Asian J Biol Life Sci., 2026; 15(1):32-36.
https://ajbls.com

Review Article

Application of Machine Learning Algorithms for Early Prediction of Cardiovascular Diseases through Laboratory Biomarkers

Prem Kumar Essgir*, Mulavagili Vijaya Simha

Department of Medical Lab Technology, SunRise University, Alwar, Rajasthan, INDIA.

ABSTRACT

Cardiovascular Diseases (CVDs) remain the leading cause of morbidity and mortality worldwide, accounting for a significant proportion of global deaths each year. Early detection and risk stratification are crucial for preventing disease progression and reducing mortality rates. Traditional diagnostic methods rely on clinical risk scores and laboratory investigations; however, these approaches may fail to capture complex nonlinear relationships between multiple biological variables. Machine Learning (ML), a subset of artificial intelligence, has emerged as a powerful tool capable of analyzing large datasets and identifying hidden patterns that contribute to disease prediction. Recent studies have demonstrated that ML algorithms integrated with laboratory biomarkers can significantly improve the early detection and prognosis of cardiovascular diseases. Biomarkers such as cardiac troponins, C-reactive Protein (CRP), lipoproteins, cytokines, and other molecular indicators provide valuable information about cardiovascular pathophysiology. When combined with ML algorithms including random forest, support vector machines, neural networks, and gradient boosting, these biomarkers enable predictive models with higher accuracy and sensitivity. This review explores the current progress in integrating machine learning with laboratory biomarkers for early cardiovascular disease detection. It discusses commonly used biomarkers, types of machine learning algorithms, model development strategies, clinical applications, challenges, and future prospects. The integration of ML-based predictive systems with routine laboratory testing has the potential to revolutionize personalized cardiovascular medicine and improve clinical decision-making.

Keywords: Artificial intelligence, Biomarkers, Cardiovascular diseases, Early detection, Machine learning, Predictive modeling.

Correspondence:
Mr. Prem Kumar Essgir
Department of Medical Lab Technology,
SunRise University, Alwar, Rajasthan,
INDIA.
Email: premessgir@gmail.com

Received: 02-01-2026;
Revised: 19-02-2026;
Accepted: 27-04-2026.

INTRODUCTION

Cardiovascular Diseases (CVDs) include a broad range of disorders affecting the heart and blood vessels, such as coronary artery disease, myocardial infarction, heart failure, and stroke. These conditions represent one of the most significant global health challenges and are responsible for millions of deaths annually (Taylan *et al*., 2023). The increasing prevalence of cardiovascular diseases is associated with lifestyle factors such as obesity, smoking, hypertension, and diabetes, along with genetic predisposition and aging populations (Ogunpola *et al*., 2024).

Early detection of cardiovascular diseases is essential because many individuals remain asymptomatic until the disease reaches an advanced stage. Traditional diagnostic approaches rely on clinical evaluation, imaging techniques, and laboratory tests; however, these methods often lack the ability to analyze complex relationships among multiple risk factors simultaneously. Machine learning techniques provide a powerful solution by enabling automated analysis of large datasets and identifying patterns that may not be evident through conventional statistical approaches (Elvas & Ferreira, 2025).

Machine learning algorithms can process diverse healthcare data sources, including electronic health records, laboratory biomarker profiles, imaging data, and genetic information. By analyzing these data collectively, ML models can predict disease risk, classify disease stages, and support personalized treatment strategies. Integration of machine learning with laboratory biomarkers has therefore become an emerging research area aimed at improving early detection and prevention of cardiovascular diseases (Ahmed & Mahesh, 2021).

DOI: 10.5530/ajbls.20260132

Publishing Partner : Manuscript Technomedia. [www.mstechnomedia.com]

CARDIOVASCULAR DISEASES AND THE NEED FOR EARLY DETECTION

Cardiovascular diseases develop gradually through pathological processes such as atherosclerosis, endothelial dysfunction, inflammation, and metabolic imbalance. These conditions often progress silently for many years before clinical symptoms appear. Early diagnosis and preventive interventions are therefore critical to reducing morbidity and mortality associated with cardiovascular disorders (Netala *et al*., 2025).

Traditional cardiovascular risk assessment models such as the Framingham Risk Score use demographic and clinical factors to estimate disease risk. Although these models are useful, they often fail to capture the complex interactions among genetic, metabolic, and environmental factors that influence cardiovascular health. Machine learning algorithms are capable of identifying nonlinear relationships among variables and improving prediction accuracy beyond traditional regression-based approaches (Ogunpola *et al*., 2024).

Recent research demonstrates that integrating routine laboratory biomarkers with machine learning models can significantly improve the prediction of cardiovascular events. For instance, ML-based predictive models using laboratory test results such as cholesterol levels, blood pressure measurements, and inflammatory markers have achieved high accuracy in predicting coronary artery disease (Weng *et al*., 2017; Pal *et al*., 2022). These findings suggest that combining biomarker analysis with advanced computational techniques could enable earlier detection of cardiovascular diseases and facilitate timely intervention.

Laboratory Biomarkers in Cardiovascular Disease Detection

Laboratory biomarkers play a crucial role in the diagnosis, prognosis, and monitoring of cardiovascular diseases. Biomarkers are measurable biological indicators that reflect physiological or pathological processes occurring in the body. They can be detected in blood, urine, or other biological fluids and provide valuable insights into disease mechanisms (Netala *et al*., 2025).

Cardiac Troponins

Cardiac troponins (cTnI and cTnT) are widely recognized biomarkers for myocardial injury. Elevated levels of these proteins indicate damage to cardiac muscle cells and are commonly used for diagnosing myocardial infarction. High-sensitivity troponin assays enable detection of even minor myocardial damage and can support early diagnosis of acute coronary syndrome (Park *et al*., 2017).

Natriuretic Peptides

B-type Natriuretic Peptide (BNP) and N-terminal pro-BNP (NT-proBNP) are biomarkers associated with heart failure. These peptides are released in response to ventricular stretching and pressure overload, making them useful indicators of cardiac dysfunction (Goetze *et al*., 2020).

Inflammatory Biomarkers

Inflammation plays a major role in cardiovascular disease progression. Biomarkers such as C-reactive Protein (CRP), interleukins, and tumor necrosis factor-alpha are frequently elevated in patients with cardiovascular disorders (Amezcua-Castillo *et al*., 2023).

Novel Biomarkers

Recent research has identified several emerging biomarkers that provide additional diagnostic and prognostic information. Examples include soluble ST2, galectin-3, and growth differentiation factor-15. The soluble ST2 protein is associated with cardiac remodeling and fibrosis and can predict outcomes in heart failure patients (Weinberg *et al*., 2003).

Additionally, cytokines and lipoprotein-related markers such as Preβ HDL have been investigated as potential predictors of coronary heart disease using machine learning models (Kim *et al*., 2022). These biomarkers provide valuable insights into disease mechanisms and can improve predictive modeling when combined with advanced analytical techniques.

FUNDAMENTALS OF MACHINE LEARNING IN HEALTHCARE

Machine learning is a branch of artificial intelligence that enables computers to learn patterns from data and make predictions without explicit programming (Johnson *et al*., 2018). In healthcare, ML algorithms analyze large clinical datasets to identify patterns associated with disease risk, diagnosis, and treatment outcomes (Deo, 2015).

Machine learning methods are generally classified into three main categories:

Supervised Learning

Supervised learning algorithms are trained using labeled datasets where the outcome variable is known. These algorithms learn relationships between input features and target outcomes (Alpaydin, 2020). Common supervised learning methods used in cardiovascular research include logistic regression, Support Vector Machines (SVM), random forest, and gradient boosting (Lee *et al*., 2017; Gaonkar & Davatzikos, 2013).

Unsupervised Learning

Unsupervised learning algorithms analyze unlabeled datasets to identify hidden patterns or clusters. Techniques such as clustering and principal component analysis are used to identify subgroups of patients with similar characteristics (Alloghani *et al*., 2020).

Deep Learning

Deep learning is an advanced form of machine learning that uses artificial neural networks with multiple layers to process complex datasets (Leijnen & Veen, 2020). Deep learning models are particularly useful for analyzing high-dimensional data such as genomic profiles, medical imaging, and multi-omics datasets (Ravi *et al*., 2017).

Machine learning algorithms can analyze complex relationships among multiple biomarkers and clinical variables, providing improved prediction accuracy compared to conventional statistical approaches (Ogunpola *et al*., 2024).

MACHINE LEARNING ALGORITHMS USED FOR CARDIOVASCULAR DISEASE PREDICTION

Several machine learning algorithms have been applied to predict cardiovascular diseases using laboratory biomarkers.

Logistic Regression

Logistic regression is one of the simplest machine learning techniques used for binary classification problems. It has been widely used in cardiovascular risk prediction models due to its interpretability and simplicity (Bashir *et al*., 2019).

Support Vector Machines (SVM)

Support vector machines are powerful classification algorithms that identify optimal boundaries between different classes in high-dimensional datasets (Vapnik, 1998). SVM models have demonstrated strong performance in cardiovascular disease prediction studies (Tabesh *et al*., 2010).

Random Forest

Random forest is an ensemble learning method that constructs multiple decision trees and combines their predictions to improve accuracy. Random forest models are particularly effective in handling large datasets with numerous features (Breiman, 2001)

Gradient Boosting and XGBoost

Gradient boosting algorithms iteratively improve prediction accuracy by correcting errors from previous models (Chen and Guestrin, 2016). XGBoost, a popular gradient boosting implementation, has achieved high predictive accuracy in cardiovascular disease detection studies (Ogunpola *et al*., 2024).

Artificial Neural Networks

Artificial neural networks mimic the structure of biological neurons and are capable of learning complex nonlinear relationships among variables. Neural networks have been widely used in medical diagnostics and predictive modeling (Hasan *et al*., 2020).

In a study by Fahd Saleh Alotaibi (Alotaibi, 2019). The experiment was applied using five algorithms, such as Decision Tree, Naïve Bayes, Random Forest, Logistic Regression, and Support Vector Machine. By comparing with other previous work, the accuracy of the model has improved the performance of the classifiers.

Integration of Machine Learning with Laboratory Biomarkers

The integration of machine learning with laboratory biomarkers enables the development of predictive models capable of identifying individuals at risk for cardiovascular diseases before clinical symptoms appear (Commandeur *et al*., 2019).

For example, a machine learning model incorporating plasma cytokines and Preβ HDL biomarkers demonstrated high accuracy in distinguishing patients with coronary heart disease from healthy individuals (Kim *et al*., 2022). Random forest algorithms were used to analyze the biomarker dataset and achieved excellent predictive performance.

Another study used machine learning models combined with routine laboratory tests to predict early-stage coronary artery disease. The integration of algorithms such as gradient boosting and random forest significantly improved diagnostic accuracy compared with traditional risk assessment models (Koloi *et al*., 2024).

Machine learning algorithms can also integrate multi-omics datasets, including genomic, proteomic, and metabolomic data, to identify novel biomarker signatures associated with cardiovascular diseases. These approaches enable comprehensive analysis of complex biological systems and support the development of precision medicine strategies (Deo, 2015).

CLINICAL APPLICATIONS OF ML-BASED CARDIOVASCULAR PREDICTION MODELS

Machine learning-based predictive models have several clinical applications in cardiovascular medicine.

Risk Prediction and Screening

ML models can identify individuals at high risk of cardiovascular diseases using routine laboratory data. Early identification allows clinicians to implement preventive strategies such as lifestyle modifications and pharmacological interventions.

Diagnosis

Machine learning algorithms can assist physicians in diagnosing cardiovascular diseases by analyzing laboratory biomarker profiles and clinical data.

Prognosis

ML models can predict disease progression and patient outcomes, enabling personalized treatment planning.

Personalized Medicine

Integration of ML with multi-omics data enables personalized risk prediction and targeted therapeutic strategies tailored to individual patients.

CHALLENGES IN IMPLEMENTING MACHINE LEARNING FOR CARDIOVASCULAR DIAGNOSTICS

Despite the promising potential of machine learning in cardiovascular disease detection, several challenges remain.

- **Data Quality and Availability:** High-quality datasets are required to train reliable machine learning models. Incomplete or biased data can reduce prediction accuracy.
- **Model Interpretability:** Many ML models function as "black boxes," making it difficult for clinicians to understand how predictions are generated.
- **Clinical Validation:** Predictive models must undergo extensive clinical validation before being integrated into healthcare systems.
- **Ethical and Privacy Issues:** The use of large healthcare datasets raises concerns regarding patient privacy and data security.

FUTURE PERSPECTIVES

The future of cardiovascular diagnostics will likely involve the integration of machine learning, biomarker discovery, and digital health technologies. Advances in multi-omics research and high-throughput laboratory techniques will generate large datasets that can be analyzed using advanced computational tools.

Artificial intelligence systems capable of analyzing genomic data, wearable device data, and laboratory biomarkers simultaneously may provide comprehensive insights into cardiovascular health. Such systems could enable continuous monitoring of patient health and facilitate early detection of cardiovascular diseases.

Furthermore, explainable AI techniques are being developed to improve transparency and trust in machine learning models. These methods allow clinicians to understand how specific biomarkers contribute to disease predictions, enhancing clinical acceptance of AI-based diagnostic tools.

CONCLUSION

The integration of machine learning algorithms with laboratory biomarkers represents a promising approach for the early detection and prediction of cardiovascular diseases. Machine learning techniques can analyze complex relationships among multiple clinical variables and biomarker datasets, enabling accurate risk stratification and personalized treatment strategies. Numerous studies have demonstrated that ML-based predictive models using routine laboratory biomarkers can achieve high diagnostic accuracy and improve clinical decision-making. However, challenges related to data quality, model interpretability, and clinical validation must be addressed before widespread implementation in healthcare settings. Continued research in machine learning, biomarker discovery, and digital health technologies will play a critical role in transforming cardiovascular medicine and improving patient outcomes in the future.

ABBREVIATIONS

CVDs: Cardiovascular Diseases; **ML:** Machine Learning; **AI:** Artificial Intelligence; **SVM:** Support Vector Machine; **BNP:** B-type Natriuretic Peptide; **NT-proBNP:** N-terminal pro–B-type Natriuretic Peptide; **cTnI:** Cardiac Troponin I; **cTnT:** Cardiac Troponin T; **CRP:** C-reactive Protein; **ST2:** Suppression of Tumorigenicity 2; **HDL:** High-Density Lipoprotein; **Preβ HDL:** Pre-beta High-Density Lipoprotein; **XGBoost:** Extreme Gradient Boosting; **PCA:** Principal Component Analysis; **EHRs:** Electronic Health Records; **TNF-α:** Tumor Necrosis Factor-alpha.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

REFERENCES

Taylan, O., Alkabaa, A., Alqabbaa, H., Pamukçu, E., & Leiva, V. (2023). Early Prediction in Classification of Cardiovascular Diseases with Machine Learning, Neuro-Fuzzy and Statistical Methods. Biology, 12(1), 117. https://doi.org/10.3390/biology12010117

Ogunpola, A., Saeed, F., Basurra, S., Albarrak, A. M., & Qasem, S. N. (2024). Machine Learning-Based Predictive Models for detection of cardiovascular diseases. *Diagnostics*, *14*(2), 144. https://doi.org/10.3390/diagnostics14020144

Elvas, L. B., Almeida, A., & Ferreira, J. C. (2025). The Role of AI in Cardiovascular Event Monitoring and Early Detection: Scoping Literature Review. JMIR medical informatics, 13, e64349. https://doi.org/10.2196/64349

Ahmed, M.Z., & Mahesh, C. (2021). A Comprehensive Review on Machine Learning Models for Medical Data Classification. 2021 2nd International Conference on Smart Electronics and Communication (ICOSEC), 1259-1264.

Netala, V. R., Hou, T., Wang, Y., Zhang, Z., & Teertam, S. K. (2025). Cardiovascular Biomarkers: Tools for Precision Diagnosis and Prognosis. International journal of molecular sciences, 26(7), 3218. https://doi.org/10.3390/ijms26073218

Weng, S. F., Reps, J., Kai, J., Garibaldi, J. M., & Qureshi, N. (2017). Can machine-learning improve cardiovascular risk prediction using routine clinical data?. PloS one, 12(4), e0174944. https://doi.org/10.1371/journal.pone.0174944

Pal M, Parija S, Bandyopadhyay A. (2022). Predicting cardiovascular disease using machine learning techniques. Sci Rep.12:11025.

Park, K. C., Gaze, D. C., Collinson, P. O., & Marber, M. S. (2017). Cardiac troponins: from myocardial infarction to chronic disease. Cardiovascular research, 113(14), 1708–1718. https://doi.org/10.1093/cvr/cvx183

Goetze, J. P., Bruneau, B. G., Ramos, H. R., Ogawa, T., de Bold, M. K., & de Bold, A. J. (2020). Cardiac natriuretic peptides. Nature reviews. Cardiology, 17(11), 698–717. https://doi.org/10.1038/s41569-020-0381-0

Amezcua-Castillo, E., González-Pacheco, H., Sáenz-San Martín, A., Méndez-Ocampo, P., Gutierrez-Moctezuma, I., Massó, F., Sierra-Lara, D., Springall, R., Rodríguez, E., Arias-Mendoza, A., & Amezcua-Guerra, L. M. (2023). C-Reactive Protein: The Quintessential Marker of Systemic Inflammation in Coronary Artery Disease-Advancing toward Precision Medicine. Biomedicines, 11(9), 2444. https://doi.org/10.3390/biomedicines11092444

Weinberg, E. O., Shimpo, M., Hurwitz, S., Tominaga, S., Rouleau, J. L., & Lee, R. T. (2003). Identification of serum soluble ST2 receptor as a novel heart failure biomarker. Circulation, 107(5), 721–726. https://doi.org/10.1161/01.cir.0000047274.66749.fe

Kim J, Lee SY, Cha BH, Lee W, Ryu J, Chung YH, Kim D, Lim S-H, Kang TS, Park B-E, Lee M-Y and Cho S (2022) Machine learning models of clinically relevant biomarkers for the prediction of stable obstructive coronary artery disease. Front. Cardiovasc. Med. 9:933803. doi: 10.3389/fcvm.2022.933803

Johnson, K. W., Torres Soto, J., Glicksberg, B. S., Shameer, K., Miotto, R., Ali, M., Ashley, E., & Dudley, J. T. (2018). Artificial Intelligence in Cardiology. Journal of the American College of Cardiology, 71(23), 2668–2679. https://doi.org/10.1016/j.jacc.2018.03.521

Deo R. C. (2015). Machine Learning in Medicine. Circulation, 132(20), 1920–1930. https://doi.org/10.1161/CIRCULATIONAHA.115.001593

Alpaydin, E. (2020) Introduction to Machine Learning. MIT Press, Cambridge. https://doi.org/10.7551/mitpress/13811.001.0001

Lee, E. J., Kim, Y. H., Kim, N., & Kang, D. W. (2017). Deep into the Brain: Artificial Intelligence in Stroke Imaging. Journal of stroke, 19(3), 277–285. https://doi.org/10.5853/jos.2017.02054

Gaonkar, B., & Davatzikos, C. (2013). Analytic estimation of statistical significance maps for support vector machine based multi-variate image analysis and classification. NeuroImage, 78, 270–283. https://doi.org/10.1016/j.neuroimage.2013.03.066

Alloghani, Mohamed & Al-Jumeily Obe, Dhiya & Mustafina, Jamila & Hussain, Abir & Aljaaf, Ahmed. (2020). A Systematic Review on Supervised and Unsupervised Machine Learning Algorithms for Data Science. 10.1007/978-3-030-22475-2_1.

Leijnen, S., & Veen, F.V. (2020). The Neural Network Zoo. Proceedings.

Ravi, D., Wong, C., Deligianni, F., Berthelot, M., Andreu-Perez, J., Lo, B., & Yang, G. Z. (2017). Deep Learning for Health Informatics. IEEE journal of biomedical and health informatics, 21(1), 4–21. https://doi.org/10.1109/JBHI.2016.2636665

Bashir, S., Khan, Z., Hassan Khan, F., Anjum, A., & Bashir, K. (2019). Improving Heart Disease Prediction Using Feature Selection Approaches. 2019 16th International Bhurban Conference on Applied Sciences and Technology (IBCAST), 619-623.

Vapnik, V. (1998) Statistical Learning Theory. John Wiley & Sons, Chichester.

Tabesh, Pooya & Lim, Gino & Khator, Suresh & Dacso, Clifford. (2010). A support vector machine approach for predicting heart conditions.

Breiman, L. (2001) Random Forests. Machine Learning, 45, 5-32. http://dx.doi.org/10.1023/A:1010933404324

Chen, T. and Guestrin, C. (2016) XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Association for Computing Machinery, San Francisco, 13-17 August 2016, 785-794.https://doi.org/10.1145/2939672.2939785

Hasan, Najmul & Bao, Yukun. (2020). Comparing different feature selection algorithms for cardiovascular disease prediction. Health and Technology. 11. 10.1007/s12553-020-00499-2.

Alotaibi, F. S. (2019). Implementation of machine learning model to predict heart failure disease. International Journal of Advanced Computer Science and Applications, 10(6). https://doi.org/10.14569/ijacsa.2019.0100637

Commandeur, F., Slomka, P. J., Goeller, M., Chen, X., Cadet, S., Razipour, A., McElhinney, P., Gransar, H., Cantu, S., Miller, R. J. H., Rozanski, A., Achenbach, S., Tamarappoo, B. K., Berman, D. S., & Dey, D. (2019). Machine learning to predict the long-term risk of myocardial infarction and cardiac death based on clinical risk, coronary calcium, and epicardial adipose tissue: a prospective study. Cardiovascular Research, 116(14), 2216–2225. https://doi.org/10.1093/cvr/cvz321

Koloi, A., Loukas, V. S., Hourican, C., Sakellarios, A. I., Quax, R., Mishra, P. P., Lehtimäki, T., Raitakari, O. T., Papaloukas, C., Bosch, J. A., März, W., & Fotiadis, D. I. (2024). Predicting early-stage coronary artery disease using machine learning and routine clinical biomarkers improved by augmented virtual data. European heart journal. Digital health, 5(5), 542–550. https://doi.org/10.1093/ehjdh/ztae049

Cite this article: Essgir PK, Simha MV. Application of Machine Learning Algorithms for Early Prediction of Cardiovascular Diseases through Laboratory Biomarkers. Asian J Biol Life Sci. 2026;15(1):32-6.