THE ANATOLIAN JOURNAL OF CARDIOLOGY

Official journal of the TURKISH SOCIETY OF CARDIOLOGY

Identifying Cardiovascular Disease Risk Factors in Adults with Explainable Artificial Intelligence

ORIGINAL INVESTIGATION

ABSTRACT

Background: The aim of this study was to evaluate the relationship between risk factors causing cardiovascular diseases and their importance with explainable machine learning models.

Methods: In this retrospective study, multiple databases were searched, and data on 11 risk factors of 70 000 patients were obtained. Data included risk factors highly associated with cardiovascular disease and having/not having any cardiovascular disease. The explainable prediction model was constructed using 7 machine learning algorithms: Random Forest Classifier, Extreme Gradient Boost Classifier, Decision Tree Classifier, KNeighbors Classifier, Support Vector Machine Classifier, and GaussianNB. Receiver operating characteristic curve, Brier scores, and mean accuracy were used to assess the model's performance. The interpretability of the predicted results was examined using Shapley additive description values.

Results: The accuracy, area under the curve values, and Brier scores of the Extreme Gradient Boost model (the best prediction model for cardiovascular disease risk factors) were calculated as 0.739, 0.803, and 0.260, respectively. The most important risk factors in the permutation feature importance method and explainable artificial intelligence—Shapley's explanations method are systolic blood pressure (ap_hi) [0.1335 ± 0.0045 w (weight)], cholesterol (0.0341 ± 0.0022 w), and age (0.0211 ± 0.0036 w).

Conclusion: The created explainable machine learning model has become a successful clinical model that can predict cardiovascular patients and explain the impact of risk factors. Especially in the clinical setting, this model, which has an accurate, explainable, and transparent algorithm, will help encourage early diagnosis of patients with cardiovascular diseases, risk factors, and possible treatment options.

Keywords: Cardiovascular disease, explainable artificial intelligence, machine learning, prediction, risk factors

INTRODUCTION

Cardiovascular diseases (CVDs) have become one of the leading diseases worldwide, with the highest rate of death and disability, regardless of the countries are developed or not. According to the World Health Organization, CVD is the cause of death of approximately 17.9 million people each year.[1] As a result of the rapidly worsening disease course, it has become complicated to examine the risk factors.[2] Many *in vitro*, *in vivo*, and *in silico* studies have been carried out on CVD risk factors. Studies have shown that systolic and diastolic blood pressures are related to[3] smoking,[3-5] aging,[5-7] weight subgroups,[8,9] cholesterol,[10,11] glucose level,[12,13] and alcohol consumption[14-16] risk factors. Due to the complex and extensive process of risk factors, early diagnosis of CVD and its treatment phase have become a necessity. In recent studies, machine learning (ML) techniques, which can explain the complex interactions of disease-related factors with these diseases, are increasingly used in prediction models.[17-20] Many studies have been conducted using ML and deep learning (DL) methods. Al'Aref et al[21] contributed to CVD studies in 2018 with non-invasive imaging methods, Ghosh et al[22] used ML models in 2021, and Aryal et al[23] used gut microbiome-based ML methods in 2020. Explainable artificial intelligence (XAI) methods have emerged as artificial intelligence (AI) raises concerns about the explainability of the model and

Kevser Kübra Kırboğa[1,2]

Ecir Uğur Küçüksille[3]

[1]Department of Bioengineering, Bilecik Seyh Edebali University, Faculty of Engineering, Bilecik, Türkiye
[2]Informatics Institute, İstanbul Technical University, İstanbul, Türkiye
[3]Department of Computer Engineering, Süleyman Demirel University, Isparta, Türkiye

Corresponding author:
Kevser Kübra Kırboğa
kubra.kirboga@bilecik.edu.tr

Received: March 7, 2023
Accepted: July 3, 2023
Available Online Date: August 25, 2023

Cite this article as: Kırboğa KK, Küçüksille EU. Identifying cardiovascular disease risk factors in adults with explainable artificial intelligence. *Anatol J Cardiol.* 2023;27(11):657-663.

DOI:10.14744/AnatolJCardiol.2023.3214

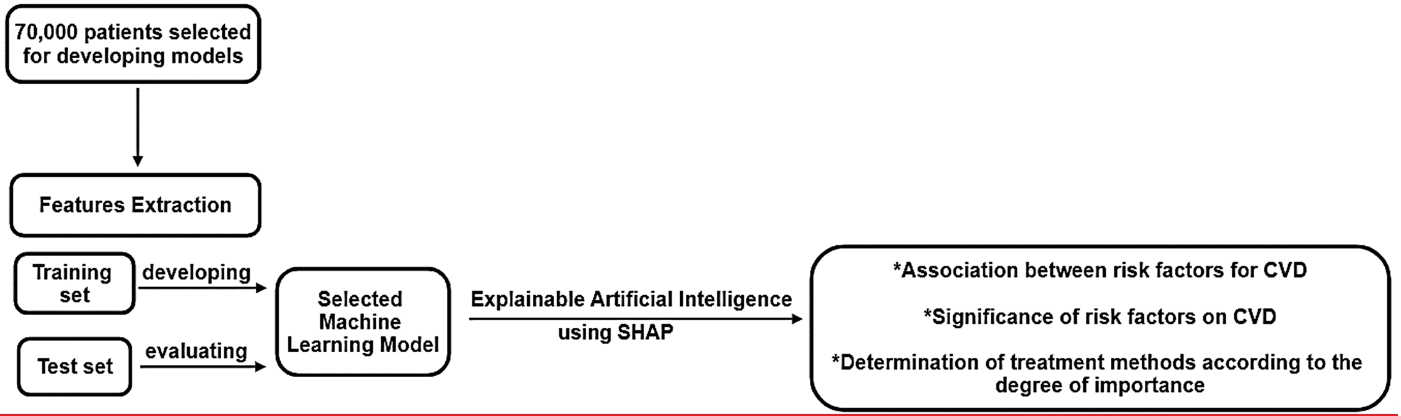

Figure 1. Flowchart illustrating the stages of this retrospective study. CVD, cardiovascular diseases; SHAP, Shapley additive description.

algorithm.[24] Explainable artificial intelligence has been used recently to help healthcare professionals understand the relationships between ML models and their predicted outcomes.[25,26] At the same time, XAI can improve the interpretative ability of algorithms and maintain the predictive accuracy of complex ML models. There is a lack of XAI studies among the studies on CVD risk factors. In this study, feature importance selection and XAI methods were used on the dataset containing CVD risk factors. We compared the dominant importance of risk factors on CVD with the standard importance method, permutation feature importance (PFI) method, and XAI method. Using the Shapley additive explanations (SHAP) method, we identified risk factors that significantly impact CVD. We revealed the relationships between CVD and risk factors with various XAI graphs and proved the accuracy of our explainable model by comparing it with the literature data.

METHODS

Data Properties and Processing

This dataset, which contains detailed information about risk factors for CVD, was obtained by searching "Exploring Risk Factors for Cardiovascular Disease in Adults" on Kaggle. Additionally, the dataset is available at https://data.world/kudem. The dataset includes information about age, gender, height, weight, blood pressure values, cholesterol and glucose levels, smoking habits, whether the person is active or not, and alcohol consumption, which are frequently among the CVD risk factors of 70 000 individuals. In addition, there is information about whether there is any CVD, which is the main factor in achieving the aim of the study and the expected results. It is thought that this dataset will provide an excellent resource for researchers to integrate various ML, DL, and AI techniques to explore potential relationships between CVD risk factors and CVD, which can ultimately lead to a better understanding of CVD risk factors and designing better preventive measures (Figure 1). For this reason, it was seen as a dataset suitable for the study and used as a compass for determining appropriate treatment options due to early diagnosis of CVD.

HIGHLIGHTS

- Systolic blood pressure, cholesterol, and age are significant predictors of cardiovascular diseases.
- Explainable artificial intelligence is a predicting tool for diagnosing predictable cardiovascular diseases.
- The created extreme gradient boosting model, with its explicable, interpretable, and transparent features, enabled the prediction to be made with high accuracy.

Permutation Feature Importance

Permutation feature importance explains which features mislead the model's performance and score by measuring the decline in model performance after mixing single feature values.[27,28] This agnostic method tells us how dependent the model is on each feature relative to the drop in the model score. In this study, the permutation_importance function is used to show the importance of each feature in the dataset that affects CVD.

Shapley Additive Explanations

Shapley additive explanations method was first used in a game theory to determine players' contributions to the game.[29] This algorithm is a great way to create synergism with engineering a black-box model. When combined with Lundberg and Lee's[30] ML algorithms, this framework, which works with Shapley values, brought transparency, explainability, and predictability to black-box models.[31] The SHAP works on all possible permutations of the features and calculates the average of the contributions. This study presents an explanatory design with SHAP algorithms for CVD risk factors. Although there have been many studies of importance related to CVD risk factors, these studies may be insufficient to reach the interactions between factors for model outputs.

For this reason, XAI integration was made using the dataset containing many risk factors. We investigated individual and collective force plots to understand how risk factor variables affect model results. Next, we compare the TreeSHAP findings with previous studies suggesting the causative factor of CVD and the influence of risk factors on CVD.

Model Development and Statistical Analysis

The statistical significance of all classifier algorithms was evaluated with several criteria derived from precision, recall, F1-score, mean accuracy, area under the curve (AUC), and Brier scores for training, validation, and test sets. We randomly divided datasets into 80% calibration (training and validation) and 20% test sets. Validation and training sets were used in the model-building process to find and correct the best combination of parameters and to make an initial assessment of the model's performance. The discriminative power was evaluated using receiver operating characteristic curve analysis. Model accuracy scores, which represent the discrepancy between the projected probability of the models and the actual observed outcomes, were used to determine the overall model accuracy. The advanced prediction capabilities and generalization performance of the optimized models were evaluated by test sets. Seven ML models were used to develop predictive models predicting patients with and without CVD. These models included Random Forest Classifier, Extreme Gradient Boost (XGBoost) Classifier, Decision Tree Classifier, KNeighbors Classifier, SVM Classifier, and GaussianNB. Many techniques were investigated to develop the best model with the highest and most accurate performance. Python 3.7.9 was used for data analysis (available from the Python Software Foundation, python.org).

RESULTS

In our study, we used a unique data set of 70 000 patient data that included clinical information on CVD risk factors and outcomes. Among the 7 algorithms, the accuracy and prediction values of the XGBoost model were calculated as 0.739 and 0.72, respectively. In addition, it was chosen as the best model with a value of 0.80 AUC. Model algorithms were evaluated and selected according to AUC (Figure 2), Brier score, and mean accuracy (Table 1). The XGBoost model

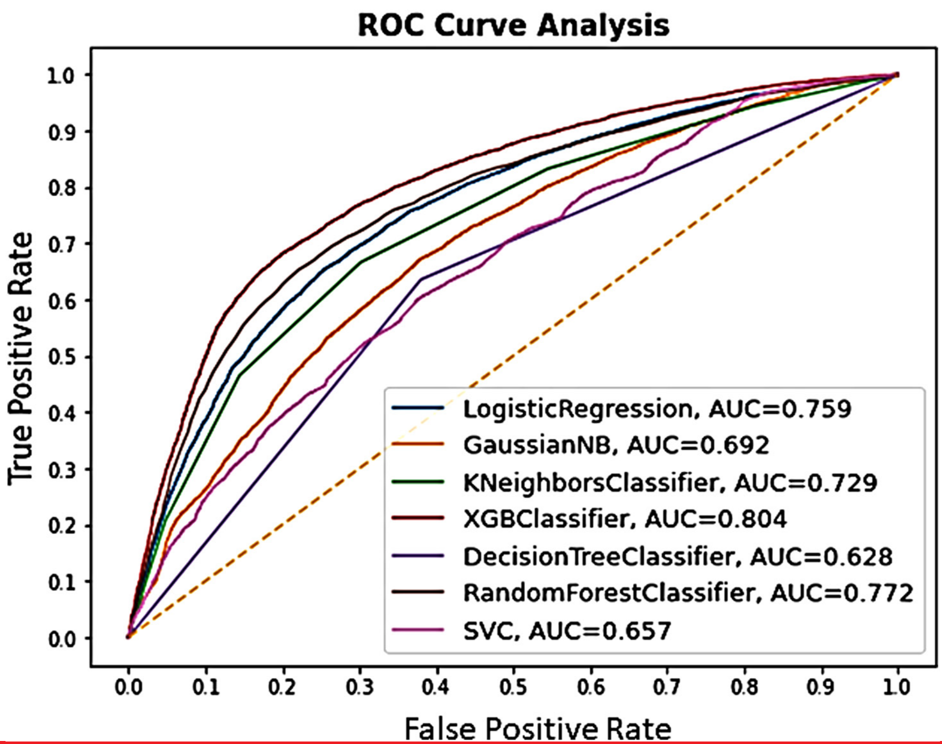

Figure 2. Receiver operating characteristic curves are analyzed using the created algorithms in the test dataset to forecast cardiovascular illnesses. AUC, area under the curve; ROC, receiver operating characteristic.

Table 1. Precision, Recall, F1-Score, Model Accuracy, AUC, and Brier Scores for Developed Algorithms

Models	Precision	Recall	F1-Score	Mean Accuracy	Mean AUC	Mean Brier
Random Forest Classifier	0.71	0.72	0.72	0.712	0.773	0.282
XGBoost Classifier	0.72	0.78	0.75	0.739	0.803	0.260
Decision Tree Classifier	0.63	0.62	0.63	0.629	0.630	0.369
KNeighbors Classifier	0.68	0.70	0.69	0.682	0.728	0.317
SVM Classifier	0.60	0.65	0.62	0.605	0.657	0.394
GaussianNB	0.56	0.87	0.68	0.594	0.692	0.406
Logistic Regression	0.69	0.73	0.71	0.698	0.758	0.301

AUC, area under the curve; XGB, Extreme Gradient Boost.

was the best with 0.8 AUC, 0.26 Brier, and 0.62 mean accuracy values. Among the variables, the most critical variable affecting the diagnosis was systolic blood pressure (ap_hi) (SHAP value = +0.80, PFI weight = 0.1335 ± 0.0045) (Table 2). Similar results were obtained using the PFI (Figure 3), a traditional method supporting the SHAP method.

The main aspects of risk factors and their impacts on the outcome were interpreted using the XAI–SHAP approach. The Shapley plot for interpreting the significance of the variables including beeswarm, summary, and force graphs show that systolic blood pressure, age, and cholesterol are the most important risk factors (Figure 4).

DISCUSSION

In this study, explicable model algorithms were created to determine the relationship between CVD patients and risk factors. For this purpose, ML models were developed to predict patients with CVD, determine the importance of risk factors, and develop new treatment perspectives for clinicians for these risk factors with datasets obtained from the database.

Many models, such as the XGBoost model, which offers the best performance in the study, are affected by data preprocessing, such as algorithms, scaling of numerical features, and coding of categorical features. Data preprocessing is not applied to a common dataset and is reserved for direct training validation and testing. This may have caused the system claimed to give the best results due to the training process affected by the extreme values to have higher

Table 2. Rankings and Importance Weights of the Risk Factors for Predicting Cardiovascular Disease

Ranking	Weight	Risk Factors
1	0.1335 ± 0.0045	ap_hi
2	0.0341 ± 0.0022	Cholesterol
3	0.0211 ± 0.0036	Age
4	0.0038 ± 0.0028	Active
5	0.0029 ± 0.0048	Weight
6	0.0028 ± 0.0027	ap_lo
7	0.0021 ± 0.0014	Gucose
8	−0.0003 ± 0.0010	Alcohol
9	−0.0020 ± 0.0014	Smoking
10	−0.0021 ± 0.0030	Gender
11	−0.0055 ± 0.0021	Height

The risk factors are described as follows: age, the person's age in days; height, the person's height; weight, the person's weight; gender, the person's gender; ap_lo, diastolic blood pressure, ap_hi, systolic blood pressure. Diastolic heart rate cholesterol level: 1, normal; 2, above normal; 3, significantly above normal. Glucose level: 1, normal; 2, above normal; 3, significantly above normal. Alcohol consumption: 0, no; 1, true. Smoking and physical exercise are both 0 (no) and 1 (true).

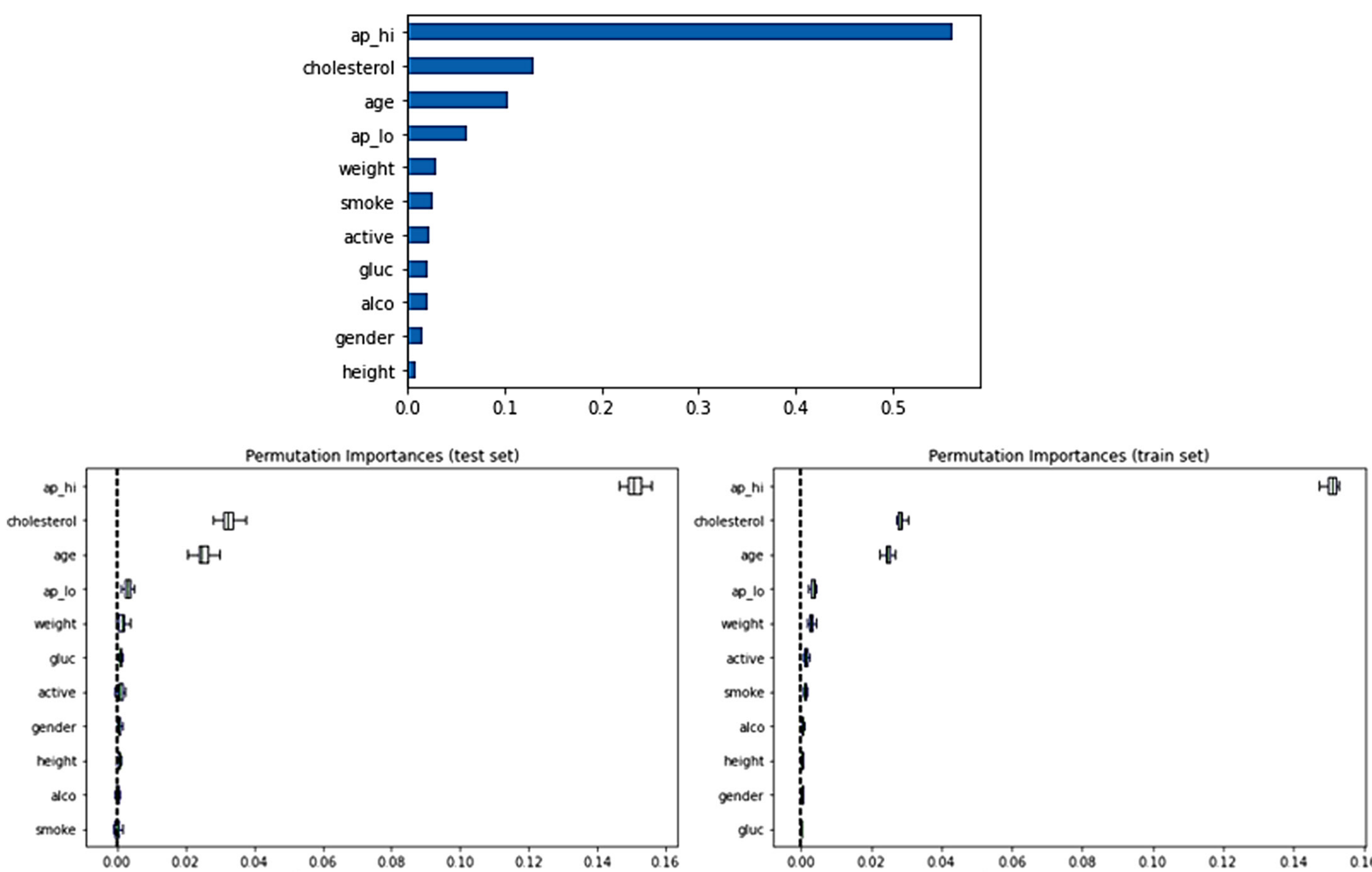

Figure 3. Permutation importance graphs of risk factors according to test and training datasets.

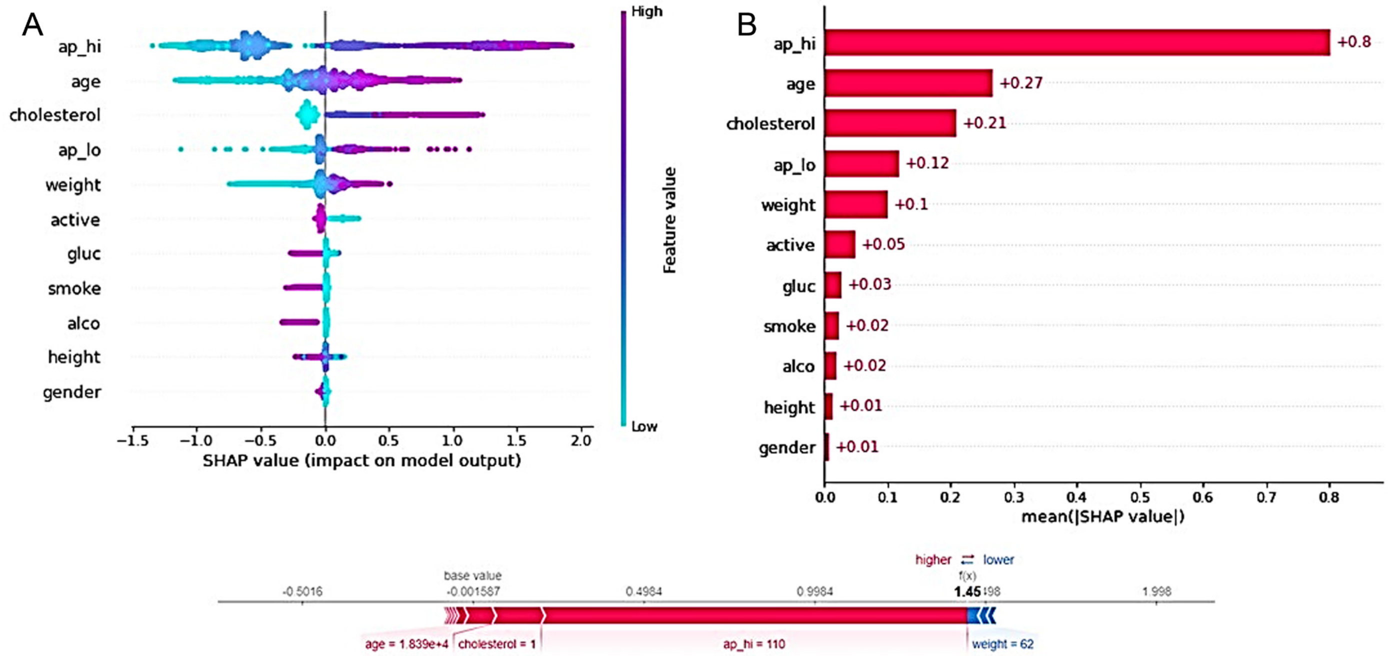

Figure 4. Shapley graphs for prediction. (A) Beeswarm importance plot listing the most significant risk factors. (B) Summary plot of risk factors that impact the prediction model's decision. (C) Force plot showing the risk factors affecting the expected values in the model algorithm.

accuracy levels or to give better results than they do. Data preprocessing should be performed on larger datasets containing CVD risk factors, and it should be considered that it may be better or worse than it should be. Also, the best-performing XGBoost is a popular tree-based learning algorithm, and Tree SHAP is a variant of SHAP specifically designed for annotating tree-based models. Therefore, the TreeSHAP method is used, considering that using Tree SHAP is an effective method to explain the predictions of our XGBoost model.

In the first stage, the feature importance of the risk factors was determined by the classical PFI method based on the training and test data. According to the PFI results, systolic blood pressure, cholesterol, and age were the most critical factors. The PFI method describes which risk factors in the dataset drive the model's performance, while techniques such as SHAP describe which CVD risk factors play a more important role in generating a prediction. Shapley additive explanations is probably one of the cutting-edge products in ML annotation capability. It has a clear interpretation and a solid foundation used in game theory. The Shapley value can be an important method that provides full disclosure, as it is based on solid theory and fairly distributes the effects by calculating the difference between the model's estimate f(x) and the expectation estimate E[f(x)] when the feature value is collected. The disadvantage of SHAP is that the Shapley value requires much computing time. This is because the training period increases exponentially in proportion to the number of traits.

One solution to minimizing computation time is calculating the contributions of only a few instances of possible coalitions.[32] The importance of the permutation property is also linked to the model's error, which is not always what we want. Permutation feature importance is also not well suited for models trained with correlated features, as adding an associated feature can divide the importance between both features, reducing the importance of the related feature. We think that methods such as SHAP, which we used in our study, may be preferred more in studies where risk factors such as CVD constantly develop. Among the models, the XGBoost model with 0.80 AUC had the best prediction model. In the next step, the importance of CVD risk factors and their effects on the prediction model results were determined by applying the XAI–SHAP method to this selected model. The high mortality and disability rate of CVD and many complex risk factors necessitated using explainable models.

Some studies have explained the relationship between risk factors with CVD using different statistical methods.[33-38] When the study's results were examined, it was seen that systolic blood pressure has great importance over CVD. In a study by Whelton et al[39] in 2020, the relationship of normal systolic blood pressure values, as currently defined, with CVD in people without traditional CVD risk factors was examined. Whelton et al[39] emphasized the importance of primary prevention for CVD risk factors of increased systolic blood pressure level, which appears to have similar gradient trajectories in CVD risk within values considered normal in this study.[39] Supporting our results, another study reported that lowering systolic blood pressure below the recommended targets can significantly reduce the risk of CVD-related death.[40] As a result of observational studies, the importance of both the magnitude and the duration of exposure to high systolic blood pressure in the assessment of CVD risk was emphasized.[41,42] Based on the literature support, the final model we predicted in the study could interpret the result with some prognostic variables from a clinical point of view. Although linear regression analysis can be used to estimate the significance of each clinical factor, we thought that the ML method could show relationships between clinical variables and predicted outcomes. In a large dataset, this study identified important clinical factors contributing to predicting risk factors associated with CVDs. This study applied the SHAP method to identify the top-ranking important factors in predicting CVDs across various clinical variables. We have shown that the most important feature is systolic blood pressure in both methods.

Study Limitations

Although it is thought that this study may provide support for other studies related to CVD, it also has some limitations. First, it does not contain information on routine clinical data collection. However, the support of the results by both expert recommendations and the literature shows that the ML model of the study still has high reliability in interpreting important clinical features. The ML model's binary classification performance to predict CVD was another limitation. More data can be used for training and validation, which might lead to variations across groups and more accurate outcomes.

CONCLUSION

This work effectively created an ML model to predict CVD patients and determine the correlation between significant risk variables and CVD patients. With this model, doctors will be able to understand the identification of patients with CVD, the association between CVD and risk factors, and the available treatments for these risk factors more transparently.

Data Availability: While the study data can be accessed at https://data.world/kudem, it can also be accessed by searching "Exploring Risk Factors for Cardiovascular Disease in Adults" on Kaggle.

Ethics Committee Approval: N/A.

Peer-review: Externally peer-reviewed.

Author Contributions: Concept – K.K.K.; Design – K.K.K.; Supervision – E.U.K.; Data Collection and/or Processing – K.K.K.; Analysis and/or Interpretation – K.K.K., E.U.K.; Literature Review – K.K.K.; Writing – K.K.K.; Critical Review – E.U.K.

Declaration of Interests: The authors have no conflict of interest to declare.

Funding: The authors declared that this study has received no financial support.

REFERENCES

1. (WHO) WHO. *Who. Cardiovascular Diseases (CVDs)*. Accessed 11 June, 2021.
2. Barua JD, Omit SBS, Rana HK, Podder NK, Chowdhury UN, Rahman MH. Bioinformatics and system biological approaches for the identification of genetic risk factors in the progression of cardiovascular disease. *Cardiovasc Ther*. 2022;2022:9034996. [CrossRef]
3. Winkleby MA, Jatulis DE, Frank E, Fortmann SP. Socioeconomic status and health: how education, income, and occupation contribute to risk factors for cardiovascular disease. *Am J Public Health*. 1992;82(6):816-820. [CrossRef]
4. Ifeanyi Chiazor E, Evans M, van Woerden H, Oparah AC. A systematic review of community pharmacists' interventions in reducing major risk factors for cardiovascular disease. *Value Health Reg Issues*. 2015;7:9-21. [CrossRef]
5. Al Rawahi AH, Lee P, Al Anqoudi ZAM, et al. Cardiovascular disease incidence and risk factor patterns among Omanis with Type 2 diabetes: A retrospective cohort study. *Oman Med J*. 2017;32(2):106-114. [CrossRef]
6. Ren J, Zhang Y. Targeting autophagy in aging and aging-related cardiovascular diseases. *Trends Pharmacol Sci*. 2018;39(12):1064-1076. [CrossRef]
7. Rudnicki M, Haas TL. Exploring risk factors at the molecular level. *eLife*. 2021;10:e68271. [CrossRef]
8. Khosravi A, Akhavan Tabib A, Golshadi I, et al. The relationship between weight and CVD risk factors in a sample population from central Iran (based on IHHP). *ARYA Atheroscler*. 2012; 8(2):82-89.
9. Sandberg C, Rinnström D, Dellborg M, et al. Height, weight and body mass index in adults with congenital heart disease. *Int J Cardiol*. 2015;187:219-226. [CrossRef]
10. Carson JAS, Lichtenstein AH, Anderson CAM, et al. Dietary cholesterol and cardiovascular risk: A science advisory from the American Heart Association. *Circulation*. 2020;141(3):e39-e53. [CrossRef]
11. Jeong SM, Choi S, Kim K, et al. Effect of change in total cholesterol levels on cardiovascular disease among young adults. *J Am Heart Assoc*. 2018;7(12) [CrossRef]
12. Lopez-Jaramillo P, Gomez-Arbelaez D, Martinez-Bello D, et al. Association of the triglyceride glucose index as a measure of insulin resistance with mortality and cardiovascular disease in populations from five continents (PURE study): a prospective cohort study. *Lancet Healthy Longev*. 2023;4(1):e23-e33. [CrossRef]
13. Tong Y, Liu F, Huang K, et al. Changes in fasting blood glucose status and incidence of cardiovascular disease: the China-PAR project. *J Diabetes*. 2023;15(2):110-120. [CrossRef]
14. Hoek AG, van Oort S, Mukamal KJ, Beulens JWJ. Alcohol consumption and cardiovascular disease risk: placing new data in context. *Curr Atheroscler Rep*. 2022;24(1):51-59. [CrossRef]
15. van de Luitgaarden IAT, van Oort S, Bouman EJ, et al. Alcohol consumption in relation to cardiovascular diseases and mortality: a systematic review of Mendelian randomization studies. *Eur J Epidemiol*. 2022;37(7):655-669. [CrossRef]
16. Krittanawong C, Isath A, Rosenson RS, et al. Alcohol consumption and cardiovascular health. *Am J Med*. 2022;135(10):1213-1230.e3. [CrossRef]
17. Maddali MV, Churpek M, Pham T, et al. Validation and utility of ARDS subphenotypes identified by machine-learning models using clinical data: an observational, multicohort, retrospective analysis. *Lancet Respir Med*. 2022;10(4):367-377. [CrossRef]
18. Naimark M, Zhang AL, Leon I, Trivellas A, Feeley BT, Ma CB. Clinical, radiographic, and surgical presentation of subscapularis tendon tears: A retrospective analysis of 139 patients. *Arthroscopy*. 2016;32(5):747-752. [CrossRef]
19. Sengupta PP, Shrestha S, Kagiyama N, et al. A machine-learning framework to identify distinct phenotypes of aortic stenosis severity. *JACC Cardiovasc Imaging*. 2021;14(9):1707-1720. [CrossRef]

20. Li C, Alike Y, Hou J, et al. Machine learning model successfully identifies important clinical features for predicting outpatients with rotator cuff tears. *Knee Surg Sports Traumatol Arthrosc*. 2023;31(7):2615-2623. [CrossRef]
21. Al'Aref SJ, Anchouche K, Singh G, et al. Clinical applications of machine learning in cardiovascular disease and its relevance to cardiac imaging. *Eur Heart J*. 2019;40(24):1975-1986. [CrossRef]
22. Ghosh P, Azam S, Jonkman M, et al. Efficient prediction of cardiovascular disease using machine learning algorithms with relief and lasso feature selection techniques. *IEEE Access*. 2021;9:19304-19326. [CrossRef]
23. Aryal S, Alimadadi A, Manandhar I, Joe B, Cheng X. Machine learning strategy for gut microbiome-based diagnostic screening of cardiovascular disease. *Hypertension*. 2020;76(5):1555-1562. [CrossRef]
24. Zhang Y, Weng Y, Lund J. Applications of explainable artificial intelligence in diagnosis and surgery. *Diagnostics (Basel)*. 2022;12(2):237. [CrossRef]
25. van der Velden BHM, Kuijf HJ, Gilhuijs KGA, Viergever MA. Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. *Med Image Anal*. 2022;79:102470. [CrossRef]
26. Wolfe JC, Mikheeva LA, Hagras H, Zabet NR. An explainable artificial intelligence approach for decoding the enhancer histone modifications code and identification of novel enhancers in Drosophila. *Genome Biol*. 2021;22(1):308. [CrossRef]
27. Altmann A, Toloşi L, Sander O, Lengauer T. Permutation importance: a corrected feature importance measure. *Bioinformatics*. 2010;26(10):1340-1347. [CrossRef]
28. Breiman L. Random forests. *Mach Learn*. 2001;45(1):5-32. [CrossRef]
29. Shapley L. *A Value for n-Person Games*. Princeton; Princeton University Press; 1953:307-317. [CrossRef]
30. Lundberg S, Lee S-I. *"A unified approach to interpreting model predictions," presented at the Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, California, USA, 2017*. Model Predictions; 2017.
31. Kamath U, Liu J. Introduction to interpretability and explainability. In: Kamath U, Liu J, eds. *Explainable Artificial Intelligence: an Introduction to Interpretable Machine Learning*. Cham: Springer International Publishing; 2021:1-26.
32. Molnar C. *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*; 2022. https://leanpub.com/interpretable-machine-learning
33. Keto J, Ventola H, Jokelainen J, et al. Cardiovascular disease risk factors in relation to smoking behaviour and history: a population-based cohort study. *Open Heart*. 2016;3(2):e000358. [CrossRef]
34. Williams PT, Fortmann SP, Farquhar JW, Varady A, Mellen S. A comparison of statistical methods for evaluating risk factor changes in community-based studies: an example from the Stanford Three-Community Study. *J Chronic Dis*. 1981;34(11):565-571. [CrossRef]
35. Iqbal R, Ahmad Z, Malik F, et al. A statistical analysis of hypertension as cardiovascular risk factor. *Middle East J Sci Res*. 2012;12(1):19-22.
36. Din S, Rabbi F, Qadir F, Khattak M. Statistical analysis of risk factors for cardiovascular disease in Malakand division. *Pak J Stat Oper Res*. 2007;8:117-124.
37. Matsushita K, Kaptoge S, Hageman SHJ, et al. Including measures of chronic kidney disease to improve cardiovascular risk prediction by SCORE2 and SCORE2-OP. *Eur J Prev Cardiol*. 2023;30(1):8-16. [CrossRef]
38. Thomas HJ, Marsh CE, Lester L, Maslen BA, Naylor LH, Green DJ. Sex differences in cardiovascular risk factor responses to resistance and endurance training in younger subjects. *Am J Physiol Heart Circ Physiol*. 2023;324(1):H67-H78. [CrossRef]
39. Whelton SP, McEvoy JW, Shaw L, et al. Association of normal systolic blood pressure level with cardiovascular disease in the absence of risk factors. *JAMA Cardiol*. 2020;5(9):1011-1018. [CrossRef]
40. Bundy JD, Li C, Stuchlik P, et al. Systolic blood pressure reduction and risk of cardiovascular disease and mortality: A systematic review and network meta-analysis. *JAMA Cardiol*. 2017;2(7):775-781. [CrossRef]
41. He J, Whelton PK. Elevated systolic blood pressure and risk of cardiovascular and renal disease: overview of evidence from observational epidemiologic studies and randomized controlled trials. *Am Heart J*. 1999;138(3 Pt 2):211-219. [CrossRef]
42. Wang N, Harris K, Hamet P, et al. Cumulative systolic blood pressure load and cardiovascular risk in patients with diabetes. *J Am Coll Cardiol*. 2022;80(12):1147-1155. [CrossRef]