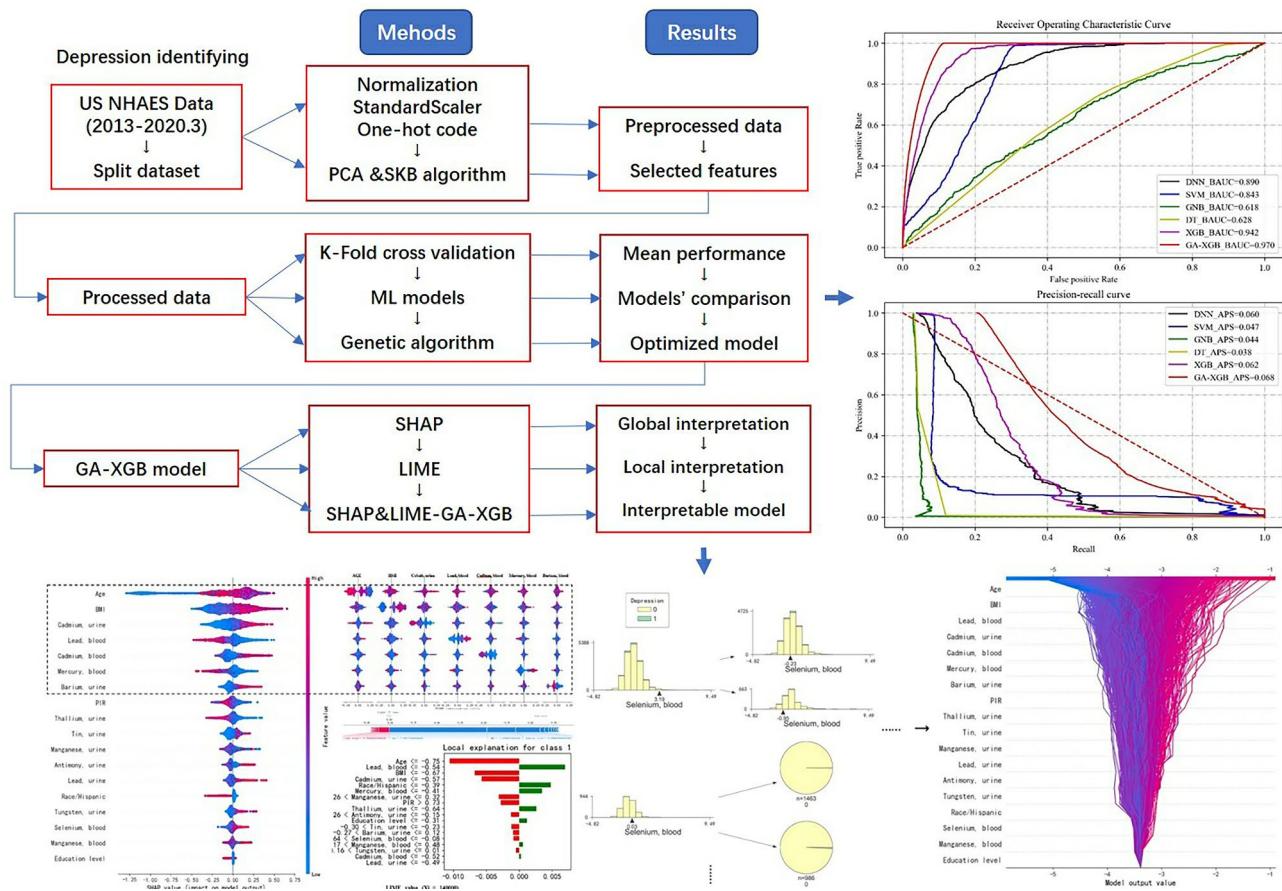
Fig. 1. Overview plot.

Characteristics	Total ($n = 19368$)	With Depression ($n = 555$)	Without Depression ($n = 18813$)	P
Gender: Percentage (%)				<0.001*
Men	9397(48.52%)	1299(49.02%)	8098(48.44%)	
Age: Median (IQR)	57(33, 69)	55(42, 66)	57(33, 69)	0.008*
BMI, kg/m2 : Median (IQR)	28.4 (23.8, 33.6)	31.4 (26.3, 35.65)	28.4 (23.8, 33.5)	<0.001*
Race/Hispanic: Percentage (%)				
Mexican American	2605(13.45%)	47(8.47%)	2558(13.6%)	
Other Hispanic	1888(9.75%)	58(10.45%)	1830(9.73%)	
Non-Hispanic White	7777(40.15%)	316(56.94%)	7461(39.66%)	<0.001*
Non-Hispanic Black	4502(23.24%)	90(16.22%)	4412(23.45%)	
Non-Hispanic Asian	1673(8.64%)	15(2.7%)	1658(8.81%)	
Other Race	923(4.77%)	29(5.23%)	894(4.75%)	
Education level: Percentage (%)				
College or above	11,603(59.91%)	339(61.08%)	11,264(59.87%)	
High school or equivalent	4027(20.79%)	122(21.98%)	3905(20.76%)	0.903
Less than high school	3738(19.3%)	94(16.94%)	3644(19.37%)	
PIR: Percentage (%)				
High	3826(19.75%)	94(16.94%)	3732(19.84%)	
Medium	11,368(58.69%)	332(59.82%)	11,036(58.66%)	0.786
Low	4174(21.55%)	129(23.24%)	4045(21.5%)	

Table 1. The study participants' characteristics in NHAES (2013–2020.3). BMI: body mass index; PIR: poverty to income ratio; *: $P < 0.05$.

Heavy metals' concentrations

The heavy metal concentrations in urine and blood for each data release cycle are described in Table 2. Across the data release cycles, significant trends were observed for Barium, Cadmium, Cobalt, Cesium, Manganese, Lead, Antimony, Tin, Thallium, and Tungsten in urine, as well as Lead, Cadmium, Mercury, Selenium, and Manganese in blood (all $P_{\text{for trend}} < 0.05$).

Models' preprocessing

In the feature selection process, PCA determined that at least 18 variables were needed to retain over 90% of the original information. The SKB feature scores ranged from 0.01 to 1083.44. We selected the top 18 features based on these scores to optimize our ML models. Five machine learning algorithms were then applied to the NHANES datasets using repeated K-Fold cross-validation for model training.

Models' performance

The XGB model exhibited optimal performance with an AAUC of 0.686 (95% CI: 0.68–0.69), a BAUC of 0.942, and an APS of 0.062, all significantly higher than the AUC values of the other four models ($P < 0.05$). To improve AAUC and APS for depression identification, we utilized a GA for parameter tuning, which resulted in the GA-XGB model achieving the best performance. The receiver operating characteristic (ROC) curves and precision-recall curves for all six machine learning models, including GA-XGB, are displayed in Fig. 2. The models demonstrated good accuracy in identifying depression: DNN (97.2%), SVM (97.2%), DT (93.6%), XGB (97.1%), and GA-XGB (97.4%).

Models' comparison

Table 3 compares the performance of the machine learning models, including metrics such as AAUC, BAUC, APS, average recall, average F1 score, average accuracy, average Brier score loss, average cross-entropy loss, average Jaccard index, and average Cohen's kappa for all 5 models. The XGB model achieved the highest scores in 6 out of the 10 performance indicators, demonstrating its superior performance in depression identification. Subsequently, we used GA to optimize the parameters of the XGB model, further enhancing its effectiveness, as shown in the far right of Table 3. Specifically, the GA-XGB model achieved the highest scores in 7 out of the 10 discrimination characteristics. The GA-XGB model's performance metrics were AAUC (AUC: 0.669; 95% CI: 0.663–0.676), BAUC (0.97), and APS (0.068).

Feature importance visualization

SHAP and LIME were employed to visualize the influence of features on depression identification in the GA-XGB model. The SHAP and LIME summary plot illustrates the impact of each selected feature on the model's performance in identifying depression (Fig. 3).

The SHAP value plot on the left side of Fig. 3 globally indicates that Cadmium (20.636) in blood positively influenced the model, while Barium (− 30.558), Thallium (− 11.242), Tin (− 12.339), Manganese (− 17.385), Antimony (− 19.088), Lead (− 23.989), Tungsten (− 21.126) in urine, and Lead (− 111.499), Cadmium (− 35.003), Mercury (− 70.835), Selenium (− 10.389), Manganese (− 16.206) in blood negatively influenced the

Heavy metal	Cycles of NHANES			P for trend
	2013–2014	2015–2016	2017–2020.3	
In urine				
Barium (ug/L)	1.6(2.68)	1.59(2.25)	1.45(2.34)	< 0.001*
Cadmium (ug/L)	0.34(0.44)	0.32(0.39)	0.39(0.55)	< 0.001*
Cobalt (ug/L)	0.68(2.09)	0.66(1.52)	0.56(1.12)	< 0.001*
Cesium (ug/L)	4.86(3.24)	4.96(3.83)	5.11(3.47)	< 0.001*
Manganese (ug/L)	51.45(49.43)	54.5(50.45)	51.12(49.21)	< 0.001*
Molybdenum (ug/L)	0.19(0.76)	0.15(0.42)	0.16(0.3)	0.424
Lead, urine (ug/L)	0.49(0.7)	0.48(0.62)	0.45(0.55)	< 0.001*
Antimony, urine (ug/L)	0.07(0.12)	0.08(0.28)	0.07(0.15)	< 0.001*
Tin, urine (ug/L)	1.89(6.21)	1.49(3.08)	1.46(3.63)	< 0.001*
Thallium, urine (ug/L)	0.17(0.14)	0.18(0.16)	0.19(0.14)	< 0.001*
Tungsten, urine (ug/L)	0.14(0.46)	0.13(0.26)	0.12(0.4)	< 0.001*
In blood				
Lead (ug/dL)	1.33(1.57)	1.24(1.12)	1.18(1.05)	< 0.001*
Cadmium (ug/L)	0.46(0.52)	0.43(0.5)	0.48(0.65)	< 0.001*
Mercury, total (ug/L)	1.37(2.46)	1.22(2.0)	1.17(2.28)	< 0.001*
Selenium (ug/L)	194.46(26.82)	189.79(25.5)	183.93(26.83)	< 0.001*
Manganese (ug/L)	9.78(3.72)	10.17(3.71)	9.58(3.69)	< 0.001*

Table 2. Mean values of heavy metal concentration by each NHANES (2013–2020.3) data release cycle. *$P < 0.05$.
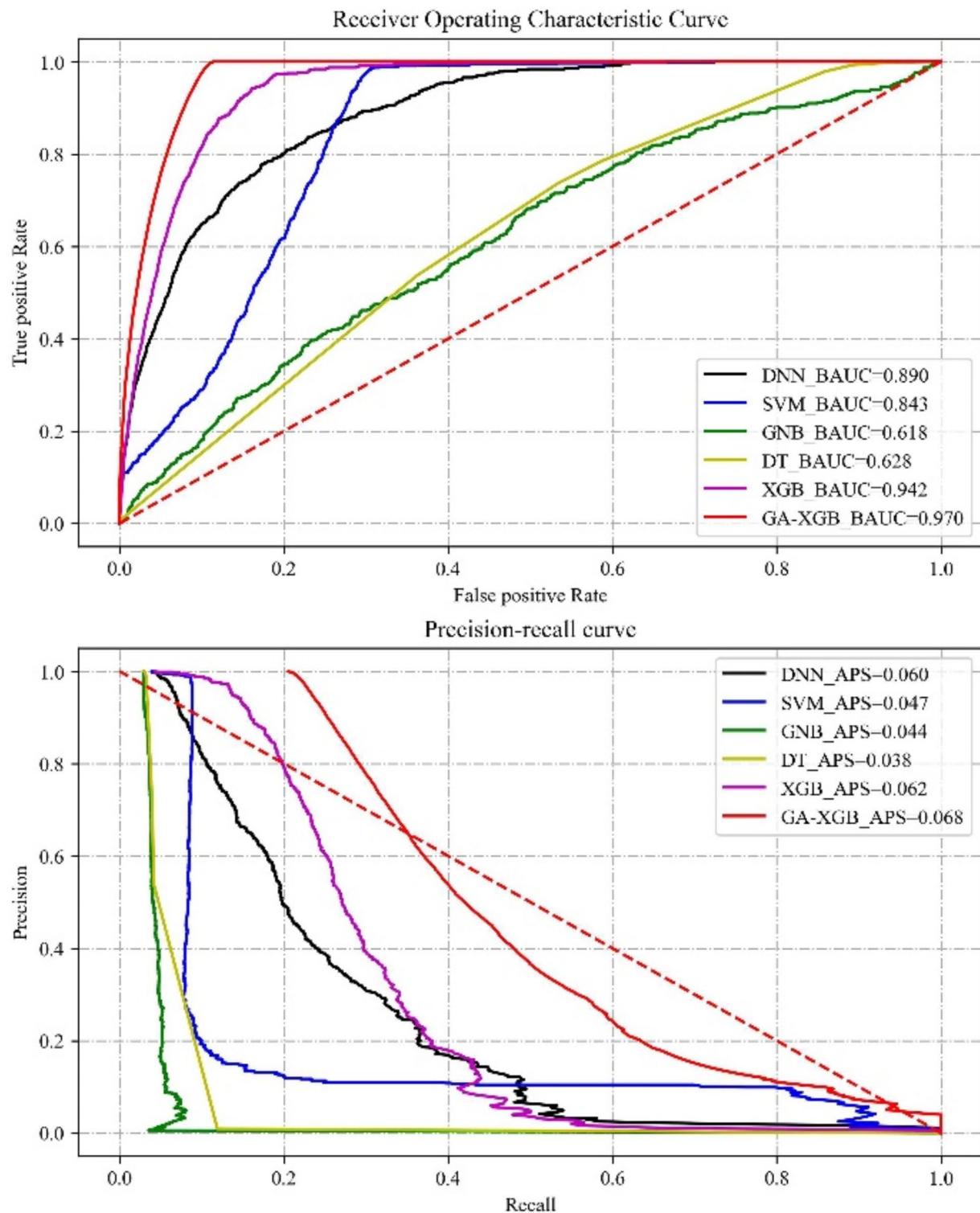
Fig. 2. The best receiver operating characteristic curve and precision-recall curve for models.

model. Additionally, the SHAP and LIME summary plot with statistical tests shows that being women, younger, non-Hispanic, and having a lower BMI are associated with a higher risk of depression. The SHAP interaction value plot, located on the upper right of Fig. 3, demonstrates the interactions between key features. The LIME value plot, on the lower right of Fig. 3, locally indicates the feature importance for a single sample (the 14,000th sample). SHAP values illustrate the contributions of each feature to the model's ability to identify depression.

Characteristics	DNN	SVM	GNB	DT	XGB	GA-XGB
AAUC	0.657 (0.65, 0.664)	0.6 (0.594, 0.606)	0.608 (0.599, 0.616)	0.614 (0.608, 0.62)	0.686 (0.68, 0.69)	0.669 (0.663, 0.676)
BAUC	0.89	0.843	0.618	0.628	0.942	0.97
APS	0.06	0.047	0.044	0.038	0.062	0.068
Average recall	0.972	0.972	0.936	0.971	0.971	0.974
Average f1 score	0.958	0.958	0.941	0.957	0.957	0.964
Average accuracy	0.972	0.972	0.936	0.971	0.971	0.974
Average Brier score loss	0.028	0.029	0.064	0.029	0.027	0.026
Average cross-entropy loss	0.976	0.984	2.21	0.99	0.99	0.908
Average Jaccard index	0.944	0.944	0.91	0.943	0.944	0.949
Average Cohen's kappa	0.031	0.011	0.04	0	0	0.178

Table 3. Comparison of ML models' performance. DNN: deep neural networks; SVM: support vector machine; GNB: Gaussian naive bayes; DT: decision tree classifier; XGB: extreme gradient boosting; AAUC: average area under the curve; BAUC: best area under the curve; APS: average precision score; NA: null.
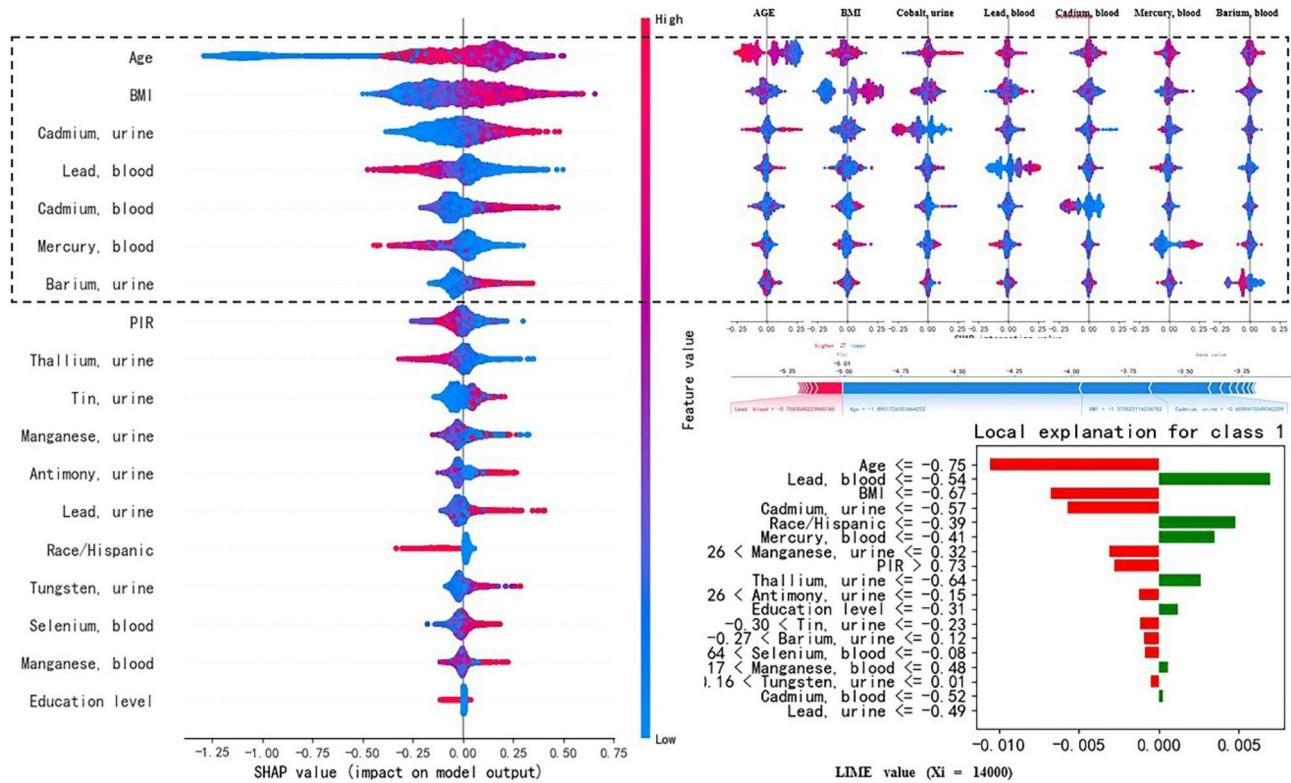
Fig. 3. The SHAP&LIME-GA-XGB summary plot.

Prediction interpretation

In the SHAP decision plot on the right side of Fig. 4, each line represents an individual participant, with all lines converging at a single point, 0.971. The features are arranged in descending order based on their impact on the observations. On the left side of Fig. 4, the tree plot illustrates the optimal decision logic used for discrimination, representing one of the fundamental trees in the model's decision-making process.

Discussion

In our study, we developed a ML strategy to identify depression in the 2013–2020.3 NHANES data, with a focus on its relationship with heavy metal exposure. The GA-XGB model was selected for its superior performance among the five ML algorithms tested, achieving an average AUC of 0.959 and an accuracy of 0.968. To improve the interpretability of these algorithms, we combined the SHAP game theory method with LIME, enabling more comprehensive feature interpretation on both global and local scales through summary and decision plots. Our
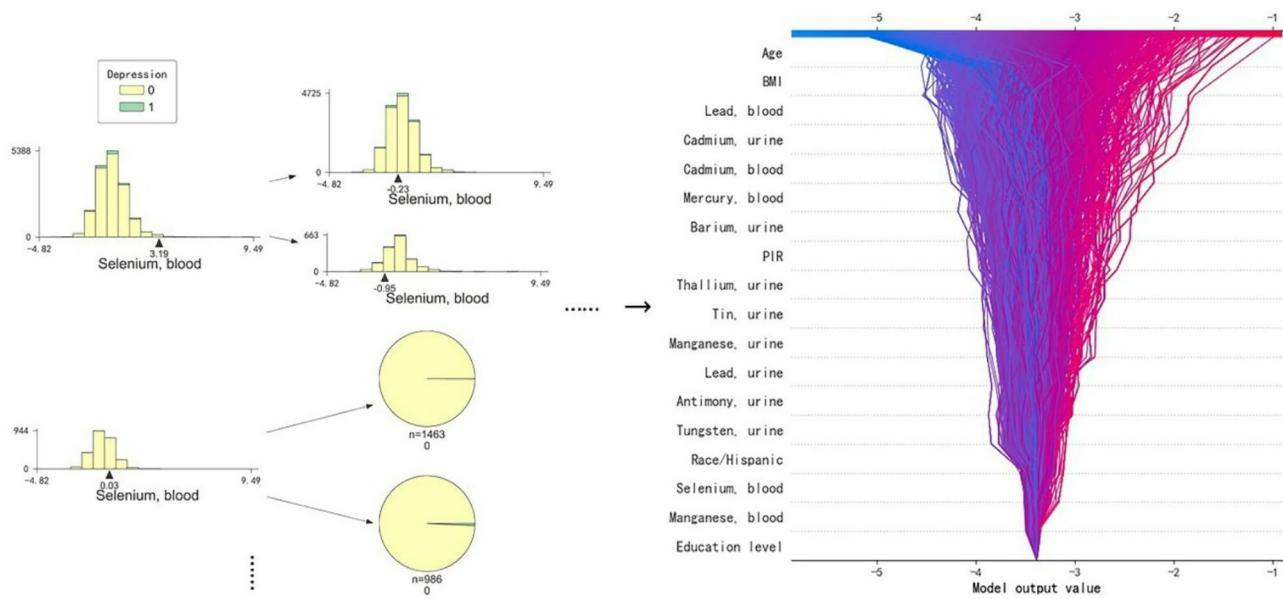
Fig. 4. The SHAP-GA-XGB decision plot.

findings suggest that the SHAP & LIME-enhanced GA-XGB model shows promising potential for identifying depression associated with heavy metal exposure.

This research builds on previous studies that used machine learning (ML) algorithms for disease prediction, highlighting the advantages of advanced classification techniques in enhancing prediction accuracy. ML, a branch of artificial intelligence, employs mathematical algorithms to detect patterns in diverse datasets, thereby aiding in the decision-making process[34]. However, the complexity of ML algorithms often limits their interpretability, making it challenging to apply them effectively in medical decision-making[35].

Our SHAP & LIME-GA-XGB model utilizes multi-source NHANES data, including demographics, examinations, laboratory results, and questionnaires, eliminating the need for additional data collection. Since 2013, significant focus has been placed on heavy metal exposure in the United States[36], coinciding with the implementation of ICD-10 for recording NHANES disease data. [39] We analyzed extensive data, particularly the concentrations of heavy metals in participants' urine and blood samples. The GA-XGB model demonstrated high efficiency, outperforming six tested ML algorithms in terms of classification robustness, supported by repeated K-Fold cross-validation to prevent overfitting[38]. SHAP and LIME analyses further enhanced the interpretability of the GA-XGB model, emphasizing the importance of various features in identifying depression.

The findings from SHAP were consistent with those of previous studies, which primarily investigated the impact of heavy metal exposure on depression. Notably, the relationship between cadmium exposure and depression is particularly significant. Cybulska[39] and Buser[40] found that higher blood cadmium levels were associated with an increased risk of depressive symptoms. However, Gao[41] and Rhee[42] found that lower levels of serum uric acid, which can be influenced by cadmium exposure, were associated with depression. These findings suggest a complex relationship between cadmium exposure, depression, and potential moderating factors.

Future research should focus on monitoring and analyzing key features to help experts draw informed conclusions, rather than relying solely on algorithmic predictions. Expanding the dataset and incorporating clinical expertise could further enhance the model's validity and performance[43].

Limitations

Our study has several limitations. Firstly, due to computational constraints, we were unable to explore other potentially dynamic correlations within the limited dataset. Secondly, the self-reported nature of depression diagnoses in the NHANES questionnaire, despite following ICD-10 standards, may introduce information bias. Thirdly, the strict inclusion criteria resulted in substantial missing data, potentially leading to bias. Fourthly, the complexity of the model's interpretation may impact the reproducibility of our findings. Fifthly, the integration of machine learning into epidemiological research offers a powerful tool for identifying patterns in large and complex datasets. However, the cross-sectional nature of the NHANES data limits the ability to establish causality, and the study's focus on heavy metals means that other potential risk factors for depression are not considered. Lastly, in this study, feature selection and preprocessing were conducted strictly within the training data, with the test set kept entirely independent to ensure unbiased evaluation. However, feature selection was not embedded within each fold of the cross-validation process, which may risk slight overestimation of performance during internal validation. While this approach improves processing consistency and efficiency, embedding feature selection within folds would provide a more rigorous methodology. Future studies could adopt this workflow to further enhance robustness. Nevertheless, the final performance evaluation relied exclusively on an independent test set, mitigating the risk of data leakage and ensuring reliable generalizability.

Conclusion

In our study among US NHANES 2013–2020.3 participants, the SHAP&LIME-GA-XGB model was identified as an interpretable, efficient, and robust machine learning model for detecting depression based on heavy metal exposure. Cadmium in blood positively contribute to depression, while Barium, Thallium, Tin, Manganese, Antimony, Lead, Tungsten in urine, and Lead, Cadmium, Mercury, Selenium, Manganese in blood negatively contribute to depression.

Data availability

The datasets that support the findings of this study are available publicly. Full lists of records identified through database searching are available on reasonable request from the corresponding author. Correspondence: schuster_ter@163.com.

References

1. Kim, K. W., Sreeja, S. R., Kwon, M., Yu, Y. L. & Kim, M. K. Association of blood mercury level with the risk of depression according to fish intake level in the general Korean population: Findings from the Korean National health and nutrition examination survey (KNHANES) 2008–2013. *Nutrients* **12** (1), 189. https://doi.org/10.3390/nu12010189 (2020).
2. Lu, J. et al. Prevalence of depressive disorders and treatment in China: A cross-sectional epidemiological study. *Lancet Psychiatry*. **8** (11). https://doi.org/10.1016/s2215-0366(21)00251-0 (2021).
3. Scinicariello, F. et al. Age and sex differences in hearing loss association with depressive symptoms: Analyses of NHANES 2011–2012. *Psychol. Med.* **49** (6), 962–968. https://doi.org/10.1017/s0033291718001617 (2018).
4. Shi, P., Jing, H. & Xi, S. Urinary metal/metalloid levels in relation to hypertension among occupationally exposed workers. *Chemosphere* **234**, 640–647. https://doi.org/10.1016/j.chemosphere.2019.06.099 (2019).
5. Qian, H. et al. Relationship between occupational metal exposure and hypertension risk based on conditional logistic regression analysis. *Metabolites.* **12**(12), 1259 (2022). https://doi.org/10.3390/metabo12121259
6. Zhong, Q. et al. Urinary metal concentrations and the incidence of hypertension among adult residents along the Yangtze river, China. *Arch. Environ. Contam. Toxicol.* **77** (4), 490–500. https://doi.org/10.1007/s00244-019-00655-4 (2019).
7. Zheng, S. et al. Associations between plasma metal mixture exposure and risk of hypertension: A cross-sectional study among adults in Shenzhen, China. *Front. Public. Health.* **10**, 1039514. https://doi.org/10.3389/fpubh.2022.1039514 (2022).
8. Xu, S. et al. Cost supervision mining from EMR based on artificial intelligence technology. *Technol. Health Care.* **31** (3), 1077–1091. https://doi.org/10.3233/THC-220608 (2023).
9. Xu, S. & Sun, M. Covid-19 vaccine effectiveness during Omicron BA.2 pandemic in Shanghai: A cross-sectional study based on EMR. *Med. (Baltim).* **101** (45), e31763. https://doi.org/10.1097/MD.0000000000031763 (2022).
10. Stafford, I. S. et al. A systematic review of the applications of artificial intelligence and machine learning in autoimmune diseases. *NPJ Digit. Med.* **3**, 30. https://doi.org/10.1038/s41746-020-0229-3 (2020). Published 2020 Mar 9.
11. Alber, M. et al. Integrating machine learning and multiscale modeling-perspectives, challenges, and opportunities in the biological, biomedical, and behavioral sciences. *NPJ Digit. Med.* **2**, 115. https://doi.org/10.1038/s41746-019-0193-y (2019).
12. Xia, F., Li, Q., Luo, X. & Wu, J. Machine learning model for depression based on heavy metals among aging people: A study with National health and nutrition examination survey 2017–2018. *Front. Public. Health.* **10**, 939758. https://doi.org/10.3389/fpubh.2022.939758 (2022).
13. Berk, M. et al. Pop, heavy metal, and the blues: secondary analysis of persistent organic pollutants (POP), heavy metals and depressive symptoms in the NHANES National epidemiological survey. *BMJ Open.* **4** (7), e5142. https://doi.org/10.1136/bmjopen-2014-005142 (2014).
14. Buser, M. C. & Scinicariello, F. Cadmium, lead, and depressive symptoms. *J. Clin. Psychiatry.* **78** (6), e515–e522. https://doi.org/10.4088/JCP.15m10383 (2017).
15. Li, W., Ruan, W., Cui, X., Lu, Z. & Wang, D. Blood volatile organic aromatic compounds concentrations across adulthood in relation to total and cause-specific mortality: A prospective cohort study. *Chemosphere* **286**, 131590. https://doi.org/10.1016/j.chemosphere.2021.131590 (2022).
16. Nordin, N., Zainol, Z., Mohd Noor, M. H. & Chan, L. F. An explainable predictive model for suicide attempt risk using an ensemble learning and Shapley additive explanations (SHAP) approach. *Asian J. Psychiatr.* **79**, 103316. https://doi.org/10.1016/j.ajp.2022.103316 (2023).
17. Peng, K. & Menzies, T. Documenting evidence of a reuse of 'why should I trust you? explaining the predictions of any classifier'. In *Proceedings of the 29th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering* (pp. 1600–1600). (2021). https://doi.org/10.1145/3468264.3477217
18. Odutayo, A. et al. Income disparities in absolute cardiovascular risk and cardiovascular risk factors in the united States, 1999–2014. *JAMA Cardiol.* **2** (7), 782–790. https://doi.org/10.1001/jamacardio.2017.1658 (2017).
19. NHANES. NHANES 2013–2014 Laboratory Methods. (2023). https://wwwn.cdc.gov/nchs/nhanes/ContinuousNhanes/. Accessed October 20.
20. Mou, C. & Ren, J. Automated ICD-10 code assignment of nonstandard diagnoses via a two-stage framework. *Artif. Intell. Med.* **108**, 101939. https://doi.org/10.1016/j.artmed.2020.101939 (2020).
21. Rodríguez, P., Bautista, M. A., Gonzalez, J. & Escalera, S. Beyond one-hot encoding: Lower dimensional target embedding. *Image Vis. Comput.* **75**, 21–31. https://doi.org/10.1016/j.imavis.2018.04.004 (2018).
22. Desyani, T., Saifudin, A. & Yulianti, Y. Feature selection based on naive bayes for caesarean section prediction. In *IOP Conference Series: Materials Science and Engineering* (Vol. 879, No. 1, p. 012091). (2020). https://doi.org/10.1088/1757-899X/879/1/012091
23. Barile, C., Casavola, C., Pappalettera, G. & Kannan, V. P. Damage progress classification in AlSi10Mg SLM specimens by convolutional neural network and k-Fold cross validation. *Mater. (Basel).* **15** (13), 4428. https://doi.org/10.3390/ma15134428 (2022). Published 2022 Jun 23.
24. Xu, S. & Sun, M. The interpretable machine learning model associated with metal mixtures to identify hypertension via EMR mining method. *J. Clin. Hypertens.* https://doi.org/10.1111/jch.14768 (2024).
25. Du, X., Liu, M. & Sun, Y. Cell recognition using BP neural network edge computing. *Contrast Media Mol. Imaging.* https://doi.org/10.1155/2022/7355233 (2022). 2022:7355233. Published 2022 Jul 12.
26. Kim, M. et al. Machine learning models to identify low adherence to influenza vaccination among Korean adults with cardiovascular disease. *BMC Cardiovasc Disord.* **21**(1), 129 (2021). https://doi.org/10.1186/s12872-021-01925-7
27. Ding, X., Zhang, H., Ma, C., Zhang, X. & Zhong, K. User identification across multiple social networks based on Naive Bayes model [published online ahead of print, 2022 Sep 14]. *IEEE Trans. Neural Netw. Learn. Syst.* **PP** https://doi.org/10.1109/TNNLS.2022.3202709 (2022).

28. Yang, S. et al. A synthesis framework using machine learning and Spatial bivariate analysis to identify drivers and hotspots of heavy metal pollution of agricultural soils. *Environ. Pollut.* **287**, 117611. https://doi.org/10.1016/j.envpol.2021.117611 (2021).

29. Zweck, E. et al. Machine learning identifies clinical parameters to predict mortality in patients undergoing transcatheter mitral valve repair. *JACC Cardiovasc. Interv.* **14** (18), 2027–2036. https://doi.org/10.1016/j.jcin.2021.06.039 (2021).

30. Xia, F., Li, Q., Luo, X. & Wu, J. Identification for heavy metals exposure on osteoarthritis among aging people and machine learning for prediction: A study based on NHANES 2011–2020. *Front. Public. Health.* **10**, 906774. https://doi.org/10.3389/fpubh.2022.906774 (2022).

31. Deng, J. et al. Automatic cardiopulmonary endurance assessment: A machine learning approach based on GA-XGBOOST. *Diagnostics (Basel).* **12** (10), 2538. https://doi.org/10.3390/diagnostics12102538 (2022). Published 2022 Oct 19.

32. El Bilali, A. et al. An interpretable machine learning approach based on DNN, SVR, extra tree, and XGBoost models for predicting daily Pan evaporation. *J. Environ. Manag.* **327**, 116890. https://doi.org/10.1016/j.jenvman.2022.116890 (2023).

33. Pruessner, J. C., Kirschbaum, C., Meinlschmid, G. & Hellhammer, D. H. Two formulas for computation of the area under the curve represent measures of total hormone concentration versus time-dependent change. *Psychoneuroendocrinology* **28** (7), 916–931. https://doi.org/10.1016/s0306-4530(02)00108-7 (2003).

34. Akyea, R. K., Qureshi, N., Kai, J. & Weng, S. F. Performance and clinical utility of supervised machine-learning approaches in detecting Familial hypercholesterolaemia in primary care. *NPJ Digit. Med.* **3**, 142. https://doi.org/10.1038/s41746-020-00349-5 (2020). Published 2020 Oct 30.

35. Srour, B. et al. Ultraprocessed food consumption and risk of type 2 diabetes among participants of the NutriNet-Santé prospective cohort. *JAMA Intern. Med.* **180** (2), 283–291. https://doi.org/10.1001/jamainternmed.2019.5942 (2020).

36. Guney, M. & Zagury, G. J. Contamination by ten harmful elements in toys and children's jewelry bought on the North American market. *Environ. Sci. Technol.* **47** (11), 5921–5930. https://doi.org/10.1021/es304969n (2013).

37. Yin, R. et al. Hypertension in China: burdens, guidelines and policy responses: A state-of-the-art review. *J. Hum. Hypertens.* **36** (2), 126–134. https://doi.org/10.1038/s41371-021-00570-z (2022).

38. Rajkomar, A. et al. Scalable and accurate deep learning with electronic health records. *NPJ Digit. Med.* **1**, 18. https://doi.org/10.1038/s41746-018-0029-1 (2018).

39. Cybulska, A. M. et al. Are cadmium and lead levels linked to the development of anxiety and depression?—A systematic review of observational studies. *Ecotoxicol. Environ. Saf.* **216**, 112211. https://doi.org/10.1016/j.ecoenv.2021.112211 (2021).

40. Buser, M. C. & Scinicariello, F. Cadmium, lead, and depressive symptoms. *J. Clin. Psychiatry.* **78** (5), e515–e521. https://doi.org/10.4088/jcp.15m10383 (2017).

41. Gao, S., Zuk, A. O., Wu, M., Tassone, V. K. & Jung, H. Association between depression and urinary heavy metal levels. *Univ. Tor. J. Public. Health.* **4** (1). https://doi.org/10.33137/utjph.v4i1.41690 (2023).

42. Rhee, S. J., Lee, H. & Ahn, Y. M. Association between serum uric acid and depressive symptoms stratified by low-grade inflammation status. *Sci. Rep.* **11** (1). https://doi.org/10.1038/s41598-021-99312-x (2021).

43. Choi, D. J., Park, J. J., Ali, T. & Lee, S. Artificial intelligence for the diagnosis of heart failure. *NPJ Digit. Med.* **3**, 54. https://doi.org/10.1038/s41746-020-0261-3 (2020).

Author contributions

All authors contributed to designing the study. Xu S was responsible for data collection and analysis. Xu S was responsible for writing the manuscript. The corresponding author Sun attested that all listed authors meet authorship criteria. No other individuals meeting the criteria have been omitted. Sun is the guarantor. All authors have read and approved the final manuscript.

Funding

Declarations

Consent for publication

Informed consent was obtained from all individual participants included in the study.

Competing interests

The authors declare no competing interests.

Ethics statement

We did not take part in the participant recruiting since this analysis was based on the US NHANES's already-available data. As far as we are aware, no patients were involved in the planning, selection, or execution of the study.

Additional information

Correspondence and requests for materials should be addressed to M.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.