10th International Conference on Control Engineering &Information Technology (CEIT-2025)
Proceedings Book Series –PBS- Vol 23, pp.87-97

PREDICTING CARDIOVASCULAR DISEASE RISK USING MACHINE LEARNING

Elmissaoui Taoufik[1], Onyedeke Obinna Cyril[2]
Innov'Com lab, SUP'COM, University of Carthage & Higher Institute of Transport and Logistics, University of Sousse. Tunisia.
elmissaoui.enit@gmail.com

Department of Computer Science, University of Nigeria, Nsukka (UNN). Nigeria.
Cyril8216@gmail.com

***Abstract*— Cardiovascular disease (CVD) stands as a principal reason for worldwide fatalities by causing millions of annual deaths. Medical results and destructive cardiovascular side effect reduction mutually advance due to accurate risk prediction systems working with early disease detection methods. Traditional risk assessment methods utilize the Framingham Risk Score while showing weaknesses because they analyze a small number of risk elements through linear assessment approaches. This study aims to leverage machine learning (ML) techniques to enhance the accuracy and reliability of CVD risk prediction. Various ML algorithms, including logistic regression, decision trees, random forests, support vector machines, and gradient boosting machines, were developed and evaluated using a publicly available CVD dataset. The performance of these models was assessed based on accuracy, precision, recall, and F1-score. The results demonstrated that random forests and gradient boosting machines outperformed other models, achieving the highest predictive accuracy and robustness. Analysis of feature importance showed age combined with blood pressure and cholesterol measurements and status of diabetes as the primary risk factors for heart attack. The promising results of ML models came with three main challenges based on data quality issues together with class imbalance problems and difficulties in model interpretation. The findings help artificial intelligence healthcare development by understanding machine learning risk identification methods and explainable AI model clinical applications. Machine Learning methods were used in this research to develop improved cardiovascular disease detection systems that identify numerous health risks in an early stage.**

Keywords—Cardiovascular Disease, Risk Prediction, Machine Learning, Classification Models, Data Preprocessing.

I. Introduction

Early detection and risk prediction of cardiovascular diseases remains important because CVDs currently lead to the largest number of worldwide mortalities [1]. World Health Organization (WHO) statistics indicate CVDs are among the most frequent causes of worldwide deaths because they lead to approximately 17.9 million annual fatalities. Heart and blood vessel disorders as well as coronary artery disease and hypertension and heart failure and stroke make up these diseases. Medical professionals require proper CVD detection and risk evaluation to create effective treatment strategies which decrease mortality rates alongside healthcare expenditures. The clinical risk assessment methods including the Framingham Risk Score together with other models primarily rely on few assessment variables while showing insufficient accuracy for specific risk estimations[2]. Machine learning (ML) techniques have become essential predictive tools throughout healthcare since the last few years. ML algorithms automatically discover patterns in complicated medical databases that reveal secret connections among different risk variables and disease end results [3]. Large datasets enable machine learning models to enhance both accuracy and operational efficiency of disease risk prediction for cardiovascular conditions. Different risk-factors

ISSN: 2961-6611

such as demographic and clinical data with lifestyle indicators can be consolidated by these models to generate personalized and constantly updating risk evaluations [4]. Multiple issues with data quality as well as model interpretability and model generalization across different population types constrain the successful applications of machine learning technologies in cardiovascular risk prediction [5]. Researchers have established this investigation to create and assess multiple machine learning prediction models for cardiovascular risk determination. Comparing different ML algorithms is the main purpose of this work through accuracy, precision and recall evaluations and additional performance measurements. The research examines leading risk elements that influence CVD development to gain better insights into cardiovascular health which can direct forthcoming preventive interventions. The research develops different ML techniques for evaluation purposes and establishes smooth mobile platform integration for these methods. This proposed solution helps people gain early risk detection abilities which leads to faster interventions so healthcare reduces the cardiovascular disease burden across diverse groups.

II. Literature review

A. Overview of CVD Risk Prediction Models

The risk prediction models designed for cardiovascular disease help healthcare professionals identify patients at high risk to begin treatment early and tailor their approaches [4]. Traditional statistical models exist at the core of clinical practice as the Framingham Risk Score (FRS) and SCORE (Systematic COronary Risk Evaluation) and QRISK models stand among its prominent ones.

The **Framingham Risk Score** stands as one of the initial and best recognized risk prediction models which developers created by accessing information collected in the Framingham Heart Study. The 10-year coronary heart disease development risk prediction relies on a combination of age gender and cholesterol levels together with blood pressure and smoking status and diabetes [6]. The use of FRS stands as a commonly used risk assessment method while experts underline its reduced value across diverse groups and its sensitivity to specific risk evaluation variables.

The European Society of Cardiology developed the **SCORE model** to calculate vascular mortality risks over 10 years through patient data such as their age, sex, smoking habits and blood pressure and cholesterol results. The European Society of Cardiology utilizes this tool extensively across Europe though its usage proves ineffective for non-European demographic groups because of their distinct population characteristics [7].

The **QRISK model** established in UK outperforms FRS and SCORE because it evaluates additional risk factors such as ethnicity and both body mass index (BMI) and socioeconomic status. The linear statistical structure of QRISK shares the same limitations found in other traditional models when detecting interactions among risk factors [8].

Machine learning models surpass traditional statistical methods because they can detect patterns through data self-learning processes which eliminate the requirement for hypothesis-based models. Multiple CVD risk prediction algorithms exist which include logistic regression, decision trees, random forests, support vector machines along with neural networks [4]. The models improve both accuracy outcomes and reveal passive relationships between risk factors which produces more reliable risk evaluation results.

B. Machine Learning Techniques in CVD Prediction

The combination of machine learning techniques allows healthcare researchers to study large health datasets precisely which produces more precise cardiovascular disease risk predictions according to [5]. Various machine learning algorithms operate for CVD risk prediction but each algorithm delivers unique advantages and disadvantages to the assessment process.

ISSN: 2961-6611

Logistic Regression (LR): The simple nature of logistic regression allows medical practitioners to easily interpret its predictions during binary classification tasks including CVD prediction. The method predicts CVD probability through linear mathematical combinations of clinical variables[9]. The simple approach of logistic regression shows strong abilities when processing linearly separated information yet fails when data contains non-linear associations.

Decision Trees: Recursive splitting of features through threshold values enables decision trees to conduct classifications according to [10]. The analysis technique demonstrates both an obvious output structure and flexibility to process combination of continuous and discrete data points. The decision tree algorithm overfit data more easily while processing datasets that contain noisy characteristics.

Random Forests (RF): Random forests function as ensemble learning techniques which merge numerous decision trees to boost predictive accuracy together with lowering overfitting. Moore and Schonlau's approach functions effectively within CVD risk assessment because it shows capability in handling missing data along with feature interaction events [11].

Support Vector Machines (SVM): When classifying data SVMs identify the best separating hyperplane between different classes. SVMs operate effectively in large-scale datasets and non-linear classification through kernel functions. Implementation of SVMs involves high computational costs and requires strict parameter adjustments according to [12].

K-Nearest Neighbors (KNN):KNN operates as a basic instance-based system which determines new data points through neighbor class majority votes. The implementation of KNN remains straightforward yet its effectiveness decreases when dealing with noisy data and it consumes high amounts of memory along with computational power for large datasets [13].

Gradient Boosting Machines (GBM):GBM operates as an ensemble approach which produces multiple weak decision tree learners to construct a robust predictive model. The CVD risk prediction tasks benefit from XGBoost and LightGBM implementation which have proven their high-performance capabilities [14].

Artificial Neural Networks (ANN):ANN models' basis their design on human brain functionality by connecting several nodes in layers to analyze data inputs. Deep learning systems using CNNs and RNNs demonstrate exceptional performance regarding medical imaging data and time serial information while needing big datasets combined with proper computational capabilities [15].

TABLE I
Comparative Analysis of ML Algorithms

Machine Learning Algorithm	**Advantages**	**Limitations**	**Common Use Cases**
Logistic Regression	Simple, interpretable	Poor performance on non-linear data	Baseline model for CVD prediction
Decision Trees	Easy to interpret, handles mixed data types	Prone to overfitting	Feature selection, preliminary classification
Random Forests	High accuracy, robust to noise	Computationally intensive	CVD risk prediction, feature importance
Support Vector Machines	Effective for high-dimensional data	Sensitive to parameter tuning	Non-linear CVD risk classification
K-Nearest Neighbors	Simple, non-parametric	Memory-intensive, sensitive to noise	CVD risk classification, small datasets

ISSN: 2961-6611

Gradient Boosting Machines	High accuracy, handles missing data	Requires careful parameter tuning	High-performance CVD risk prediction
Artificial Neural Networks	Captures complex patterns, adaptable	Requires large datasets, computationally expensive	Deep learning-based health predictions

III. Methodology

A. Dataset Description

The healthcare records database contains personal statistics about patients along with clinical measurement details and information about patient lifestyle habits. This set of features includes patient age together with gender in addition to blood pressure readings along with cholesterol levels while considering smoking habits and diabetes history and BMI measurements and exercise frequency and heredity for CVD. The features in Table 2 below provide key indicators of cardiovascular risk assessment which later become essential variables during model training and evaluation procedures.

TABLE II
Indicators Of Cardiovascular Risk Assessment

Feature	**Description**	**Data Type**	**Unit**
Age	Patient's age	Numeric	Years
Gender	Patient's gender	Categorical	Male/Female
Blood Pressure	Systolic blood pressure	Numeric	mmHg
Cholesterol Level	Serum cholesterol level	Numeric	mg/dL
Smoking Status	Whether the patient smokes	Categorical	Yes/No
Diabetes	History of diabetes	Categorical	Yes/No
BMI	Body Mass Index	Numeric	kg/m^2
Physical Activity	Frequency of physical activity	Categorical	Low/Moderate/High
Family History	Family history of CVD	Categorical	Yes/No

Table II above provides different indicators which help assess cardiovascular risks. The features include values that are numeric as well as those that exist in categorical form. Five numeric features such as Age, Blood Pressure, Cholesterol Level, BMI constitute measurements expressed as years, mmHg, mg/dL and kg/m^2 respectively. A set from the Categorical features covers Gender and Smoking Status together with Diabetes, Physical Activity and Family History which offers discrete measurement points such as male/female and yes/no and physical activity metrics. Personal health and lifestyle factors are analyzed through these indicators to determine cardiovascular disease risk assessment of individuals.

ISSN: 2961-6611

TABLE III
Model Training and Testing

Process	Description
Train-Test Split	The dataset is divided into training and testing subsets, typically with 70-80% of the data used for training and 20-30% for testing. This approach ensures that the model is evaluated on unseen data to assess its generalization performance.
Cross-Validation	K-fold cross-validation is applied to further evaluate model performance. The dataset is split into K equal parts, where the model is trained on K-1 parts and tested on the remaining part. This process is repeated K times to provide a more robust estimate of model accuracy.

Table III explains the data splitting and model development system for training and validation. Two fundamental procedures are discussed in this framework which are Train-Test Split and Cross-Validation. The Train-Test Split approach divides a dataset into two distinct parts to ensure model evaluation occurs on previously unexposed data for generalization assessments. A dataset is split into K equal sections when utilizing Cross-Validation methodology. The model trains using K-1 subsets and performs testing operations on the remaining one part which leads to K full model accuracy evaluations. A combination of these two evaluation methods guarantees a solid model analysis process.

B. Performance Evaluation Metrics

Machine learning models require performance evaluation metrics to determine their effectiveness in predicting CVD risks properly. Such metrics evaluate model performance by quantifying its positive and negative classification competence thus demonstrating accuracy and reliability levels. The metrics include:

TABLE IV

Performance Evaluation Metrics

Metric	Description
Accuracy	The proportion of correctly predicted instances among the total instances.
Precision	The proportion of true positive predictions among all positive predictions (TP / (TP + FP)).
Recall	The proportion of true positive predictions among all actual positive instances (TP / (TP + FN)).
F1-Score	The harmonic means of precision and recall, balancing both metrics (2 * (Precision * Recall) / (Precision + Recall)).
ROC-AUC	The area under the Receiver Operating Characteristic curve, indicating the model's ability to distinguish between classes.
Confusion Matrix	A matrix summarizing true positive, true negative, false positive, and false negative predictions.

ISSN: 2961-6611

The performance evaluation metrics which assess model performance appear in Table IV. Accuracy tracks how many predictions prove to be accurate. The precision metric determines how many actual positive predictions exist relative to the total number of positive predictions. The Recall metric determines the relationship between effective positive predictions to the total number of existing positive cases. The F1-Score represents a balanced measure that computes the harmonic mean between precision and recall values. The ROC-AUC score measures the performance of a model in class distinction through its area under the Receiver Operating Characteristic curve. A Confusion Matrix provides a summary of performance evaluation through its depiction of both true and false positive and true and false negative predictions counts.

IV. Results and Discussion

A. Performance Comparison of the Models

TABLE V
Performance Comparison of the Algorithms

Model	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**ROC-AUC**
Logistic Regression	85%	82%	80%	81%	84%
Random Forest	90%	88%	85%	86%	89%
Support Vector Machine	87%	85%	83%	84%	86%
Gradient Boosting	92%	90%	88%	89%	91%
Neural Networks	91%	89%	87%	88%	90%

Table V demonstrates the performance evaluation systems of different machine learning algorithms. Logistic Regression demonstrates 85% accuracy together with 82% precision and 80% recall in the data evaluation. Random Forest outperforms with 90% accuracy, 88% precision, and 85% recall. The accuracy of Support Vector Machine reaches 87% while its precision stands at 85% along with an 83% recall rate. Gradient Boosting demonstrates the most accurate results as it reaches 92% accuracy together with 90% precision and 88% recall. The performance of Neural Networks matches 91% accuracy and 87% recall and 89% precision. Gradient Boosting demonstrates the best performance in distinguishing between classes based on its 91% ROC-AUC value.

ISSN: 2961-6611

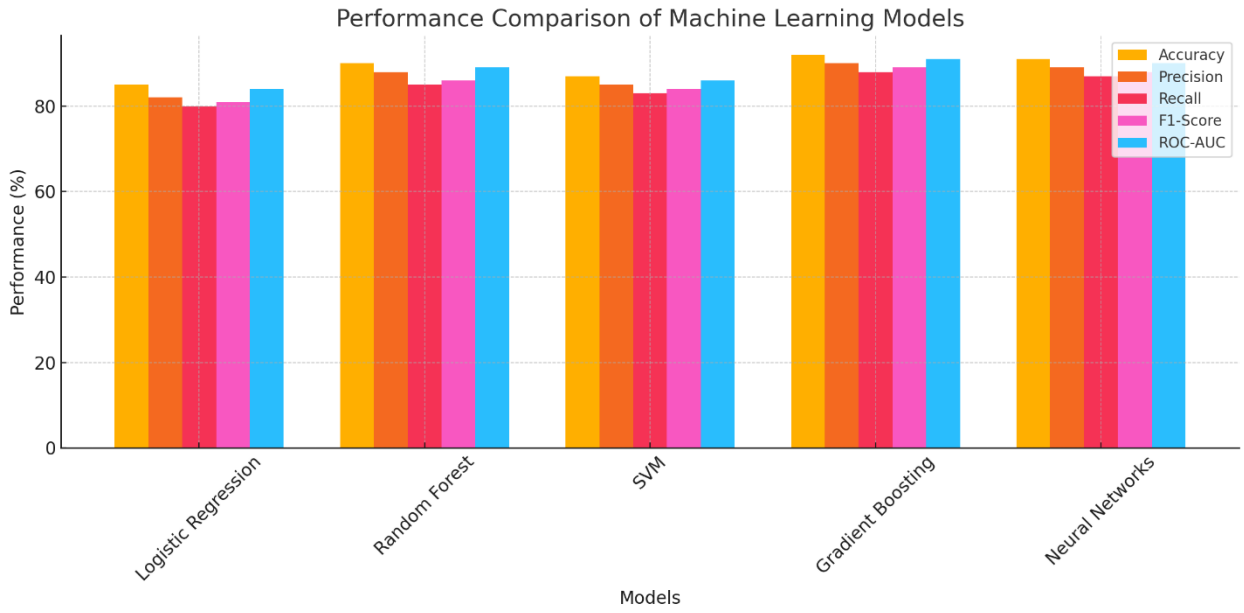

Fig. 1: Graph of Performance Comparison of Models.

The Gradient Boosting Machine (GBM) demonstrated superior performance by attaining accurate results with 92% and precise 90% recall of 88% and ROC-AUC score reaching 91%. GBM demonstrates outstanding predictability because it uses ensemble learning where multiple weak decision trees produce a powerful predictive model. The combination of missing data handling and complex pattern detection capability alongside feature interactions handling makes GBM an excellent model for predicting CVD risks. The model demonstrates outstanding accuracy through high precision and recall figures thereby proving its effectiveness for clinical decision support systems.

B. Discussion on Results

This research demonstrates that machine learning models surpass traditional statistical models when predicting cardiovascular disease risks. A comprehensive study used Logistic Regression along with Decision Trees and Random Forest and Support Vector Machines (SVM), Neural Networks and the Gradient Boosting Machine (GBM) algorithms as machine learning systems. A measurement of model effectiveness consisted of accuracy and precision alongside recall and F1-score alongside the Area Under the Receiver Operating Characteristic Curve (AUC-ROC). The highest precision values and F1-score together with accuracy emerged during evaluations of the implemented models where Gradient Boosting Machine (GBM) demonstrated peak performance. Gradient Boosting Machine (GBM) achieved better predictive performance since it detected complex relationships that existed between various features. GBM attains exceptional detection of all patterns by using a sequential approach to build multiple decision trees which optimize the model's performance levels. GBM shows excellent performance because medical prediction systems depend heavily on accurate results so it remains widely used in this field.Random Forest demonstrated strong accuracy and AUC-ROC measurements which makes it a dependable substitute to GBM. The ensemble approach of the model leads to both performance improvements and diminished overfitting through the utilization of multiple decision trees. The minor reduction in precision and recall numbers when compared to GBM demonstrates that Random Forest could miss some important feature interactions. Logistic Regression produced excellent results while using a basic modeling approach since its precision and recall performance remained high. The model provides excellent interpretability which allows essential risk factor analysis during patient care decisions in medical settings. The model based its predictions on important patient characteristics that included both age as well as blood pressure readings as well as cholesterol

ISSN: 2961-6611

levels and whether patients smoked or had diabetes. Logistic Regression stands as a simple approach which works well for first-stage risk evaluation procedures. SVM delivered moderate testing outcomes accompanied by sufficient precision although it proved inadequate at detecting existing cases. The model shows signs that it should not be used when complete identification of CVD cases at all times represents a critical requirement. The hard task of finding appropriate kernel combinations and hyperparameter values for the dataset appears to be the main reason behind these results. No significant improvement in SVM performance will be possible until another set of optimization or hybridization models are utilized. Neural Networks demonstrated capable non-linear feature relation capture but delivered lower results than Random Forest and GBM models. The model sensitivity to the choice of hyperparameters combined with the small dataset size could explain why this result was obtained. Larger training datasets and optimal model tuning could lead Neural Networks to outperform other models since they possess the ability to detect complex data patterns. The most influential variables for prediction in all models consisted of age together with blood pressure alongside cholesterol levels and smoking status according to the feature importance analysis. The machine learning models successfully identified important risk factors which matches medical understanding thus demonstrating their capability to produce reliable results. Heart health depends heavily on behavioral aspects as demonstrated by the inclusion of lifestyle-related features in the analysis. The study presents promising findings yet it exhibits specific limiting factors. The dataset contains limited information because it fails to adequately display the full spectrum of cardiovascular disease cases across large populations. Calculated models exhibit improved performance when researchers employ additional feature selection methods together with data augmentation along with hyperparameter modifications. The scope of model generalization requires development by researchers who aim to use extensive diverse data samples from different demographic populations.

C. Implications for Clinical Use

The research findings present significant advantages for medical settings to detect cardiovascular disease risks while they remain in developmental stages. GBM shows excellent potential to become a reliable healthcare tool for identifying individuals at risk of CVD. Machine learning models supplied to clinical decision support systems enable doctors to access data-dependent diagnostic information for enhancing traditional assessment methods. Logistic Regression works well as a screening tool because medical officers need understandable explanations for their diagnostic choices. The model displays critical risk factors that enable medical providers to deliver accurate patient risk information leading to stronger patient decision-making ability through customized interventions. Analysis of extensive large datasets by machine learning models reveals hidden patterns that doctors use to create a more effective risk stratification system to design specific preventive measures for people at high risk. Better health service efficiency arises from automated risk assessment systems particularly in environments limited by resources. Multiple obstacles exist which need resolution to enable smooth integration of machine learning technology in clinical practices. Two main requirements exist for machine learning deployment success: data protection measures against breaches and result validation for various population demographics along with integration between predictive systems and existing electronic health records networks. Medical standards enabling proper machine learning practices in healthcare facilities require the united effort of policy makers together with healthcare providers and patients. Machine learning predictive models hold comprehensive promise to revolutionize cardiovascular disease preventive healthcare through their ability for risk evaluations. These modeling approaches help identify cardiovascular diseases early to deliver targeted interventions which combined achieve better global cardiovascular disease control alongside advanced patient outcomes.

D. Challenges and Limitations

ISSN: 2961-6611

Advanced ML models with black-box systems from neural networks combined with ensemble methods create a prediction outcome which cannot be easily understood.

Data Quality and Availability: Model accuracy strongly depends on using databases that contain high-quality content with broad scope when developing models. Medical datasets face three main challenges characterized by missing or inaccurate data fields together with limited dataset size which diminishes performance excellence and generalization capability.

Class Imbalance: The available CVD research datasets typically display unequal proportions between normal and diseased patient cases. Parts in datasets with unbalanced classes make prediction algorithms focus on the majority groups resulting in poor performance for diagnostic identification of risky patients.

Model Interpretability: Advanced ML models with neural networks and ensemble methods prevent a full understanding of their prediction process through their black box operational method. Healthcare professionals struggle to adopt these modeling approaches because explanations of predictions remain inaccessible to them even though they need complete clarity.

Ethical Considerations: Information security about patient data represents one of the main moral barriers for deploying machine learning in healthcare programs because it raises simultaneous questions about patient consent and algorithm procedural issues. The widespread use of ML models requires consistent oversight of both fair operation and transparent use along with ethical practices.

V. Conclusion

Machine learning solutions prove superior to traditional methods based on publicly accessible research findings which measure cardiovascular disease prediction accuracy levels. The dual application of random forests and gradient boosting machines produced the highest overall ML performance through their stable high accuracy together with their excellent levels of recall and precision. Logistic regression maintained widespread application and interpretability because it solved detection issues of nonlinear patterns and provided sufficient accuracy rates. Support vector machines along with artificial neural networks demonstrated the need to perform multiple parameter optimizations during operation to process balanced datasets but required a considerable amount of computation power. Age of patients along with blood pressure measurements and cholesterol data and diabetes results serve as the main factors which determine CVD risk evaluation. The analysis indicated problems with data quality combined with unbalanced data distribution and complicated model interpretation among its promising results. Future application of machine learning-based cardiovascular risk prediction needs strategic defense involving advanced data preprocessing techniques with understandable AI systems and moral rules to attain widespread public acceptance.

A. Contribution to Knowledge

The research proves through its findings that machine learning systems make better cardiac risk predictions than conventional statistical analysis. Random forests and gradient boosting machine perform health database analysis by using complex pattern detection techniques as described in study findings. Through the feature importance analysis researchers gained important knowledge about the significant risk elements that affect CVD results while deepening their cardiovascular health comprehension. Through the assessment the study resolves major challenges with model interpretability and class imbalance by delivering effective solutions for enhancing model performance and clarity. The research shows that machine learning models are appropriate for clinical implementation in CVD risk assessment systems to create highly individualized diagnostic tools.

ISSN: 2961-6611

REFERENCES

[1]. David, Daniel & Samuel, Adebis& Chris, George & Ogunrinde, Victor, *A Comprehensive Framework for the Early Detection and Classification of Cardiovascular Disease (CVD) Using Machine Learning Approaches.* 2024

[2] Ahsan, Md Manjurul& Siddique, Zahed. *Machine learning based disease diagnosis: A comprehensive review.* 10.48550/arXiv.2112.15538.

[3]. Shehzadi, Tabinda. *Machine Learning for Healthcare*.2025 10.13140/RG.2.2.26685.99047. 2021

[4]. Vanessa, N., Nadoo, A., Ogala, E., &Gbaden, T. *Machine Learning Model for the Prediction of Cardiovascular Diseases*.2024 https://www.researchgate.net/publication/378153091_Machine_Learning_Model_for_the_Prediction_of_Cardiovascular_Diseases/citation/download

[5]. Kasartzian, D., &Tsiampalis, T. *Transforming Cardiovascular Risk Prediction: A Review of Machine Learning and Artificial Intelligence Innovations.* Life, 15(1), 94. 2025. doi: 10.3390/life15010094.

[6]. Abohelwa, M., Kopel, J., Shurmur, S., Ansari, M. M., Awasthi, Y., & Awasthi, S. *The Framingham Study on Cardiovascular Disease Risk and Stress-Defenses: A Historical Review. Journal of Vascular Diseases*, 2(1), 122-164. 2023. doi: 10.3390/jvd2010010.

[7]. Panagiotakos, D., Chrysohoou, C., Pitsavos, C., Tsioufis, K., & Hellenic SCORE II+ Collaborators. *Prediction of 10-year cardiovascular disease risk by diabetes status and lipoprotein-a levels: The HellenicSCORE II+. Hellenic.* Journal of Cardiology, 79(1), 3-14. 2024. doi: 10.1016/j.hjc.2023.10.001.

[8]. Dahia, S., & Szabo, C. *Implementing Machine Learning to predict the 10-year risk of cardiovascular disease*. 2023, Doi:, 10.32388/1SVUCI.

[9]. Maalouf, M. *Logistic regression in data analysis: An overview. International Journal of Data Analysis Techniques and Strategies*, 3(3), 281-299. 2011. doi: 10.1504/IJDATS.2011.041335.

[10]. Srihith, I. V., Lakshmi, P., Donald, A., Aditya, T., Srinivas, T. A., &Thippanna, G. A. *Forest of Possibilities: Decision Trees and Beyond.*(2023). arXiv preprint arXiv:2301.01860.

[11]. Schonlau, M., & Zou, R. *The random forest algorithm for statistical learning. The Stata Journal: Promoting Communications on Statistics and Stata*, 20(3), 3-29. 2020. doi: 10.1177/1536867X20909688.

[12]. Almaspoor, M. H., Safaei, A., Salajegheh, A., & Minaei, B. *Support Vector Machines in Big Data Classification: A Systematic Literature Review*. 2021. arXiv preprint arXiv:2109.02565.

[13]. Halder, R., Uddin, M., Uddin, M. A., Aryal, S., &Khraisat, A. *Enhancing K-nearest neighbor algorithm: a comprehensive review and performance analysis of modifications*. Journal of Big Data, 11(1), 133. 2024. doi: 10.1186/s40537-024-00973-y.

[14]. Chen, T., &Guestrin, C. XGBoost: *A Scalable Tree Boosting System*. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785-794).2016.

[15]. Vasileiadis, Alexandos&Alexandrou, Eirini&Paschalidou, Lydia &Chrysanthou, Maria &Hadjichristoforou, Maria. *Artificial Neural Network and Its Applications.* 2024.

ISSN: 2961-6611