IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

Leveraging Clinical Data for Machine-Learning-Based Heart Detection

Arth Bhardwaj, Janani Sekar

1) St. Francis High School, 1885 Miramonte Ave, Mountain View, CA 94040, USA; arthbhardwaj1234@gmail.com
2) Harvard University, Cambridge, MA 02138, USA

ABSTRACT: This study uses machine learning to evaluate demographics, cholesterol, and cardiovascular outcomes. We investigated cholesterol's prognostic power in two settings: incident heart disease and death among cardiac patients. Two datasets were used: UCI Cleveland (n=303; yes/no heart disease) and Faisalabad Institute of Cardiology, Pakistan (n=299; longitudinal heart failure outcomes). We hypothesized that elevated cholesterol would better detect new heart disease cases than in predicting mortality among individuals already diagnosed. Logistic regression and multilayer perceptron models were trained; AUC, recall, accuracy, and precision were used for evaluation. In Cleveland, a cholesterol-only logistic regression predicted heart disease with 53% accuracy; multivariable models adding age, ST-segment depression, and maximum heart rate reached 85% accuracy (AUC 0.91). In Faisalabad, cholesterol predicted death with 78% accuracy; after adding demographics and comorbidities, the multilayer perceptron achieved 97% accuracy, and logistic regression 82%. Best logistic models indicated diabetics had 2.6× higher death risk and that mortality increased by ~62% per unit of cholesterol. This comparative study shows that machine learning improves cardiovascular risk assessment in data-limited settings and that risk variables vary across illness stages.

KEYWORDS: Medical and Health Sciences, Cardiology, Machine Learning, Cholesterol, Risk Stratification.

Introduction

Nearly 18 million people die annually from cardiovascular disease (CVD), a significant cause of death worldwide. Thus, identifying and understanding the primary CVD risk factors is essential.[1] Cholesterol is one of numerous biochemical indicators, although its role in CVD is well known. High LDL cholesterol causes arterial plaques and atherosclerosis. These factors reduce blood flow and increase the risk of myocardial infarction and stroke. However, HDL cholesterol removes excess cholesterol from the circulation, preventing cardiovascular events.[2–4]

The risk of cardiovascular disease is not limited to high cholesterol. Age, sex, blood pressure, BMI, smoking, diabetes, and physical exercise interact complexly to produce cardiovascular disease.[5–7] These effects are worsened by urban stress, inactivity, and poor diet.[8,9] Numerous studies have demonstrated that smoking and diabetes considerably increase cardiovascular morbidity and mortality.

Physicians now use AI and ML to analyze large datasets, improving diagnosis and patient outcomes.[10] Machine learning algorithms can handle multiple risk factors at once, detecting subtle, non-linear relationships that statistical models miss. In the UCI Cleveland Heart Disease dataset, one of the most widely used benchmarks, clinical and demographic characteristics predict heart disease well.[11,12] ML approaches can increase cardiovascular risk-assessment accuracy and personalization, expanding on conventional instruments like the Framingham Risk Score.[13] ML-driven models must balance interpretability, prediction performance, transparency, explainability, and clinical actionability.

This study on cholesterol's predictive role in cardiovascular events aims to improve risk classification by adding clinical and demographic characteristics.[14] We wanted to identify people at risk of heart disease, predict mortality in cardiovascular patients, and compare the predictive power of cholesterol levels in these two contexts to see if they were enough. We hypothesized that elevated cholesterol levels would detect incident heart disease better than fatal cases. Our novel strategy separates variables affecting illness, beginning from mortality progression, to better understand the intricate web of processes that lead to cardiovascular risk.

We tested our hypotheses with two datasets. First was the UCI Cleveland Heart Disease dataset, which included 303 American patients with yes/no heart disease results. The Pakistani Faisalabad Heart Failure dataset includes 299 patients' longitudinal data, including mortality outcomes. We examined demographic and endpoint prediction tendencies using these datasets. With logistic regression and MLP models, survival and classification results may be assessed. In addition to cholesterol levels, we examined demographics, vital signs, comorbidities, and lifestyle factors for predictive potential.

For illness incidence and death, multivariate models with age, sex, blood pressure, comorbidities, and other clinical factors outperformed cholesterol-only models. Contrary to our expectations, cholesterol alone was better at distinguishing patients with and without a history of mortality than at predicting incident heart disease. Once the disease has progressed, cholesterol and other risk factors may be strong prognostic indicators.

This study shows how machine learning may improve tailored predictions and how crucial multidimensional risk modelling is for cardiovascular treatment. The work improves our under-

DOI: 10.36838/IJHSR88.56

standing of cardiovascular disease risk factors. It emphasises the need to include clinical, demographic, and biochemical data in prediction models by distinguishing disease-causing variables from death-causing factors. Such methods can inform targeted therapies and tailored management regimens in diverse and resource-constrained healthcare settings. This study reinforces the therapeutic value of combining classical markers like cholesterol with cutting-edge computational methods to improve patient care.

Methods

We compared the Faisalabad Institute of Cardiology's Pakistani clinical dataset and the University of California, Irvine's Cleveland Heart Disease dataset for cardiovascular outcome prediction. Since both datasets included demographic and clinical information, machine-learning predictive modeling was possible. We tested logistic regression and multilayer perceptron (MLP) neural networks to predict heart disease and death. Comparing linear and nonlinear modeling methods allowed us to evaluate their ability to capture complex risk-factor–outcome relationships.

We evaluated all datasets for completeness before modeling. Preprocessing was relatively straightforward in the UCI Cleveland dataset because most variables were complete; however, a small number of cholesterol entries are recorded as 0, which may represent missing values in the original source, and is discussed as a limitation. To ensure predictive analysis validity in the more complex and heterogeneous Faisalabad dataset, similar filters were employed to maintain only complete records. For machine-learning training, one-hot encoding transformed categorical data, including sex, chest-pain type, smoking status, and comorbidities. Standardizing continuous variables—heart rate, blood pressure, cholesterol, and age—to zero mean and unit variance enhanced model convergence and interpretability.

We evaluated the models using an 80/20 train–test split to verify that the test set appropriately reflected fresh data. Class balance was preserved by stratifying this split by outcome to reduce performance-metric bias. We selected the most informative variables through iterative feature selection using domain expertise and performance-guided testing. To assess prediction accuracy, we examined cholesterol-only models and models that gradually added demographic and clinical variables—Table 1. Summary of models, input features, and predicted outcomes for each dataset.

Two main models were created. Logistic regression provided a linear baseline for coefficient interpretation and revealed predictor relevance. Multilayer perceptrons—feedforward neural networks with hidden layers—can express nonlinear interactions and complicated correlations.

We examined how well algorithms predicted outcomes across the two datasets and how much cholesterol alone affected mortality and heart-disease risk relative to a multidimensional feature collection. This approach showed how machine learning can improve cardiovascular risk assessment by evaluating the utility of additional demographic and clinical data.

Table 1: Comparison of input variables and performance metrics for logistic regression (LR) and multilayer perceptron (MLP) models across the Cleveland and Faisalabad cohorts. Models with more features significantly outperformed cholesterol-only baselines: in Cleveland, multivariable LR achieved 85% accuracy (AUC 0.91), and in Faisalabad, the MLP achieved 97 % accuracy for mortality prediction—showing the power of multidimensional clinical data.

Model	Dataset	Input Features	Predicted Outcome
Logistic Regression	UCI Cleveland	Cholesterol (mg/dL)	Heart Disease
Logistic Regression	UCI Cleveland	Age (years), Sex, Resting Blood Pressure (mmHg), Cholesterol (mg/dL), Maximum Heart Rate (bpm), ECG ST Segment Depression (mm), Resting ECG, ECG ST Slope Value, Chest Pain, Exercise Angina	Heart Disease
Logistic Regression	Faisalabad	Cholesterol (mg/dL)	Mortality
Logistic Regression	Faisalabad	Age (years), Sex, Urban, Marital Status, Resting Blood Pressure (mmHg), Cholesterol (mg/dL), Maximum Heart Rate (bpm), ECG ST Segment Depression (mm), Resting ECG, ECG ST Slope Value, Chest Pain, Exercise Angina	Mortality
Multilayer Perceptron (Hidden Layers: [20, 20])	Faisalabad	Age (years), Sex, Urban, Marital Status, Resting Blood Pressure (mmHg), Cholesterol (mg/dL), Maximum Heart Rate (bpm), ECG ST Segment Depression (mm), Resting ECG, ECG ST Slope Value, Chest Pain, Exercise Angina, Diabetes, Smoking, Allergies, Family History	Mortality

Table 1 summarizes the study's datasets—UCI Cleveland and Faisalabad—machine-learning models, input attributes, and anticipated outcomes. As shown, model complexity and feature richness affected predictive performance for incident heart disease in the UCI dataset and mortality in the Faisalabad dataset. Initial models with cholesterol as the only predictor tested cholesterol's independent effect on risk prediction.

The efficacy of cholesterol as a sole outcome discriminator was assessed in these preliminary analyses using logistic regression and MLP models.

Age, sex, blood pressure, maximum heart rate, electrocardiogram (ECG) readings, chest-pain type, exercise-induced angina, and comorbidities, including diabetes and smoking status, were then added to the models. Using this multivariable strategy to measure each component's predictive value helped us build more accurate risk models. To ensure interpretability, logistic regression coefficients were used to calculate odds ratios (ORs) for each predictor. Using a 2×2 contingency table, ORs were calculated to investigate the relationship between cholesterol and heart disease. In our calculation, A and B represent high-cholesterol groups with and without heart disease, and C and D represent normal-cholesterol groups with and without heart disease; the OR was computed as $(A \times D) / (B \times C)$. A larger value indicates a higher likelihood of heart disease associated with elevated cholesterol.

We tested neural-network models with one, two, or three hidden layers and varying neuron counts. A grid search revealed that a two-hidden-layer network with 20 neurons per layer and ReLU activation performed best. Classification networks were trained using the Adam optimizer and binary cross-entropy loss. Training was discontinued when validation performance stopped improving to avoid overfitting. Using

feature selection, model modification, and regularization, we maximized expected accuracy and model generalizability.

We learned more about the multidimensional characteristics' additive value and the intricate, likely nonlinear relationships between clinical variables and cardiovascular outcomes by comparing linear logistic regression and nonlinear MLPs.

Results and Discussion

We tested the predictive power of cholesterol and other clinical parameters using logistic regression (LR) and multilayer perceptron (MLP) models trained on the UCI Cleveland and Faisalabad datasets. The original baseline model recognized heart disease in the UCI Cleveland dataset using cholesterol as a predictor with 53% accuracy. Although cholesterol is a helpful biomarker, this suggests it is not a strong diagnostic indicator alone. To improve prediction performance, we added age, maximum heart rate, ST-segment depression, resting blood pressure, and additional clinical variables. With this larger set, LR accuracy reached 85%, demonstrating how multidimensional clinical data improve risk assessment. In addition, the extended LR model obtained an AUC = 0.91, indicating accurate and consistent discrimination across thresholds.

The MLP model, which captures nonlinear variable interactions, achieved 84% accuracy with AUC = 0.88. According to these findings, adding several relevant features to either a linear or nonlinear model enhances heart-disease prediction. LR provides interpretability via odds ratios and coefficients, while MLPs model complex interactions. Figure 1 illustrates the frequency distributions of key variables in the Cleveland dataset.

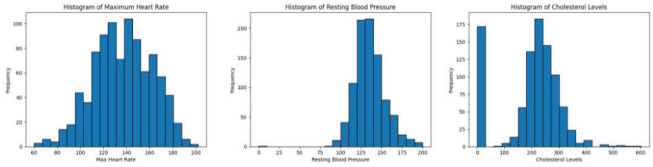

Figure 1: Histograms and density plots show distributions of key numeric variables in the Cleveland dataset. Because cholesterol(md/dl) alone yielded only 53% accuracy, these balanced distributions justify including additional variables such as age(years), ST-depression, and max heart rate, enabling an AUC of 0.91 with the multivariable LR model.

The cholesterol, resting blood pressure, and maximum heart rate histograms were approximately normal, while resting blood pressure showed a slight right skew. Skewed distributions can change the relative importance of variables in linear models; therefore, they should be examined carefully during model building and clinical interpretation. To ensure feature-representation uniformity and cross-dataset comparability, Figure 2 shows the Faisalabad dataset distributions. Visualizing these distributions helps explain population-specific demographic and clinical factors that affect model generalizability. Overall, these findings suggest that multidimensional machine-learning methods are beneficial in clinical practice and that adding clinical characteristics to cholesterol enhances cardiovascular risk prediction.

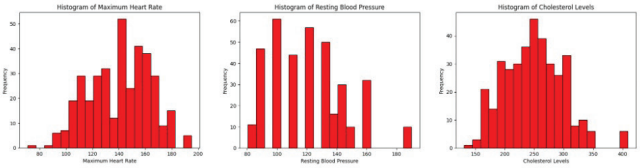

Figure 2: Visualization of key feature distributions in the Faisalabad cardiology dataset. Cohort differences visible here correspond to markedly higher mortality-prediction accuracy (up to 97%) once comorbidities and demographics are included, underscoring the impact of richer clinical features on prognostic models.

In the Faisalabad dataset, cholesterol-only models predicted mortality outcomes with 78% accuracy in cardiovascular disease patients. Adding demographic and comorbidity-related factors such as age, diabetes, smoking, and urban dwelling improved predictive performance. By combining multidimensional clinical data, LR achieved 82% accuracy. The MLP captured complex nonlinear interactions among variables with 97% accuracy, as seen in Figure 7. These findings show that machine learning can improve prognostic evaluations beyond single-variable models and that high-risk populations require comprehensive feature sets.

Outcome-segregated feature distributions support these findings. Figures 3 and 4 show that deceased patients were older and had higher cholesterol, while survivors tended to have higher maximum heart rates. These trends imply that cholesterol is a meaningful predictor on its own but becomes considerably stronger when combined with demographic and clinical characteristics. These results suggest that neural networks and other multidimensional modeling methods can improve risk classification and personalized treatment regimens by predicting poor cardiovascular outcomes in diverse patient groups. Importantly, these visual shifts are clinically plausible and help interpret why certain predictors receive larger learned weights in the multivariable models.

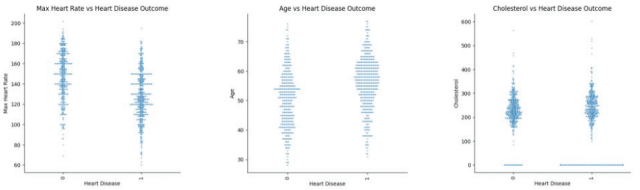

Figure 3: Boxplots showing how age (years), cholesterol (mg/dl), and max heart rate (bpm) differ between healthy and diseased patients. Individuals with disease are generally older, have higher cholesterol, and lower max heart rate—patterns that allowed LR to improve from 53 % (cholesterol-only) to 85 % accuracy with multiple predictors.

Figure 4: Heart-rate(bpm) response, matching the model's learned coefficients. Diabetes increased risk approximately 2.6×, explaining the superior performance of multifeature MLP models.

DOI: 10.36838/IJHSR88.56

To understand variable relationships, we examined correlations with age (Figures 5 and 6). We evaluated the statistical significance of individual predictors and model performance across outcome groups using one-way ANOVA for continuous variables (cholesterol, age) and chi-square tests for categorical variables (smoking, diabetes). Paired t-tests compared cholesterol-only and multivariable configurations across five cross-validation rounds. All tests were two-tailed with $\alpha = 0.05$. Including demographic and clinical parameters beyond cholesterol significantly improved prediction performance in both LR ($p = 0.002$) and MLP ($p < 0.001$).

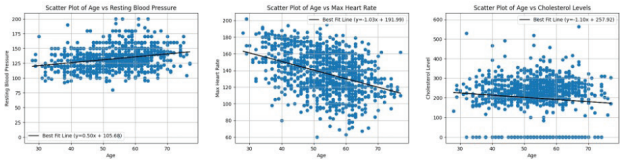

Figure 5: Scatter plots showing relationships of age with resting blood pressure (mmHg, cholesterol (mg/dl), and max heart rate (bpm). The upward trend in BP/cholesterol and decline in heart rate with age produced an additive predictive signal, helping LR reach AUC 0.91 when age was combined with ECG and exercise-related features.

Figure 6: Analogous age-related trends for the Faisalabad cohort. The synergy of age with comorbidities such as diabetes and smoking strengthened mortality prediction, explaining why nonlinear MLPs achieved the highest prognostic accuracy.

An odds-ratio (OR) analysis assessed these links. Diabetics had 2.61× higher risk of death, and mortality increased by ~62% per unit of cholesterol. In addition, higher ECG ST-segment slope values approximately doubled to tripled the risk of death. In this cohort, top mortality models had precision and recall above 0.90, supporting reliable identification of high-risk patients. In clinical practice, risk markers beyond cholesterol should therefore be considered when making prognostic predictions.

Comparison with prior work: Our Cleveland results are consistent with previous machine-learning studies that report substantial gains when moving from single-biomarker baselines to multivariable models on the UCI Cleveland benchmark.[11,12] Likewise, prior survival-prediction work in cardiology has shown that combining clinical variables can markedly improve mortality-risk stratification, aligning with our finding that multidimensional models outperform cholesterol-only approaches in the Faisalabad cohort. Because our goal was to compare cholesterol's role across two clinical contexts rather than exhaustively optimize algorithms, we limited model families to logistic regression and an MLP. Future work can extend this comparison to additional model classes (e.g., random forests, support vector machines, gradient boosting) using the same train/test protocol for fair benchmarking.

Finally, ROC curves summarize discrimination for both datasets and models (Figure 7). On Cleveland, LR AUC = 0.91 and MLP AUC = 0.88; on Faisalabad, LR AUC = 0.77 and MLP AUC = 1.00; models were evaluated on held-out test splits across five runs. Overall, results partly revise our initial expectation: cholesterol alone was more prognostic for mortality (Faisalabad) than diagnostic for incident disease (Cleveland). However, adding broader clinical information consistently improved performance for both endpoints. LR offers clinician-friendly interpretability via coefficients and ORs, while MLPs can capture complex, nonlinear relations and deliver higher accuracy for survival prediction. A dual-model strategy—transparent LR for screening and higher-capacity MLPs for prognosis—can balance interpretability and performance across clinical contexts.

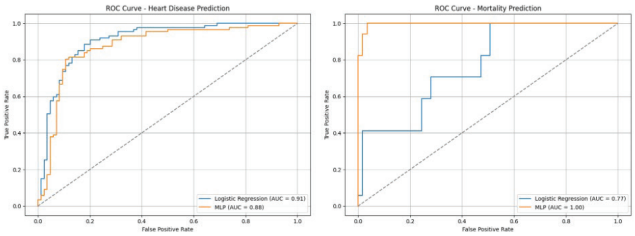

Figure 7: Receiver-Operating-Characteristic (ROC) curves comparing LR and MLP across both datasets. LR slightly outperformed MLP for incident disease in Cleveland (AUC 0.91 vs 0.88), whereas MLP dominated for mortality in Faisalabad (AUC 1.00 vs 0.77). This highlights a dual-model strategy: interpretable LR for screening and high-capacity MLP for prognosis.

Conclusion

Logistic regression and MLP models are effective machine-learning tools for predicting cardiovascular outcomes using clinical, demographic, and cholesterol data. To compare the predictive power of multidimensional feature sets with cholesterol alone, we used the Faisalabad Institute of Cardiology dataset for mortality and the UCI Cleveland Heart Disease dataset for incident heart disease. Our data show that cholesterol alone cannot reliably predict heart disease; in the Cleveland sample, it achieved a moderate accuracy of 53%. By adding age, sex, blood pressure, maximum heart rate, ST-segment depression, comorbidities, and lifestyle factors, logistic regression reached AUC = 0.91 and 85% accuracy.

In the Faisalabad cohort of previously diagnosed patients, cholesterol alone predicted mortality with 78% accuracy. Adding demographic and clinical data increased logistic-regression accuracy to 82% and MLP accuracy to 97%. Diabetes raised mortality risk 2.61×, while mortality increased by ~62% per unit of cholesterol; higher ECG ST-segment slope values were also associated with increased risk. These findings emphasize multidimensional risk models for cardiovascular disease diagnosis and prognosis. Logistic regression provides interpretability and clinical reasoning through odds ratios and coefficients, while MLP models improve survival-prediction accuracy and capture nonlinear interactions. This trade-off supports a two-model approach: transparent LR for initial screening and higher-capacity MLPs for individualized prognostication.

This study provides cardiovascular risk-stratification insights despite limitations in dataset size, single-center sam-

pling, and the absence of lifestyle detail and lipid subtypes. Future work should incorporate temporal modeling, additional biochemical and lifestyle markers, and multi-site cohorts. Overall, combining traditional biomarkers like cholesterol with modern computational methods can enable more accurate, resource-aware care and better patient outcomes.

Limitations and Future Work

This study provides cardiovascular risk-stratification insights despite limitations in dataset size, single-center sampling for the Faisalabad cohort, and the absence of lifestyle detail, lipid subfractions (e.g., LDL/HDL), and medication information. Additionally, a small number of cholesterol values recorded as 0 in the Cleveland dataset may reflect the absence in the source data and could introduce bias in cholesterol-only analyses. To avoid overgeneralization, the reported performance should be interpreted as a comparative benchmark within these cohorts rather than a universally generalizable clinical tool.

Future work should incorporate external validation on independent cohorts, stronger validation practices (e.g., nested cross-validation and calibration analysis), and richer biochemical and lifestyle markers. Where available, future analyses should replace total cholesterol with full lipid panels and evaluate sensitivity to handling of zero/placeholder values (e.g., treating them as missing and applying imputation or exclusion). Finally, additional model families (e.g., random forests, SVMs, gradient boosting) can be compared under the same evaluation protocol to assess whether performance improvements are robust across algorithms.

Acknowledgments

The author extends sincere gratitude to his mentors and teachers for their invaluable guidance throughout this research. The author also thanks Saint Francis High School, Mountain View, California, for providing academic support and resources that made this study possible.

References

1. Roth, G. A.; Mensah, G. A.; Johnson, C. O.; Addolorato, G.; Ammirati, E.; Baddour, L. M.; Barengo, N. C.; Beaton, A. Z.; Benjamin, E. J.; Benziger, C. P.; Bonny, A. Global Burden of Cardiovascular Diseases and Risk Factors, 1990–2019: Update from the GBD 2019 Study. *J. Am. Coll. Cardiol.* **2020**, *76* (25), 2982–3021. https://doi.org/10.1016/j.jacc.2020.11.010
2. Vourakis, M.; Mayer, G.; Rousseau, G. The Role of Gut Microbiota on Cholesterol Metabolism in Atherosclerosis. *Int. J. Mol. Sci.* **2021**, *22* (15), 8074. https://doi.org/10.3390/ijms22158074
3. Bentzon, J. F.; Otsuka, F.; Virmani, R.; Falk, E. Mechanisms of Plaque Formation and Rupture. *Circ. Res.* **2014**, *114* (12), 1852–1866. https://doi.org/10.1161/CIRCRESAHA.114.302721
4. Casula, M.; Colpani, O.; Xie, S.; Catapano, A. L.; Baragetti, A. HDL in Atherosclerotic Cardiovascular Disease: In Search of a Role. *Cells* **2021**, *10* (8), 1869. https://doi.org/10.3390/cells10081869
5. Mozaffarian, D.; Afshin, A.; Benowitz, N. L.; Bittner, V.; Daniels, S. R.; Franch, H. A.; Jacobs, D. R., Jr.; Kraus, W. E.; Kris-Etherton, P. M.; Krummel, D. A.; Popkin, B. M.; Whitsel, L. P.; Zakai, N. A. Population Approaches to Improve Diet, Physical Activity, and Smoking Habits: A Scientific Statement from the American Heart Association. *Circulation* **2012**, *126* (12), 1514–1563. https://doi.org/10.1161/CIR.0b013e318260a20b
6. Enkhmaa, B.; Berglund, L. Lipoprotein(a) and Effects of Diet: Time for Reassessment. *Nutrients* **2025**, *17* (10), 1714. https://doi.org/10.3390/nu17101714
7. Quispe, R.; Elshazly, M. B.; Zhao, D.; Toth, P. P.; Puri, R.; Virani, S. S.; Blumenthal, R. S.; Martin, S. S.; Jones, S. R.; Michos, E. D. Total Cholesterol/HDL-Cholesterol Ratio Discordance with LDL-Cholesterol and Non-HDL-Cholesterol and Incidence of Atherosclerotic Cardiovascular Disease in Primary Prevention: The ARIC Study. *Eur. J. Prev. Cardiol.* **2020**, 27 (15), 1597–1605. https://doi.org/10.1177/2047487319862401
8. Cushman, M.; Shay, C. M.; Howard, V. J.; Jiménez, M. C.; Lewey, J.; McSweeney, J. C.; Newby, L. K.; Poudel, R.; Reynolds, H. R.; Rexrode, K. M.; Sims, M. Ten-Year Differences in Women's Awareness Related to Coronary Heart Disease: Results of the 2019 American Heart Association National Survey. *Circulation* **2021**, *143* (7), e239–e248. https://doi.org/10.1161/CIR.0000000000000907
9. Kopčeková, J.; Holovičová, M.; Gažarová, M.; Mrázová, J.; Habánová, M.; Mečiarová, L.; Bronkowska, M. Association between Selected Dietary Habits and Lipid Profiles of Patients with Cardiovascular Disease. *Int. J. Environ. Res. Public Health* **2020**, *17* (20), 7605. https://doi.org/10.3390/ijerph17207605
10. Chicco, D.; Jurman, G. Machine Learning Can Predict Survival of Patients with Heart Failure from Serum Creatinine and Ejection Fraction Alone. *BMC Med. Inform. Decis. Mak.* **2020**, *20*, 16. https://doi.org/10.1186/s12911-020-1023-5
11. Yahaya, L.; Oye, N. D.; Adamu, A. Performance Analysis of Some Selected Machine Learning Algorithms on Heart Disease Prediction Using the Noble UCI Datasets. *Int. J. Eng. Appl. Sci. Technol.* **2020**, *5* (1), 36–46.
12. Varun, L. S.; Saivardhini, M.; Srivarsha, P.; Balaram, V. V. Predictive Analytics for Heart Disease Using Machine Learning. *Int. J. Adv. Res. Appl. Sci. Eng. Technol. (IJARASEM)* **2024**, *11* (6).
13. D'Agostino, R. B., Sr.; Vasan, R. S.; Pencina, M. J.; Wolf, P. A.; Cobain, M.; Massaro, J. M.; Kannel, W. B. General Cardiovascular Risk Profile for Use in Primary Care: The Framingham Heart Study. *Circulation* **2008**, *117* (6), 743–753. https://doi.org/10.1161/CIRCULATIONAHA.107.699579
14. Kontush, A. *Cholesterol, Lipoproteins, and Cardiovascular Health: Separating the Good (HDL), the Bad (LDL), and the Remnant; Wiley*: Hoboken, NJ, **2024**. https://doi.org/10.1002/9781394158393

Authors

Arth Bhardwaj is a research collaborator with the University of Illinois Urbana-Champaign and a Saint Francis High School student in Mountain View, California. His research integrates machine learning, huge language models, and cybersecurity systems to study AI safety, automation, and ethical deployment. Arth co-authored an IEEE threat intelligence conference article, demonstrating his technical expertise and interest in new security paradigms. He continues to work on durable, transparent, and secure AI-driven infrastructures for society.

DOI: 10.36838/IJHSR88.56