

Review

A Powerful Paradigm for Cardiovascular Risk Stratification Using Multiclass, Multi-Label, and Ensemble-Based Machine Learning Paradigms: A Narrative Review

Jasjit S. Suri ^{1,*}, Mrinalini Bhagawati ², Sudip Paul ², Athanasios D. Protogerou ³, Petros P. Sfikakis ⁴, George D. Kitas ⁵, Narendra N. Khanna ⁶, Zoltan Ruzsa ⁷, Aditya M. Sharma ⁸, Sanjay Saxena ⁹, Gavino Faa ¹⁰, John R. Laird ¹¹, Amer M. Johri ¹², Manudeep K. Kalra ¹³, Kosmas I. Paraskevas ¹⁴ and Luca Saba ¹⁵

- ¹ Stroke Diagnostic and Monitoring Division, AtheroPoint™, Roseville, CA 95661, USA
 - ² Department of Biomedical Engineering, North-Eastern Hill University, Shillong 793022, India; bhagawatimrinalini07@gmail.com (M.B.); sudip.paul.bhu@gmail.com (S.P.)
 - ³ Research Unit Clinic, Laboratory of Pathophysiology, Department of Cardiovascular Prevention, National and Kapodistrian University of Athens, 11527 Athens, Greece; aprotog@med.uoa.gr
 - ⁴ Rheumatology Unit, National Kapodistrian University of Athens, 11527 Athens, Greece; psfikakis@med.uoa.gr
 - ⁵ Arthritis Research UK Centre for Epidemiology, Manchester University, Manchester 46962, UK; george.kitas@gmail.com
 - ⁶ Department of Cardiology, Indraprastha APOLLO Hospitals, New Delhi 110020, India; drnnkhanna@gmail.com
 - ⁷ Department of Internal Medicines, Invasive Cardiology Division, University of Szeged, 6720 Szeged, Hungary; zruza25@gmail.com
 - ⁸ Division of Cardiovascular Medicine, University of Virginia, Charlottesville, VA 22903, USA; as8ah@hscmail.mcc.virginia.edu
 - ⁹ Department of CSE, International Institute of Information Technology, Bhubaneswar 751003, India; sanjay@iiit-bh.ac.in
 - ¹⁰ Department of Pathology, A.O.U., di Cagliari-Polo di Monserrato s.s., 09045 Cagliari, Italy; gavinofaa@gmail.com
 - ¹¹ Cardiology Department, St. Helena Hospital, St. Helena, CA 94574, USA; lairdjr@ah.org
 - ¹² Department of Medicine, Division of Cardiology, Queen's University, Kingston, ON K7L 3N6, Canada; amerjohri@gmail.com
 - ¹³ Department of Radiology, Massachusetts General Hospital, Boston, MA 02114, USA; mkalra@mgh.harvard.edu
 - ¹⁴ Department of Vascular Surgery, Central Clinic of Athens, N. Iraklio, 14122 Athens, Greece; paraskevask@hotmail.com
 - ¹⁵ Department of Radiology, A.O.U., di Cagliari-Polo di Monserrato s.s., 09045 Cagliari, Italy; lucasabamd@gmail.com
- * Correspondence: jasjit.suri@atheropoint.com; Tel.: +1-(916)-749-5628



Citation: Suri, J.S.; Bhagawati, M.; Paul, S.; Protogerou, A.D.; Sfikakis, P.P.; Kitas, G.D.; Khanna, N.N.; Ruzsa, Z.; Sharma, A.M.; Saxena, S.; et al. A Powerful Paradigm for Cardiovascular Risk Stratification Using Multiclass, Multi-Label, and Ensemble-Based Machine Learning Paradigms: A Narrative Review. *Diagnostics* **2022**, *12*, 722. <https://doi.org/10.3390/diagnostics12030722>

Academic Editor: Dechang Chen

Received: 21 February 2022

Accepted: 13 March 2022

Published: 16 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Background and Motivation: Cardiovascular disease (CVD) causes the highest mortality globally. With escalating healthcare costs, early non-invasive CVD risk assessment is vital. Conventional methods have shown poor performance compared to more recent and fast-evolving Artificial Intelligence (AI) methods. The proposed study reviews the three most recent paradigms for CVD risk assessment, namely multiclass, multi-label, and ensemble-based methods in (i) office-based and (ii) stress-test laboratories. **Methods:** A total of 265 CVD-based studies were selected using the preferred reporting items for systematic reviews and meta-analyses (PRISMA) model. Due to its popularity and recent development, the study analyzed the above three paradigms using machine learning (ML) frameworks. We review comprehensively these three methods using attributes, such as architecture, applications, pro-and-cons, scientific validation, clinical evaluation, and AI risk-of-bias (RoB) in the CVD framework. These ML techniques were then extended under mobile and cloud-based infrastructure. **Findings:** Most popular biomarkers used were office-based, laboratory-based, image-based phenotypes, and medication usage. Surrogate carotid scanning for coronary artery risk prediction had shown promising results. Ground truth (GT) selection for AI-based training along with scientific and clinical validation is very important for CVD stratification to avoid RoB. It was observed that the most popular classification paradigm is multiclass followed by the ensemble, and

multi-label. The use of deep learning techniques in CVD risk stratification is in a very early stage of development. Mobile and cloud-based AI technologies are more likely to be the future. **Conclusions:** AI-based methods for CVD risk assessment are most promising and successful. Choice of GT is most vital in AI-based models to prevent the RoB. The amalgamation of image-based strategies with conventional risk factors provides the highest stability when using the three CVD paradigms in non-cloud and cloud-based frameworks.

Keywords: CVD; multiclass; multi-label; ensemble; cloud; COVID; bias; gold standard

1. Introduction

Cardiovascular disease (CVD) results in 18 million deaths worldwide [1]. In 2020, the financial burden due to CVD was \$237 billion USD [2]. With COVID-19 still not subsided, rising inflation costs, loss of families due to migration, depression on the rise, and comorbidities increasing, the risk of CVD risk is likely to go up. The main cause of CVD is atherosclerotic deposition in the heart's coronary arteries [3]. Due to different types of comorbidities such as diabetes [4], chronic kidney disease (CKD) [5,6], rheumatoid arthritis [7,8], hypertension [9], high lipids [10], and brain diseases [11–13], the risk of CVD is increasing, putting patients at a higher risk of heart disease and stroke. It is estimated that by 2030, the financial burden due to CVD will reach about \$3T USD [2]. Therefore, the need for an early CVD risk detection system will alleviate the mortality and morbidity rates.

CVD risk assessment can take two forms, namely (a) in the doctor's office or pathology laboratory or both, (b) in the stress test centers or signal processing clinics [14–16]. The calculators used in the office-based scenario are conventional CVD calculators that use laboratory-based biomarkers (LBBM) and office-based biomarkers (OBBM) [17], while the CVD risk assessment in stress test centers uses electrocardiograms (ECG) [18–20]. There are multiple conventional tools for assessment of risk due to CVD, namely (i) QRISK3 [21], (ii) Framingham risk score (FRS) [22], (iii) the systematic coronary risk evaluation score (SCORE) [23], (iv) the Reynolds risk score (RRS) [24], and (v) the atherosclerosis cardiovascular disease (ASCVD) [25]. Specific guidelines like the American College of Cardiology/American Heart Association (ACC/AHA) [26], the European Society of Cardiology (ESC) [27,28], and the Canadian society [29,30] are followed for predicting the CVD risk when using these calculators.

The conventional CVD calculators offer several challenges [26,27], which include (i) not being able to deal with the non-linearity between the covariates (or risk factors) [31] and the gold standard (outcomes); (ii) does not reflect a direct representation of plaque build-up in the arteries [17,32,33]; (iii) usage of ad hoc threshold for CVD risk stratification and lack granularity for CVD [34,35]; and (iv) finally, the lack of usage of cohort's knowledge. All the abovementioned reasons put pressure to investigate a more accurate CVD risk classification tool that can assess the proper non-invasive atherosclerotic plaque burdens by using LBBM and OBBM.

When it comes to a non-invasive framework, the risk of coronary artery disease can be estimated via the carotid artery network, because of the same genetic composition of these two arteries (see Appendix H, Figure A8: Top). Carotid artery imaging also provides an advantage to both CVD and stroke risk predictions and is often adapted to act as a surrogate type of biomarker for CVD risk classification [36]. Generally, for imaging, the carotid arteries, the popular three medical imaging modalities used are magnetic resonance imaging (MRI) [37], computed tomography (CT) [38], and ultrasound (US) [39].

Carotid B-mode ultrasound (cBUS) offers several benefits, namely cost-effectiveness, user-friendliness, easy reach through the neck window, high-resolution via compound, and harmonic imaging [39–41]. Carotid videos can be also generated in the form of movies (so-called CINIE loop with cardiac gating) during imaging, which can then be used for better carotid plaque vulnerability. This can be accomplished by correlations and

characterization [42] by taking the advantage of image registration paradigms between the slices. The phenotypes for carotid ultrasound image-based (CUSIP) technique are carotid intima-media thickness (cIMT) [43–47], intima-media thickness variability (IMVT) [48–51], maximum plaque height (MPH) [52–54], and total plaque area (TPA) [55–57] and can be obtained using cBUS frozen scans. The classification of risk for CVD can be improved in terms of more reliable results by fusing CUSIP biomarkers along with the OBBM, LBBM as shown by AtheroEdge 2.0 (Roseville, CA, USA) [36]. Though it is fully automated and statistically based, it does not use cohort’s knowledge and Artificial Intelligence (AI) framework. Therefore, a more accurate solution is needed to handle this challenge to ensure reliable and superior CVD risk prediction.

With the advancement of AI in the field of healthcare [19,58–62], especially in machine learning (ML), deep learning (DL), combined with mobile solutions such as e-health and cloud-based technologies, CVD risk assessment has shown promising signs. The main focus of the proposed study is the ML paradigm however, we very briefly touch on DL strategies due to their infancy stage. Recently, we have seen research showing that ML can handle non-linearity between the input covariates and target outcomes (or gold standard), while DL automates the feature extraction process from the input data for classification. We therefore hypothesize that CVD classification paradigms such as multiclass, multi-label, and ensemble are more accurate and reliable. Due to the amalgamation of the linear and non-linear covariates along with the gold standard, there is no clear-cut defined strategy when adapting these three paradigms for CVD risk stratification. This can sometimes lead to over-performance inaccuracies and under-performance in clinical outcomes leading to bias in AI [63]. The proposed study also presents the bias measurements in these three paradigms independently, and further when all the three sets of techniques are jointly taken into consideration for CVD risk stratification. The pseudo-code for each technique is discussed in Appendices A–C. With the evolution of fast-growing telecommunication technology, these CVD techniques can be applied in e-health frameworks such as mobile or cloud settings, which provide access to the patient population for rural areas of the world. This review further dwells in the above-mentioned area. Lastly, due to changing environmental conditions such as COVID-19, it is important to understand how the CVD risk assessment integrates into the COVID-19 framework. Several CVD reviews are already available [64–69], but none of these consider the recent advanced methods like using ML and DL in office-based, mobile/cloud-based set-ups.

The design of the proposed review is as follows. Section 2 shows the PRISMA strategy used for study selection along with the statistical distribution of AI attributes. Section 3 presents the biological link between atherosclerosis and CVD risk. Section 4 represents the heart of the system discussing the three paradigms, namely multiclass, multi-label, and ensemble-based CVD risk stratification along with performance evaluation (PE) metrics for these techniques. Section 5 presents the bias in AI for these three methods. The CVD risk assessment through mobile, e-Health, and cloud-based techniques is presented in Section 6. The critical discussion of the review is in Section 7, while the study concludes in Section 8.

2. Search Strategy and Statistical Distributions

The statistical distribution of the literature is necessary to understand the types of CVD methods, the gold standard adapted for these AI-based solutions, the participation of the feature extraction methods, and bias in the AI-based solutions. Thus, we adapt the PRISMA model for the selection of the studies for the CVD risk assessment. This section is therefore divided into two parts: Section 2.1 discusses the study selection criteria and Section 2.2 presents the statistical distributions.

2.1. PRISMA Model

The PRISMA model was used for searching and selecting the final studies for the review. The search was done using Science Direct, Google Scholar, IEEE Xplore, and PubMed by adapting the following keywords “multiclass classification for CVD risk”, “multi-label

classification for CVD risk”, “ensemble classification for CVD risk”, “CVD risk using Machine Learning/Artificial Intelligence for multiclass”, CVD risk using Machine Learning/Artificial Intelligence for multi-label, “CVD risk using Machine Learning/Artificial Intelligence for ensemble”, “CVD risk assessment in ML/AI framework”, and “Bias in ML/AI”. The total number of ML/AI-based CVD studies is shown in Figure 1. An exhaustive search resulted in a total of 19,454 studies. The three criteria used for exclusion were (a) non-relevant studies (b) articles removed after search and screening of the studies (c) records rejected due to insufficient data. The implementation of exclusion criteria provides 19,084, 88, and 17 studies for exclusion showed by E1, E2, and E3 (Figure 1). The important scientific knowledge from these final studies was gained and the statistical classification was drawn. Further, a comprehensive analysis of the studies was done between the three techniques with the determination of AI bias.

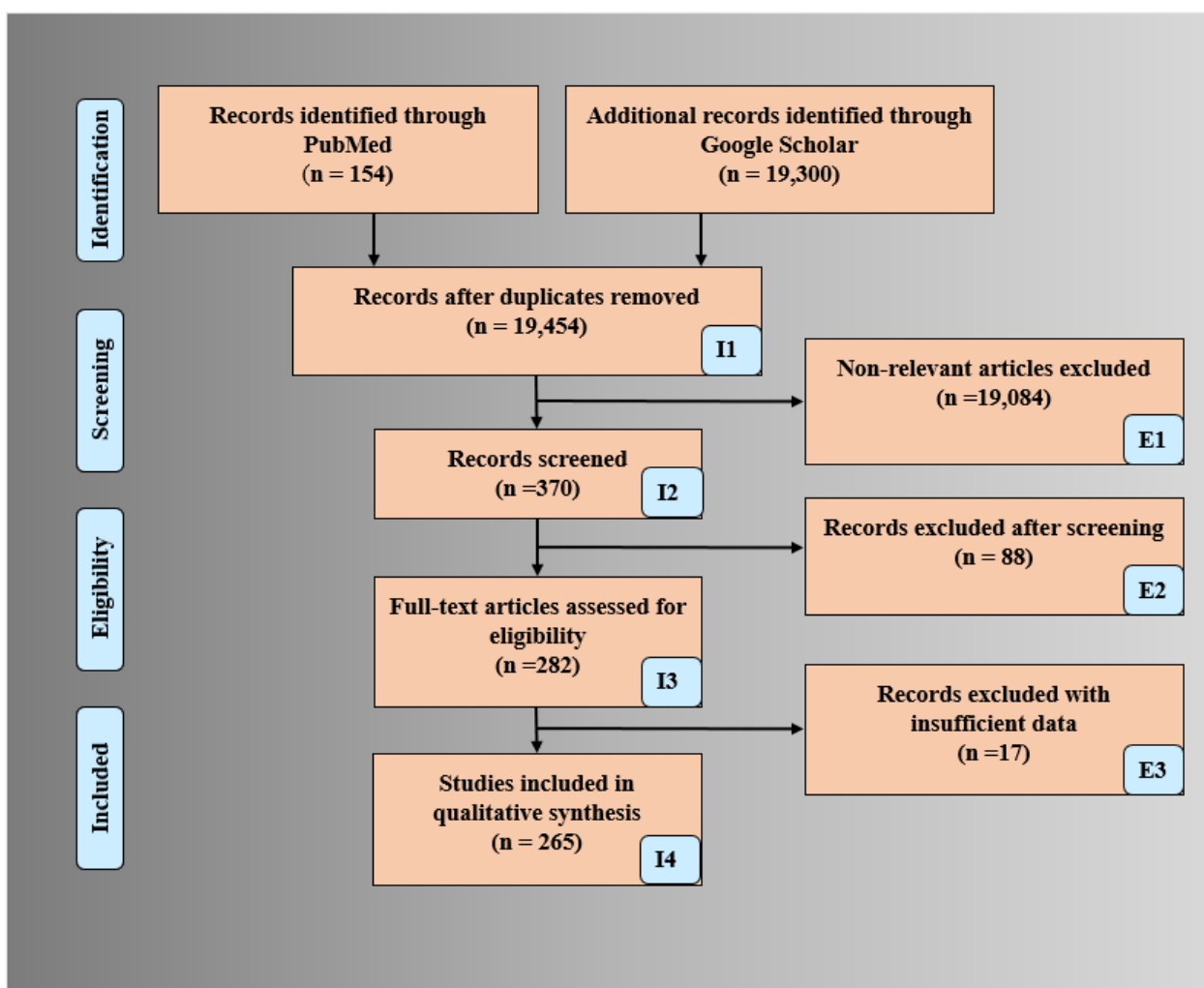


Figure 1. PRISMA model for selection of studies for CVD risk assessment.

2.2. Statistical Distribution

The statistical distributions derived from the selected studies are shown in Figure 2. The following attributes were used for the statistical distribution (a) types of CVD paradigms, (b) types of risk classes in multiclass CVD (c) ML-based CVD systems without/with feature extraction, (d) # GTs in multi-label-based CVD, (e) feature selection techniques, and (f) ML-based CVD publications.

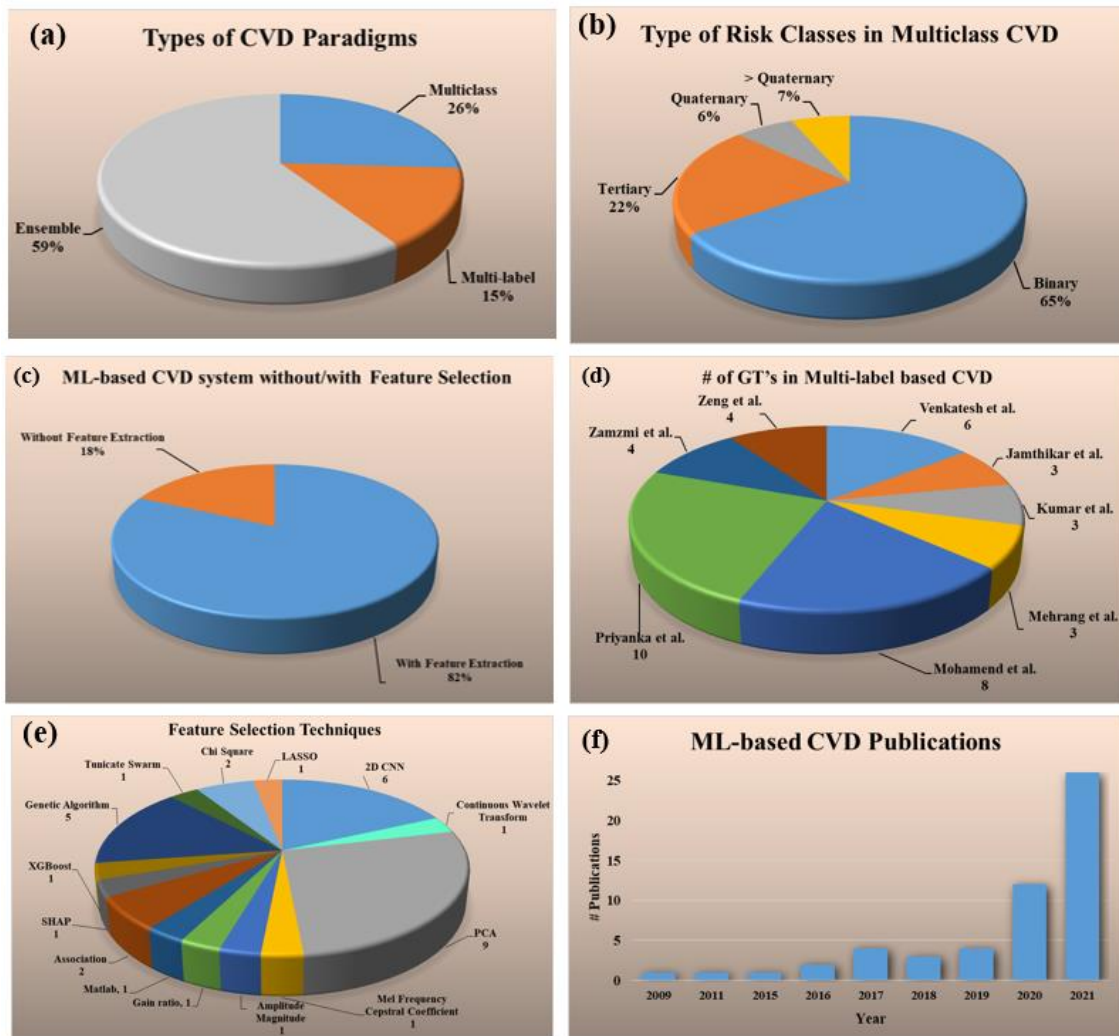


Figure 2. Statistical distribution (a) types of CVD paradigms, (b) types of risk classes in multiclass CVD (c) ML-based CVD systems without/with feature selection, (d) # GT's in multi-label based CVD, (e) feature selection techniques, (f) trend of the ML-based CVD publications by year.

The percentage of studies for each of the three kinds of CVD risk prediction had the following distributions: multiclass (26%) [69–82], multi-label (15%) [83–90], and ensemble (59%) [80,91–121] (Figure 2a). Several different kinds of risk classes were identified in multiclass CVD framework, namely binary (65%), tertiary (22%), quaternary (6%), and greater than quaternary (7%) (Figure 2b). The distribution of the ML-based CVD studies with and without feature selection are shown in Figure 2c. It was found that almost 82% of ML-based CVD studies performed feature selection for risk prediction whereas only 18% [69,70,73,75,83,94,96,110,120] did not perform it. For the ML-based multi-label CVD (Figure 2d), the total number of GT's used for each study were as follows and given in the ground braces: Venkatesh et al. (6) [83], Jamthikar et al. (3) [84], Kumar et al. (3) [85], Mehrang et al. (3) [86], Mohamend et al. (8) [87], Priyanka et al. (10) [88], Zamzmi et al. (4) [89], and Zeng et al. (4) [90]. There were eight sectors in the pie chart and each sector represents a study (publication) in the area of multi-label-based ML system. Below the study shows the number of gold standards used for the design of the multi-label ML system paradigm. For example, Venkatesh et al. had 6 types of gold standards ((death, stroke, coronary heart disease (CHD), CVD, heart failure (HF), atrial fibrillation (AF)) during the design of their multi-label-based ML system. Similarly, Jamthikar et al. had three types of gold standard (coronary artery disease (CAD), acute coronary syndrome (ACS), composite cardiovascular event (CVE)) during the design of the multi-label ML system. Since the number

of gold standards are important during the multi-label paradigm, the pie-chart shows the statistical distribution of the different studies using the number of gold standards. The number of studies (given in curly braces) that used the following feature selection techniques were 2D convolutional neural network (CNN) (6) [71,79,81,89,101,111], continuous wavelet transform (1) [72], principal component analysis (PCA) (9) [76,79,84,98,102,112,114,119,121], Mel frequency cepstral coefficient (1) [77], amplitude magnitude (1) [78], gain ratio (1) [80], Matlab (1) [86], association technique (2) [87], SHAP (1) [90], extreme gradient boost (XG-Boost) (1), genetic algorithm (5) [91,103,104,122,123], Tunicate Swarm (1) [116], chi-Square (2) [117], least absolute shrinkage and selection operation (LASSO) (1) [99] (Figure 2e). The increasing trend of CVD publications from the year 2009 to 2021 is shown in Figure 2f.

3. Biological Link between Atherosclerosis and Cardiovascular Disease

The fundamental cause of CVD is the disease of atherosclerosis [124]. The process of plaque formation is known as atherogenesis as shown in Figure 3a(A–I) [125]. It is a process when the plaques develop in the arteries where there is low endothelial shear stress [126]. The shear stress depends on the flow velocity characteristics like type of flow, direction, and velocity. Leukocytes attack the epithelium in this region (Figure 3(bA)) [126]. Mainly there is the migration of monocytes into the sub-epithelial layer where it is oxidized by the low amount of low-density lipoprotein (LDL) cholesterol and turns into macrophage (Figure 3(bB)) [127,128]. Eventually, these macrophages become large foam cells with oxidized LDL cholesterol leading to the formation of necrotic core (Figure 3(bC)). Microscopic calcium granules expand in the necrotic cells and forms lumps of calcium deposits. This necrotic core is separated from the blood vessel by a fibrous cap [129]. The blood remains uninterrupted when the plaque is small as the arteries do remodeling by themselves [130]. However, when the plaques increase, the lipid-core volume decreases leading to structural stabilization of plaque (Figure 3a) [131].

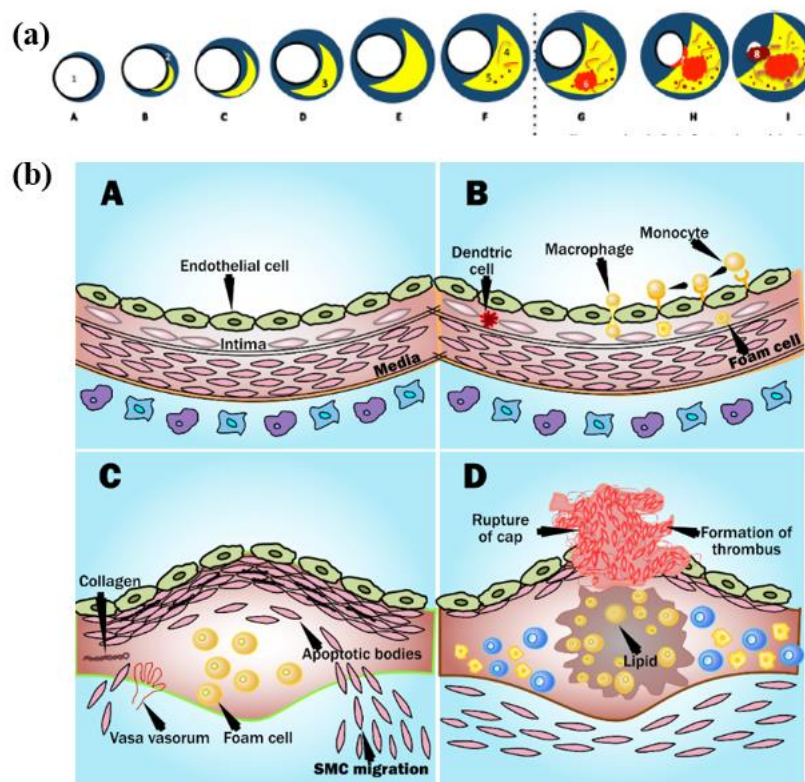


Figure 3. (a) Plaque formation in the coronary artery and (b) process of plaque rupture in coronary artery (Courtesy of AtheroPoint™, Roseville, CA, USA) [131].

Progressive deposition of lipids results in the thinning of the fibrous cap leading to rupturing the plaque [132]. The ruptures of the cap result in healing by the platelets in the bloodstream, which leads to the formation of the clot of blood or thrombus which yields blocking of artery than atrial stiffness [133]. Due to this, the tissues become deprived of blood supply, causing cell death. If the coronary artery gets blocked, causing a myocardial infarction or CVD (Figure 3(bD)) [3,7].

4. Three Paradigms for Cardiovascular Risk Stratification

The core aim of this review is to understand the three kinds of paradigms for CVD risk stratification. This allows understanding the (a) types of gold standards used for different kinds of applications, (b) types of fundamental architectures used, and (c) finally the comparison between the three different types of paradigms.

4.1. Multiclass-Based Cardiovascular Disease Risk Stratification System

The most fundamental type of CVD risk stratification is the multiclass framework [134]. There are three main characteristics in multiclass framework, namely (i) it divides the outcome into two or more granular risk classes, (ii) the drug prescription is better controlled for CVD treatments based on which class the disease stage or risk lies, and (iii) the risk of CVD is better understood when divided into several stages such as low, mild, low-of-moderate, high-of-moderate, low-of-high, and high-of-high.

4.1.1. CVD-Based Multiclass Risk Assessment System

For any CVD system, there are two most important attributes: (a) the types of the covariates used and (b) the gold standard adopted. Accordingly, in the multiclass framework, there are 14 published studies (see Table 1). It shows the three attributes represented in three columns: covariates, gold standard, and the AI category, namely ML or DL. The types of covariates considered for the multiclass systems were OBBM [71,76,80,82], LBBM [71,76,80,82], CUSIP [76,80] for *office-based setups* (Table 1: row 1–5), and Electrocardiogram (ECG) [79,81,82], PCG [77], Acceleration Plethysmogram (APG) [78] signals for *cardiac stress test laboratories* (Table 1: row 6–9), and coronary artery calcium (CAC) for CT-based CVD models [135]. The ground truths considered for CVD risk assessment (Table 1) were death [80], coronary heart disease (CHD) [82], chronic heart conditions (CHC) [79], cardiovascular event (CVE) [76], sudden cardiac death (SCD) [72], heart failure (HF), myocardial infarction (MI) [75], coronary artery calcification (CAC) score [69], fatal/non-fatal CVD [73], joint CVD and diabetes [70]. Note that these gold standard choices along with AI attributes, scientific and clinical validations are key to preventing bias in AI.

4.1.2. Comparison between CVD Application and Non-CVD Application

The comparison between CVD and non-CVD applications [136] is shown in Table 2. Seven attributes were used for this comparison. The image modalities used in the CVD-based system were US, CT, MRI, and ECG (Table 2: row 4, CVD column). The architectures applied were ML and DL. DL provided better results due to its unique automated feature selection process (Table 2: row 6, CVD column). The defined number of classes was in the range of 3–9 (Table 2: row 5, CVD column) [69–82]. The multiclass approach for classification has been applied to non-CVD applications such as Alzheimer's prediction or different cancer types. The interpretation of multiclass in the non-CVD system can be thought of as different stages of the diseases, for example, in the case of Alzheimer's disease (AD), it can be categorized as different stages of memory loss with age. Similarly, in the case of cancer, it can be different stages or grades of cancer.

Table 1. Multiclass 14 CVD studies and their characteristics in ML/DL framework.

SN	Studies	Input Covariates	Gold Standard Types	#RC	ML/DL
1	Chao et al. [71]	OBBM, LBBM	CVD Event	K	DL
2	Lui et al. [79]	ECG parameters	CHC	3	ML
3	Wiharto et al. [82]	OBBM, LBBM, ECG	CHD	3	ML
4	Jamthikar et al. [76]	OBBM, LBBM, CUSIP	CVE	4	ML
5	Nakanishi et al. [80]	OBBM, LBBM, CUSIP	Death	3	ML
6	Devi et al. [72]	ECG Parameters	SCD	3	ML
7	Khan et al. [77]	PCG Signals	CVE	3	ML
8	Krupa et al. [78]	APG signals	BCVD	3	ML
9	Ni et al. [81]	ECG Signals	CVD, No CVD	4	DL
10	Hedman et al. [74]	OBBM, LBBM	Heart Failure	3	ML
11	Hussain et al. [75]	OBBM, LBBM, ECG	MI	3	ML
12	Sanchez et al. [69]	OBBM, LBBM	CAC score	9	ML
13	Emaus et al. [73]	OBBM, CAC (CT)	F/NF CVD	3	DL
14	Buddi et al. [70]	OBBM, LBBM	CVD, Diabetes	4	ML

SN: Serial number; APG: Acceleration plethysmogram; CHD: Coronary heart disease; CVE: Cardiovascular events; CHC: Chronic heart conditions; SCD: Sudden cardiac death; BCVD: Binary CVD (Healthy, diseased); F/NF CVD: Fetal/Non-fetal CVD; CT: Computed tomography; #RC: Risk classes; OBBM: Office-based biomarkers; LBBM: Laboratory-based biomarkers; CUSIP: Carotid ultrasound image phenotypes; CAC: Coronary artery calcium; ECG: Electrocardiogram; MI: Myocardial infarction.

Our observations show that the gold standard types in the non-CVD system are very different from the CVD systems. For example, for the early detection of AD/Mild Cognitive Impairment (MCI), the classification is done between (1) AD vs. normal control (NC), (2) MCI vs. NC, (3) AD vs. MCI, and (4) progressive MCI (PMCI) vs. Significant Memory Concern (SMCI) for Alzheimer's. In the case of breast cancer, GTs can be proliferation and non-proliferation cancer types.

Note that the number of classes considered for multiclass differs from disease-to-disease. The different architecture followed for CVD are mainly ML and DL, whereas for non-CVD it ranges from deep learning retinal CAC score (RetiCAC) [137], pooled cohort equation (PCE) [138,139], support vector machine (SVM) [70,75–77,140], convolutional neural networks (CNN), decision tree (DT) [71,79], random forest (RF), logistic regression (LR), naive Bayesian (NB), K-nearest neighbor (KNN), and ensemble. The different types of covariates for no-CVD-based systems were breast histopathology images (BHI), OBBM, and LBBM (Table 2: row 2, column non-CVD). Modalities for the non-CVD-based system were EEG, MRI, CT [137,139] (Table 2: row 4, non-CVD column), and the number of risk classes varied from 5–14 [137–139,141,142] (Table 2: row 5, non-CVD column).

4.1.3. Multiclass CVD Architecture for Office-Based CVD Risk Stratification

The architectures opted for multiclass prediction of CVD risk has very basic components (a) data collection (b) training system, and (c) testing system. The training system is basically used for training the ML system based on different covariates (or risk factors) [143,144], with the support of different ground truths while using the training-based classifiers. The system can be trained to identify the granular risk classes from no, low, and medium, to high class. Feature selection is also performed during the training of the system [145,146]. For prediction, the training model is applied to transform the testing features either in Seen AI framework or the Unseen AI framework [147]. Two types of architectures were described in this section in terms of the above-mentioned factors. A typical online system for multiclass CVD risk stratification is shown in Appendix A, Appendix A.1.

Table 2. Multiclass in CVD vs. non-CVD using seven attributes.

SN	Attributes	Multiclass CVD	Multiclass Non-CVD
1	Ground truth types	CVE [69–73,76–79,81,82], HF [74], MI [75], Death [80]	AD, NC, MCI, PMCI vs. SMCI [141], Proliferation, NP [139], ADH, DCS, IC [137,138,142]
2	Covariates types for the ML design	OBBM [69,70,73–76,80,82], LBBM [69,70,73–76,80,82], CUSIP [71,72,76–82], MU [76]	BHI [139], OBBM [137,138,141,142], LBBM [137,138,141,142]
3	Disease Type	CVD [69–82]	Diabetes [142], Cancer (Breast, Lung, Brain) [138,139], Alzheimer’s [138,141], Retinal [137]
4	Image Modalities	ECG, CT, US [71,72,76–82]	EEG, MRI, CT [137,139]
5	# Classes	3–9 [69–82]	5–14 [137–139,141,142]
6	Architecture Type	ML [70,72,76–80,82], DL [71,81]	ML, rMLTFL [141]
7	Classifiers used	SVM [70,75–77], DT, RF, LR, NB, KNN, CNN [71,79]	RetiCAC [137], PCE, SVM, CNN, DT, LR, NB, SVM, KNN, ensemble [138,139]

SN: Serial number; CVE: Cardiovascular event; AD: Alzheimer’s; NC: Normal control; MCI: Mild Cognitive impairment; PMCI: progressive MCI; SMCI: Significant memory concern; HF: Heart failure; MI: Myocardial infarction; OBBM: Office-based biomarkers; LBBM: Laboratory-based biomarkers; CUSIP: Carotid ultrasound image phenotype; ECG: Electrocardiogram; CT: Computed tomography; US: Ultrasound; MRI: Magnetic resonance imaging; BHI: Breast histopathology images; MU: MedUse; IM: Image modalities; SVM: Support vector machine; KNN: K-nearest neighbor; DT: Decision tree; RF: Random forest; LD: Logistic regression; NB: Naive Bayesian. RetiCAC: Deep learning retinal CAC score; PCE: Pooled cohort equation; rMLTFL: robust multi-label transfer feature learning.

A generalized ML system is applied to *office-based* CVD or *stress-test-based* CVD systems as shown in Figure 4. Considering the *office-based* CVD system, the covariates were collected from OBBM, LBBM, CUSIP, and MedUSE [76], while for the CVD-based stress-test system, EEG was the input. The rest of the configuration remains the same which consists of four parts: Part A is the preprocessing of the input data (covariates) and augmentation for balancing the classes. Part B consists of a training system, Part C consists of a prediction system, and Part D consists of a performance evaluation system (Appendix E). In Part A, the objective is to balance the classes if there is a multiclass scenario, Part B consists of two subparts: (i) selection of the best feature given the set of covariates and (ii) model generation using (a) classifier, (b) selected features, and the (c) gold standard. Part C consists of the application of the trained model on the selected set of best features from the test data set by transforming the test features to compute the predicted label. Part D is used for performance evaluation of the ML system where the predicted labels are compared against the gold standard labels. Note that during the training system, the two ingredients are the classifier bank and the gold standard used. The classifier bank, for example, can be classifiers like SVM, XGBoost, KNN, NB, etc., while the gold standard is the coronary artery disease syndrome, such as coronary artery disease stages that include the four types of risk stages. Note that since the system is a K-fold (either of the K types such as K2, K3, K4, K5, and K10 can be used), every patient gets to be in the test pool, and then at the end of all the folds, the complete set can be used for performance evaluation. Further to note a classifier bank can be used during the design of the training model, that uses the gold standard (such as coronary risk scores derived from coronary angiography) and training covariates. The CVD example in Figure 4 uses four sets of covariates, which can be flipped to ECG signals [148–150] when using the stress test-based system for CVD risk assessment. The longitudinal ultrasound model is used typically for the collection of the CUSIP risk factors such as cIMT (max., min., and ave.), intima-media thickness variability (cIMTV), maximum plaque height (MPH), and total plaque area (TPA).

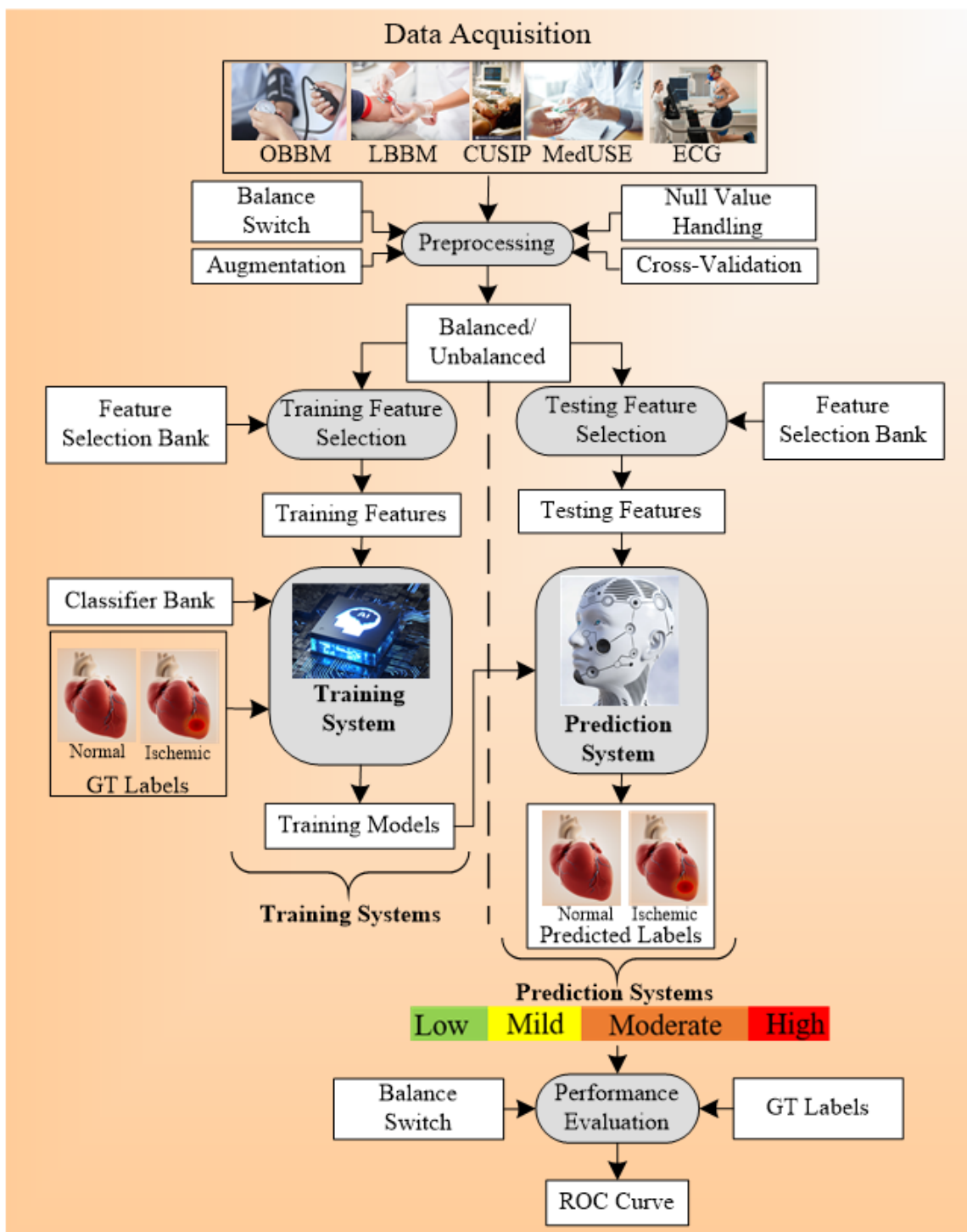


Figure 4. Multiclass architecture for CVD risk stratification (AtheroEdge 3.0_{ML}).

4.1.4. Multiclass CVD Architecture for Cardiac Stress Laboratories

Another set of architecture for multiclass CVD risk prediction was used by Hussein et al. [75] (Figure 5). The ECG signals [151–153] are obtained from the stress test laboratory for the analysis of CVD risk. The model uses the multiclass SVM classifier that takes the ECG signals as risk factors or covariates. And the ground truth used for the training system is myocardial infarction (MI). The multiclass outcomes that were identified were normal, low MI, and high MI. The feature of ST (it is the interval between ventricular depolarization and repolarization, and PR (the flat line that runs from the end of the P-wave till the start

of the QRS complex) were extracted from the time-frequency (TF) power spectrum. The created training model was the input to the prediction systems along with the test data and the final classifications were made into the normal, low MI, and high MI.

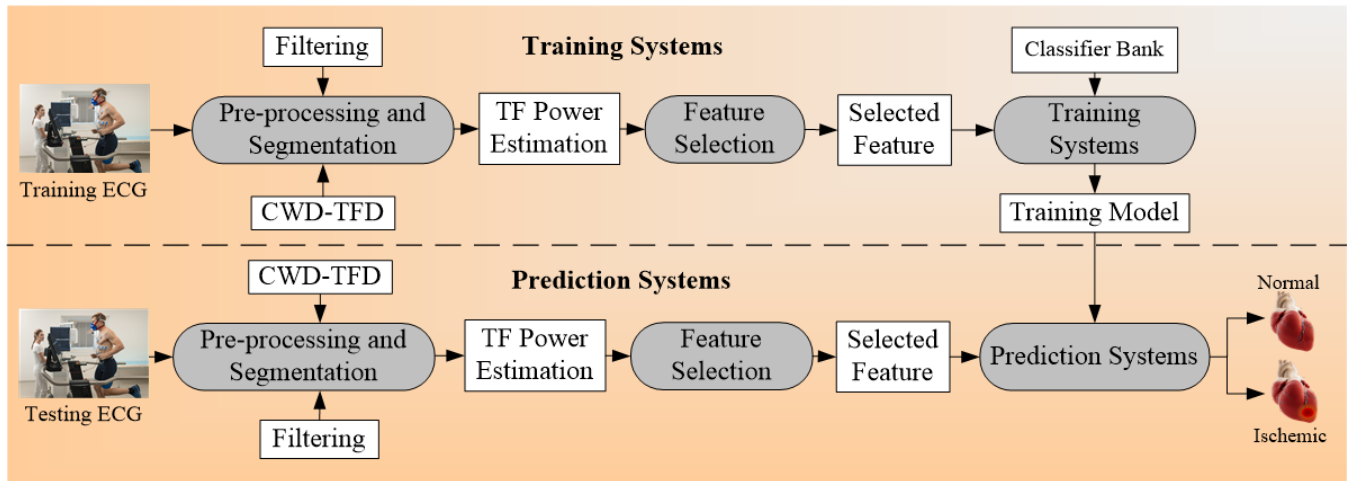


Figure 5. Example of multiclass architecture; CWD: Choi-William’s time-frequency distribution; TF: time-frequency.

The general algorithm for multiclass CVD risk stratification is explained in form of pseudo-code. A detailed explanation is provided in Appendix A, Appendix A.2.

4.2. Multi-Label-Based Cardiovascular Disease Classification

The second technique used for CVD risk stratification is multi-label-based [154–156]. The ground truth is very important for the proper classification of CVD risk [157–159]. CVD risk prediction systems were said to be multi-label-based depending on the number of ground truth (GT) used in the system [160–162]. The paradigm was considered as a multi-label-based classification if more than one number of GT is used for CVD risk detection [90,163–167]. The GTs, risk factors, and the architecture used were discussed in the next sub-sections. The pseudo-code that represents a multi-label-based risk stratification process can be referred to in Appendix B.

4.2.1. Covariates and Risk Factors for Multi-Label-Based CVD Classification

Eight multi-label-based studies for CVD risk prediction were considered in this review [83–90]. Different types of ground truths used in these studies were death, stroke, CHD, CVD, HF, atrial fibrillation (AF) [83], CAD, ACS, composite CVE [84], large vessel disease (LVD), small vessel disease (SVD) [168], intracerebral hemorrhage (ICH) [85], non-AFib-non-ADHF, AFib-non-ADHF, AFib-ADHF [86], systolic heart failure (acute, chronic type), diastolic heart failure (acute and chronic type) [87], congestive heart failure, hypertension, AF, acute kidney failure, diabetes type II, acute respiratory failure, hyperlipidemia, coronary atherosclerosis, urinary tract infection, esophageal reflux [88], CAD, dilated cardiomyopathy (DCM), MI [89], lung complication, cardiac, infectious and rhythmic complication [90].

The risk factors used were OBBM, LBBM, CUSIP, MRI, and CT image phenotypes (input covariates column, Table 3). The algorithms used for the multi-label classifications were namely binary recursive (BR), label powerset (LP), multi-label adaptive resonance associative map (MLARAM), random k-labelset (RakEL), classifier chain (CC), multi-label k-nearest neighbor (MLkNN), seismocardiography (SCG-Z), gyrocardiography (GCG-Z), principal component analysis (PCA), DCT, consensus-based risk model. Other characteristics of this classification technique were described in Table 3.

Table 3. Multi-label 8 studies and their characteristics.

SN	Studies	Input Covariates	Ground Truth	ML/DL
1	Venkatesh et al. [83]	OBBM, LBBM	Death, Stroke, CHD, CVD, HF, AF	ML
2	Jamthikar et al. [84]	OBBM, LBBM, CUSIP	CAD, ACS, Composite CVE	ML
3	Kumar et al. [85]	OBBM, LBBM, ECG	LVD, SVD, ICH	ML
4	Mehrang et al. [86]	OBBM, LBBM, CUSIP	Non-AFib-Non-ADHF, Afib-Non-ADHF, Afib-ADHF	ML
5	Mohamend et al. [87]	OBBM, LBBM, CUSIP	SHF, ASHF, CSHF, ACSHF, DHF, ADHF, CDHF, ACDHF	ML
6	Priyanka et al. [88]	OBBM, LBBM	HT, CHF, AF, CA, AKF, Dia-TII, HL, ARF, UTI, ER	ML
7	Zamzmi et al. [89]	MRI, CT Signals	HF, CAD, DCM, MI	DL
8	Zeng et al. [90]	OBBM, LBBM	LC, CC, IC, RC	ML

SN: Serial number; HF: Heart failure; AF: Arterial fibrillation; LVD; Large vessel disease; SVD: Small vessel disease; ICH: Intracerebral hemorrhage (ICH); SHF: Systolic heart failure; ASHF: Acute systolic heart failure; CSHF: Chronic systolic heart failure; ACSHF: Acute on chronic systolic heart failure; DHF: Diastolic heart failure; ADHF: Acute diastolic heart failure; CDHF: Chronic diastolic heart failure; ACDHF: Acute on chronic diastolic heart failure; HT: Hypertension; CHF: Congestive heart failure; CA: Coronary atherosclerosis, AKF: Acute kidney failure; HL: Hyperlipidemia; Dia-TII: Diabetes Type II; ARF: Acute respiratory failure; UTI: Urinary tract infection; ER: Esophageal reflux; DCM: Dilated cardiomyopathy LC: Lung complication, CC: Cardiac complication; IC: Infectious complication, RC: Rhythmia complication.

4.2.2. Multi-Label-Based Architectures for CVD Risk Stratification

The architecture design for the multi-label plays an important in the outcome results of the system. The basic component of the architecture for the CVD prediction system is training and testing. The proper choice of GT leads to non-biased results in the risk prediction of CVD. The architecture system used by Jamthikar et al. [84] is shown in Figure 6 below. The total number of ground truths considered for this system were three, namely (a) coronary artery disease, (b) acute coronary syndrome, and (c) a composite CVE, and the covariates used were OBBM, LBBM, and the CUSIP phenotype. Six types of classification techniques used include (i) four problem transformation methods (PTM) and (ii) two algorithm adaptation methods (AAM) are used for multi-label CVE prediction. The four PTM techniques were binary relevance (BR), label powerset (LP), classifier chain (CC), and random k-labelset (RAkEL). Under AAM-based, two techniques, namely multi-label k-nearest neighbor (MLkNN), and multi-label adaptive resonance associative map (MLARAM) were used. The details can be seen in Appendix B. Evaluation was performed by calculating the accuracy, sensitivity, specificity, F1-score, and AUC for all the classification techniques. The BR classification was found to be the best performer with the values for accuracy, sensitivity, specificity, F1-score, and AUC as 81.2%, 76.5%, 83.8%, 75.37, and 0.89 ($p < 0.0001$), respectively.

Another architecture [86] used for multi-label CVD classification is described in Figure 7. The mechanocardiography (MCG) data were used by the system. Four kinds of ground truth were used, namely AFib, non-AFib, ADHF, and non-ADHF. The covariates were gender, age, height, weight, BMI, given for the training and testing system. The ML classification algorithm used were random forest (RF), Xtrem Gradient Boost (XGB), and logistic regression (LR). RF gave the best performance among all the three ML classifiers. The system was validated by nested cross-validation. In this system, feature extraction was also performed using a feature vector. The hierarchal classification was also adapted in this system. Another paradigm that can use multiple classifiers at the same time is under the ensemble framework as presented in the next section.

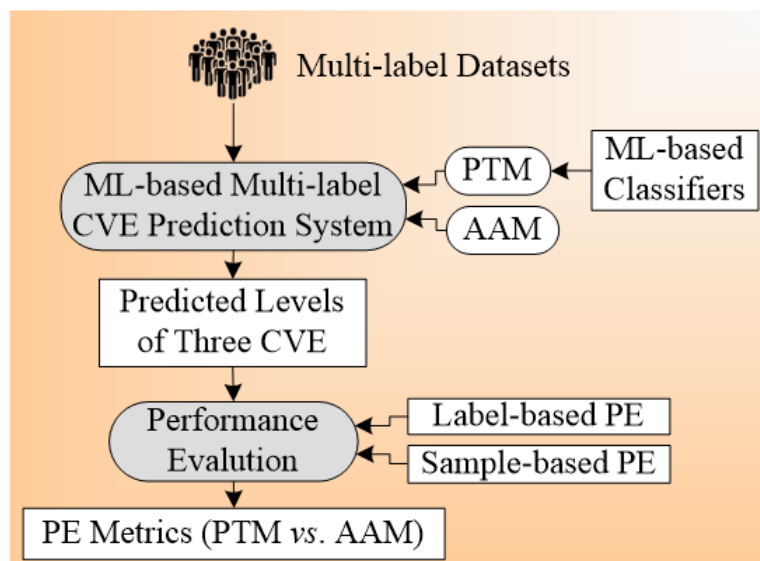


Figure 6. Architecture for multi-label-based CVD risk classification using carotid ultrasound.

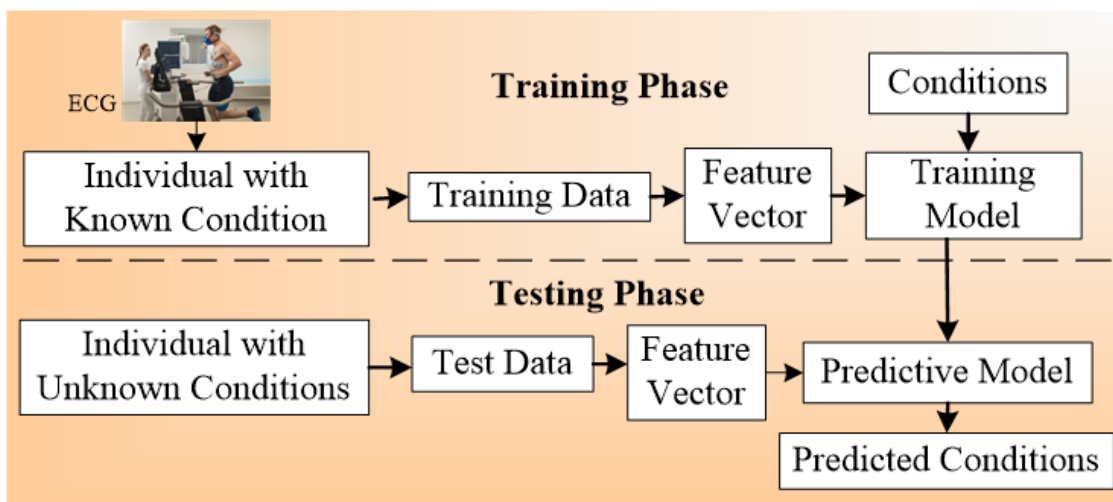


Figure 7. ECG architecture for multi-label-based CVD classification.

4.3. Ensemble-Based Cardiovascular Disease Classification

The ensemble-based technique was the third type of technique considered for CVD risk classification [169–171]. This classification was characterized by the fusion of different types of ML or DL classifiers (Table 4). It can be used with multiclass and multi-label classification [172–174]. Figure 8 shows the concept of the ensemble paradigm. There are two sets of strategies, namely homogeneous ensemble and heterogeneous ensemble (see the separation shown by dotted line). In homogenous ensemble, the conventional classifier techniques are combined using homogeneous ensemble algorithm to yield homogeneous ensemble classifier, which when trained using classifier A while using the gold standard. This homogeneous system yields the trained model A. The same protocol can be adapted for the heterogeneous ensemble paradigm yielding the trained model B. These trained models can be used by the prediction system on the test feature to produce prediction labels. Finally, the performance can be evaluated by comparing predicted labels to gold-standard labels yielding performance parameters. The key benefit of using an ensemble classifier is its superior performance compared to either multiclass or multi-label strategies. The pseudo-code that represents the ensemble-based risk stratification process can be seen in Appendix C. The ensemble technique can be applied to the CVD field, as well as to other fields, such as education, Alzheimer’s, etc.

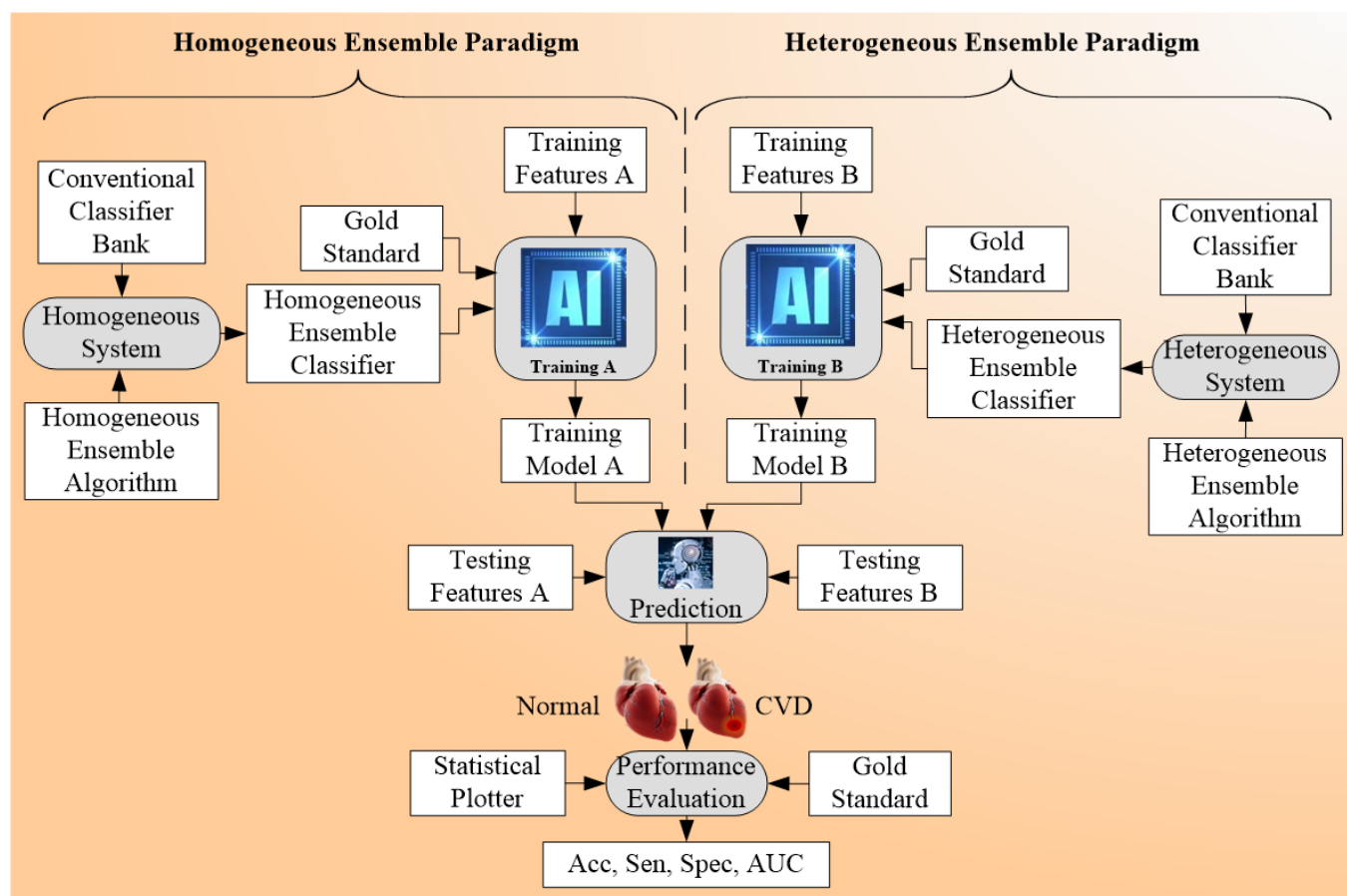


Figure 8. Ensemble-based Architecture for CVD risk stratification.

4.3.1. Different Classifier Combination for Ensemble-Based CVD Risk Stratification

The different classifiers used in ensemble techniques were kNN, Reglog, GaussNB (GNB), linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), random forest (RF) [91,95–98], multilayer perceptron (MLP), SVM [91,94,95,97,101,103,104], CNN, long short term memory network (LSTM), gated recurrent unit (GRU), bidirectional LSTM, bidirectional GRU [92], bagging, XGBoost, Adaboost [93,99], DNN [94], generalized additive models (GAMs), elastic net, penalized logistic regression (PLR), gradient boosted machines (GBMs), Bayesian logistic regression [96], K-NN [98,99,102,104,121], NB [101,104], light GBM, GBDT, LR, BPNN, DT [98,99,104,109], GB [99], Adaboost ensemble [100], ANN [101,104], GNB, LDA, LR, QDA, AdaBoost [105,113,118], XGBoost [102,118], ensemble SVM [104], CART [106], bagging, VS, LASSO, boosting, Bassian, MARS, logistic [107], ensemble boosting [80], ensemble learning, deep learning [108], ET, sequential minimal optimization (SMO), IBk, AdaBoostM1 with decision stump (DS), AdaBoostM1 with LR, REPTree, [109], neural network (NN), GB [110,114], linear Cox model [110], ensemble gradient boosting [111], ET [112], NB, multi-layer defense system (MLDS) [114], average-voting (AVen), majority-voting (MVEn), weighted-average voting (WAVen) [115], HTSA, ensemble deep learning [116], XGBoost Meta [117,119], SOM [120], extreme learning machine (ELM) [121].

4.3.2. Comparison between the Three Types of CVD Risk Assessment Systems

All the architecture can be combined to achieve the functionality of all the three models, namely multiclass, multi-label [13], and ensemble. Both multiclass, multi-label modalities can be combined with the ensemble to acquire a better accuracy in the prediction of CVD risk. The comparison between the three has been shown in Appendix D, Table A1. The data size varies from 212–66,363 (for multiclass) [69–82], 300–46,520 (for multi-label) [83–90],

459–823,627 (for ensemble) [80,91–121]. The number of risk factors for multiclass is low, multi-label is more, and for the ensemble is moderate. The risk factors considered for multiclass are family history and BMI. For multi-label-based studies and ensemble-based studies, the risk factors considered were BMI, ethnicity, hypertension, and smoking. The image modalities used for multiclass and multi-label were MRI [175,176], ECG [177–179], and CUSIP whereas ECG is not used in ensemble-based studies. The range of performance evaluation parameters used for the multiclass, multi-label, and ensemble was 1–5, 1–8, and 1–8, respectively. The different types of classifiers used for these three techniques were SVM [91,94,95,97,101,103,104], RF [91,95–98], CNN, DT, k-NN, Agatston classifier, Elastic Net, NN, NB, XGBoost, SVM, ELM, one against one (OAO), one against all (OAA), decision direct acyclic graph (DDAG), exhaustive output error correction code (ECOC) [69–82]. The power analysis is also done on more multi-label and ensemble-based techniques. The detailed description can be seen in Appendix F. The general presentation of the NN algorithm was made in Appendix H.1 Right. The ML-based systems also lead to bias as it lacks clinical evaluation which is discussed in the next section.

Table 4. Ensemble-based 33 and their characteristics of ML-based.

SN	Studies	Input Covariates	Ground Truth	ML/DL
1	Abdar et al. [91]	OBBM, LBBM	CAD	ML
2	Baccouche et al. [92]	OBBM, LBBM	HHD, IHD, MHD, VHD	DL
3	Chu et al. [93]	OBBM, LBBM, ECG	CVD, Dia	ML
4	Cai et al. [94]	OBBM, LBBM	CR	ML
5	Esfahani et al. [95]	OBBM, LBBM	CVD	ML
6	Gibson et al. [96]	OBBM, LBBM	ACS	ML
7	Gao et al. [97]	OBBM, LBBM, ECG	CVD, BC	ML
8	Gao et al. [98]	OBBM, LBBM	CVD	ML
9	Gosh et al. [99]	OBBM, LBBM, ECG	CVD	ML
10	Honsi et al. [100]	OBBM, LBBM	CVD	ML
11	Jan et al. [101]	OBBM, LBBM, ECG	HD	ML
12	Jamthikar et al. [102]	OBBM, LBBM, CUSIP	CAD, ACS	ML
13	Jothiprakash et al. [103]	OBBM, LBBM	CVD	ML
14	Liu et al. [104]	OBBM, LBBM	CA	ML
15	Miao et al. [105]	OBBM, LBBM, ECG	CHD	ML
16	Mienye et al. [106]	OBBM, LBBM	HD	ML
17	Negassa et al. [107]	OBBM, LBBM	HF	ML
18	Nakanishi et al. [80]	OBBM, LBBM, CT	Death	ML
19	Plawiak et al. [108]	OBBM, LBBM, ECG	Arrhythmia	DL
20	Puvar et al. [180]	OBBM, LBBM, ECG	HD	ML
21	Reddy et al. [109]	OBBM, LBBM	HD	ML
22	Rousset et al. [110]	OBBM, LBBM	CVD	ML
23	Sherly et al. [111]	OBBM, LBBM, ECG	HD	ML
24	Sherazi et al. [112]	OBBM, LBBM	CVE	ML
25	Tan et al. [113]	OBBM, LBBM	CVD	ML

Table 4. Cont.

SN	Studies	Input Covariates	Ground Truth	ML/DL
26	Uddin et al. [114]	OBBM, LBBM	CVD	ML
27	Velusamy et al. [115]	OBBM, LBBM	CAD	ML
28	Wankhede et al. [116]	OBBM, LBBM	HD	DL
29	Yadav et al. [117]	OBBM, LBBM	HD	ML
30	Ye et al. [118]	OBBM, LBBM	HYT	ML
31	Yekkala et al. [119]	OBBM, LBBM	CVD	ML
32	Zarkogianni et al. [120]	OBBM, LBBM	CVD, Dia	ML
33	Zhenya et al. [121]	OBBM, LBBM, ECG	HD	ML

SN: Serial number; HHR: Hypertensive heart disease; IHD: Ischemic heart disease, MHD: Mixed heart disease; VHD: Valvular heart disease; CR: Cardiac resynchronization; ACS: Acute coronary syndrome; CVD: Cardiovascular disease; CA: Cardiac arrhythmia; BC: Breast cancer; HD: Heart disease; HF: Heart failure; CVE: Cardiovascular event; Dia: Diabetes.

4.4. Performance Evaluation Metrics for Multiclass, Multi-Label, and Ensemble Techniques

Performance evaluation (PE) strategies are very vital for understanding the reliability of the ML-based CVD risk stratification systems. The main metrics used by the PE systems are sensitivity, specificity, accuracy, precision, F1-score, positive predictive value (PPV), negative predictive value (NPV), false-positive rate (FPR), false-negative rate (FNR), p -value, hamming loss, C-index in multiclass, multi-label, and ensemble-based CVD risk assessment systems. The formulae used for determining these parameters are described in Appendix E. These different PE strategies were analyzed in different techniques. It was found that PE for multi-label-based CVD is different as compared to multiclass and ensemble. There are two types of PE techniques for multi-label, namely label-based and instance-based PE. The label-based is done using micro and macro-averaging techniques. Details of these techniques can be seen in Appendix E. Figure 9 (top) shows the label-based and instance-based performance evaluation. The number of studies that used this PE parameter is the accuracy (46) followed by sensitivity (32), precision (27), F1-score (27), specificity (26), p -value (10), PPV (8), NPV (6), FPR (6), FNR (5), c-index (4), Hamming Loss (1). Hamming Loss has opted only for the ensemble-based CVD risk stratification [181–184]. The PE metrics used in the stress test-based (ECG) [185–187] techniques are area-under-the-curve (AUC), sensitivity, specificity, PPV, and NPV [188–192].

As seen from the above discussion, the most important characteristic of the multiclass paradigm is the selection of gold standards having greater than two classes. The highest flexibility in the multiclass framework is the amalgamation of different sources of covariates, namely OBBM, LBBM, CUSIP, and MedUSE. We could take characteristics of plaque in the carotid ultrasound such as information about plaque symptomatology. The same principle holds in the stress test-based CVD paradigm or non-CVD framework. The ML systems sometimes overestimate the accuracies in prediction and underestimate the scientific validation, which results in bias in the prediction systems that we discuss in Section 5.

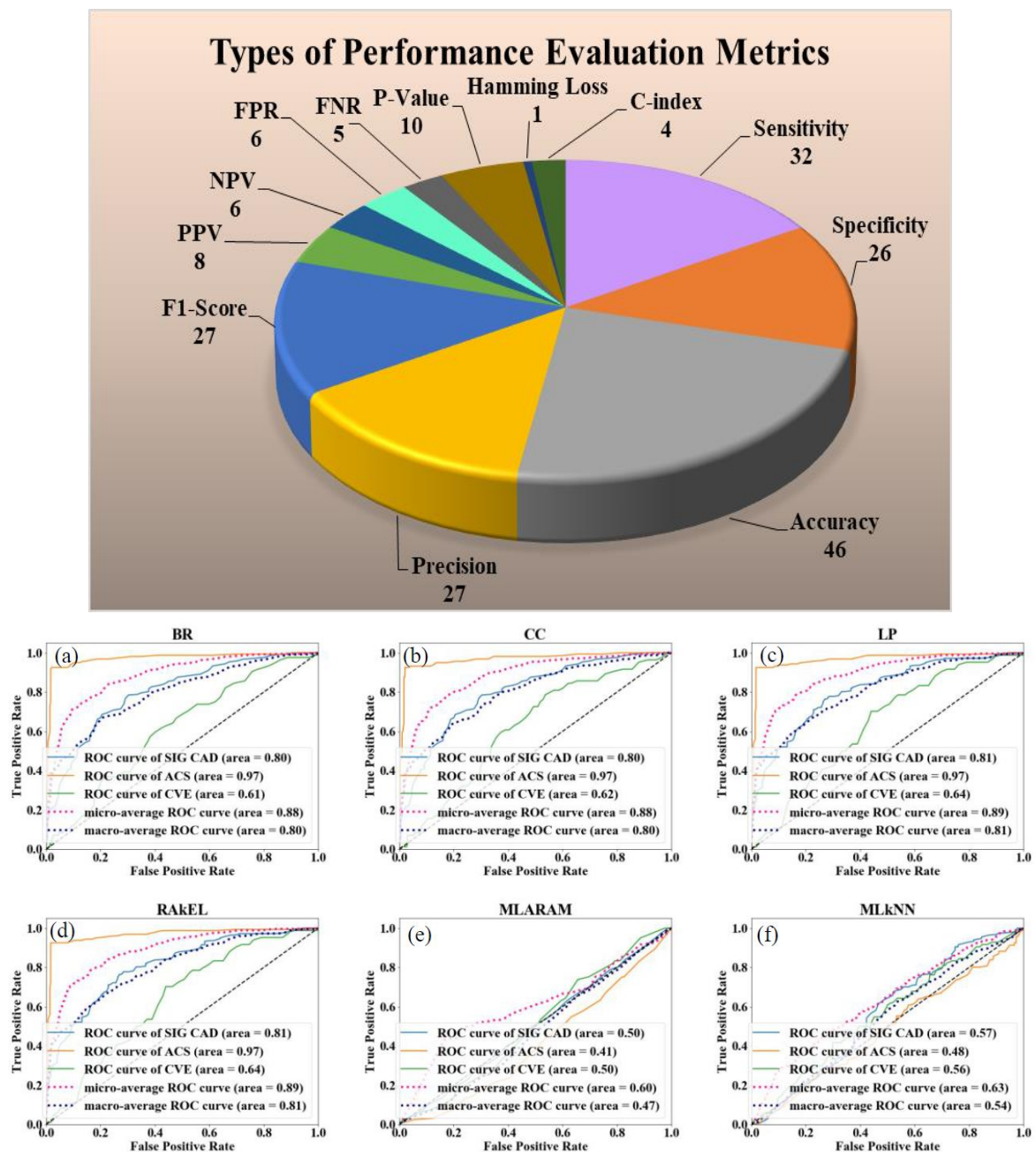


Figure 9. (Top) Types of performance evaluation metrics for ML-based CVD systems, (Bottom) Example of a ROC for multi-label-based CVD systems (Courtesy of AtheroPoint, Roseville, CA, USA) [84], PPV: positive predictive value; NPV: negative predictive value; FPR: false positive rate; FNR: false negative rate; BR: binary relevance; CC: classifier chain; LP: label powerset; MLARAM: multi-label adaptive resonance associative map; RakEL: random k-labelset; MLkNN: multi-label k-nearest neighbor; CVE: cardiovascular events; CAD: coronary artery disease; ACS: acute coronary syndrome; ROC: receiver operating characteristic; (a–f): different en-points used in the multi-label studies.

5. Bias Distribution in the ML System for Multiclass, Multi-Label, and Ensemble

The ML-based systems for CVD risk classification generate a bias due to various reasons [193–195]. Thus, it is important to understand the risk of bias (RoB) in these ML-based systems. As the ML systems were clustered in three different clusters, namely multiclass, multi-label, and ensemble, the bias nature was compared in three independent categories, and finally by considering all the three mixed together. For the RoB in the ML-based systems, the ML systems were ranked on the basis of the average mean score

along with cumulative mean values (Table 5). The mean and the cumulative score were generated by scoring the ML attributes for each study. There were 52 ML studies (14 in multiclass, 8 in multi-label, 30 in ensemble cluster) with 41 attributes each. The score was given to each AI attribute using a grading scheme [196]. In this grading scheme, a high-score was assigned to the AI attribute, if the AI attribute was adopted (used) in a particular study (publication). The score is between 0 and 5. For example, a high-score was given if the attribute “data size” had a value higher than 1000 patients, else a low-score was assigned. Similarly, as another example, a high-score of 5 was given to the attribute “feature extraction”, if it was implemented in a study, else a score of 0 was assigned, if not implemented. Later the ML-based studies were clustered into low-bias, moderate-bias, and high-bias groups. The distributions were done on the basis of the two cut-offs values. The low-moderate (LM) and moderate-high (MH) cutoff values for each cluster of ML studies were determined based on the mean values along with the cumulative-mean values. The cutoffs values obtained for the multiclass cluster are 1.8 and 1.35 for LM and MH respectively (Figure 10a). The studies belonging in the low-bias, the moderate-bias, and the high-bias bins are 4, 5, and 5, respectively. Similarly, the cutoffs for the multi-label cluster are LM: 1.9 and MH: 1.4 (Figure 10b). Multi-label-based CVD ML studies in low-bias group are 3, moderate-bias group are 3 and high-bias group is 2. The values of LM cutoff for the ensemble cluster are 1.8 and HM cutoff value is 1.6. The studies in low-bias bin are 8, in moderate-bias are 16 and high-bias bin is 6 respectively for ensemble-based ML studies (Figure 10c). Alternatively, as all the studies are based on CVD risk prediction, the LM and MH cutoffs were determined by combining all the 52 studies. The LM, HM cutoff for the combined approach is 1.9 and 1.7 respectively (Figure 10d). Thus, we see that the ensemble-based ML CVD risk estimation systems are low-biased among all the selected studies followed by multiclass-based (moderate-biased) while the multi-label-based was found to be low-biased. The AI-based CVD risk stratification systems can be further improved by incorporating the mobile, cloud, and e-health infrastructure as discussed in the next Section 6.

Table 5. Ranking table (a) multiclass studies, (b) multi-label studies, (c) ensemble studies.

(a) Multiclass Studies	Sum	Mean	Rank	(c) Ensemble Studies	Sum	Mean	Rank
Chao et al. [71]	78	1.9	1	Jamthikar et al. [102]	120.5	2.9	1
Lui et al. [79]	76.5	1.9	2	Sherazi et al. [112]	98	2.4	2
Wiharto et al. [82]	76	1.9	3	Uddin et al. [114]	94	2.3	3
Jamthikar et al. [76]	75.5	1.8	4	Velusamy et al. [115]	89.5	2.2	4
Nakanishi et al. [80]	74	1.8	5	Gao et al. [97]	85	2.1	5
Devi et al. [72]	72.5	1.8	6	Jan et al. [101]	85	2.1	6
Khan et al. [77]	71.5	1.7	7	Miao et al. [105]	84.5	2.1	7
Krupa et al. [78]	64.5	1.6	8	Gosh et al. [99]	83	2	8
Ni et al. [81]	59	1.4	9	Wankhede et al. [116]	81	2	9
Hedman et al. [74]	55.5	1.4	10	Esfahani et al. [95]	74	1.8	10
Hussain et al. [75]	53.5	1.3	11	Reddy et al. [111]	72	1.8	11
Sanchez et al. [69]	43	1	12	Rousset et al. [110]	71	1.7	12
Emaus et al. [73]	41	1	13	Yekkala et al. [119]	71	1.7	13
Buddi et al. [70]	33.5	0.8	14	Abdar et al. [91]	70.5	1.7	14
(b) Multi-label Studies	Sum	Mean	Rank	Cai et al. [94]	70	1.7	15
Jamthikar et al. [84]	111.5	2.7	1	Nakanishi et al. [80]	70	1.7	16
Venkatesh et al. [83]	108	2.6	2	Mienye et al. [106]	69	1.7	17
Mehrang et al. [86]	96.5	2.4	3	Zhenya et al. [121]	68.5	1.7	18
Zeng et al. [90]	76.5	1.9	4	Liu et al. [104]	67	1.6	19
Zamzmi et al. [89]	69.5	1.7	5	Puvar et al. [180]	67	1.6	20
Mohamend et al. [87]	60	1.5	6	Baccouche et al. [92]	65.5	1.6	21
Kumar et al. [85]	59	1.4	7	Sherly et al. [109]	64.5	1.6	22
Priyanka et al. [88]	59	1.4	8	Jothiprakash et al. [103]	64	1.6	23

Table 5. Cont.

(a) Multiclass Studies	Sum	Mean	Rank	(c) Ensemble Studies	Sum	Mean	Rank
				Negassa et al. [107]	64	1.6	24
				Ye et al. [118]	64	1.6	25
				Gao et al. [98]	63.5	1.5	26
				Honsi et al. [100]	59.5	1.5	27
				Gibson et al. [96]	55	1.3	28
				Zarkogianni et al. [120]	54.5	1.3	29
				Plawiak et al. [108]	53.5	1.3	30
				Yadav et al. [117]	53.5	1.3	31
				Chu et al. [93]	52.5	1.2	32
				Tan et al. [113]	52.5	1.2	33

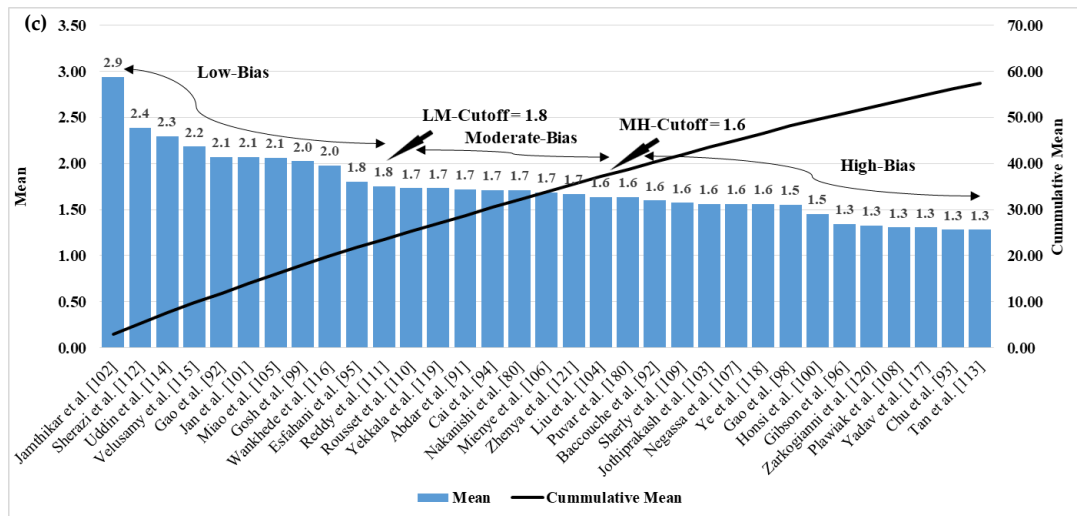
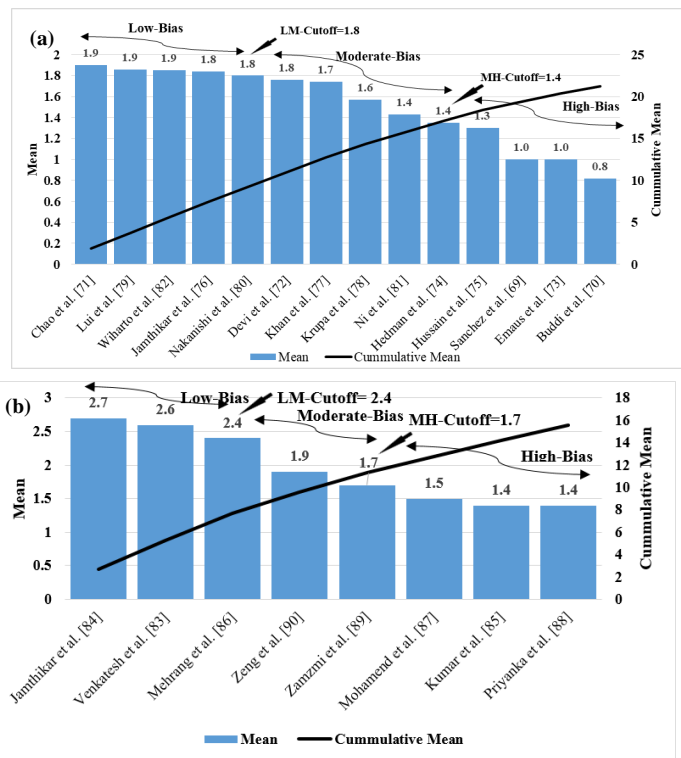


Figure 10. Cont.

based settings. Further, the scope of the study had a limited discussion on (a) CVD risk prediction using ECG signals-based settings and (b) deep learning (DL) techniques for CVD risk prediction. Therefore, the main or principal findings from this review were (i) three types of CVD risk stratification techniques, namely (a) multiclass (b) multi-label, and (c) ensemble; (ii) types of covariates used where OBBM, LBBM, MedUSE, and CUSIP. The OBBM, LBBM, MedUSE were used widely when compared to image-based phenotypes (CUSIP), which is now evolving more rapidly since is a surrogate marker for coronary artery disease; (iii) ground truth is a very vital factor so as to avoid the risk of bias (RoB) during the ML-based CVD risk prediction; (iv) popularity of the classification techniques used in the field of CVD were in the order as multiclass-based, ensemble-based, multi-label-based; (v) clinical and scientific validation is another set of AI attributes that must be accompanied in any ML-based CVD risk prediction systems to prevent the AI bias from in such systems; (vi) the performance evaluation metrics used for the three techniques were analyzed. It was found that the most commonly used PE parameter was accuracy. The cloud-based AI techniques comprising all the three classifications techniques are more likely to be the future for CVD risk prediction. In the future, advanced computer-aided diagnosis techniques can be applied based on image processing [210]. Edge devices with mobile and cloud-based AI infrastructure are now highly emerging in the medical industry as it provides remote facility and is a much faster, the most necessary feature in the COVID-19 era.

7.2. Benchmarking Table

Table 6 shows the benchmarking table with a comparison between eighteen review studies that focused on multiclass, multi-label, and ensemble techniques for CVD risk prediction. This table shows thirteen attributes (column C1 to column C13) for each of the eighteen studies [35,211–226] corresponding to the rows R1 to R18. These thirteen attributes presented were the Author (C1), year of the study (C2), name of the journal (C3), data size (C4), the study belongs to CVD or not (C5), the domain of the study (C6), machine learning (C7), classifier type (C8), cross-validation protocol (C9), the studies are multiclass (C10), multi-label study (C11), ensemble study (C12), and finally the summary of the study (C13). The data size for each study is shown in column C4, which is ranging from 8 to 86,155, whereas our study (row R18) has used 94 studies. Column C5 describes whether the study is of CVD type or not. Studies (rows R2, R3, R5, R9, R10, R11, R12, R16, and R17) along with our study (row R18) are in the field of CVD while the rest are not. Column C6 describes the different domains for the studies (rows R1, R4, R6, R7, R8, R13, R14, and R15) which does not belong to CVD. The domains are EEG, blood pressure, education, statistics, software, chronic fatigue, and sickle cells. The technical approach of the studies is shown in column C7, i.e., whether machine learning (ML) or not. Most of the studies including our proposed study are ML (rows R1, R3, R4, R6, R7, R8, R9, R10, R11, R12, R13, R14, R15, R16, R17, and R18). Column C8 indicates the classifier types for the studies ranging from SVM, NN, LDA, OVO (row R1), RF, SVM, DT, KNN, LR, GNB (row R3), SMOTE (row R4) [227], Adaboost, KNN, BPSO (row R6), XG-Boost (row R7), RF, NBC, KNN (row R8), K-Star (row R9), SVM, RF, CNN (row R10), KNN, RF, DT (row R11), LDA (row R13), MULAN (row R14), LDA, MDDM (row R15), Probabilistic (row R16), LogitBoost (row R17). The cross-validation protocols used are shown in column C9 which are K5 (rows R3, R4, R17), K7 (row R6), Open (row R7), K10 (rows R8, R11), and K* (row R9). The multiclass studies were (rows R1, R3, R6, R7, R9, R11, R12, R17) shown in column C10 along with our study (row R18). Column C11 shows multi-label studies (rows R8, R13, R14, R15, and R18) likewise column C12 shows the ensemble studies (rows R4, R6, R10, and R18). The last column C13 describes the keyword objectives of each study. The studies' objectives were classification and CVD risk prediction or stratifications.

Table 6. Benchmarking table for the multiclass, multi-label, and ensemble studies in CVD/non-CVD field.

SN	C1 Author	C2 Yr	C3 JOU	C4 DS	C5 CVD	C6 Domain	C7 ML	C8 CT	C9 CVP	C10 MC	C11 MLB	C12 Ensbl	C13 Summary
R1	Boernama et al. [211]	'21	IEEE	30	✗	EEG	✓	SVM, NN, LDA, OVO	✗	✓	✗	✗	EEG Classification
R2	Collins et al. [212]	'16	BMJ	122	✓	BP	✗	✗	✗	✗	✗	✗	CVD Meta-analysis
R3	Dissanayake et al. [213]	'21	Hindawi	CHDD	✓	✗	✓	RF, SVM, DT, KNN, LR, GNB	K5	✓	✗	✗	CVD risk
R4	Galar et al. [214]	'12	IEEE Tran.	Imb D	✗	✗	✓	SMOTE	K5	✗	✗	✓	Ensemble Classification
R5	Stewart et al. [215]	'17	JRSMCD	✗	✓	✗	✗	✗	✗	✗	✗	✗	CVD risk
R6	Mathew et al. [216]	'21	IEEE	6	✗	Edu	✓	Adaboost, KNN, BPSO	K7	✓	✗	✓	Teaching Quality
R7	Uike et al. [217]	'21	IEEE	8	✗	SC	✓	XG-Boost	Open	✓	✗	✗	SC Classification
R8	Wang et al. [218]	'14	Plos One	736	✗	CF	✓	RF, NBC, KNN	K10	✗	✓	✗	CF Classification
R9	Wiharto et al. [219]	'16	HIR	303	✓	✗	✓	K-Star	K*	✓	✗	✗	CHD Classification
R10	Boi et al. [220]	'18	CAR	126	✓	✗	✓	SVM, RF, CNN	✗	✗	✗	✓	OCT-based risk stratification
R11	Jamthikar et al. [35]	'20	CBM	208	✓	✗	✓	KNN, RF, DT	K10	✓	✗	✗	CVD risk
R12	Bianchini et al. [221]	'08	IEEE	10	✓	✗	✓	✗	✗	✓	✗	✗	Cardiovascular Risk Markers
R13	Liu et al. [222]	'12	IEEE	15	✗	Statistics	✓	LDA	✗	✗	✓	✗	Statistical Classification
R14	Charte et al. [223]	'20	IEEE	✗	✗	Software	✓	MULAN	✗	✗	✓	✗	Comparison
R15	Siblini et al. [224]	'15	IEEE	156	✗	DM	✓	LDA, MDDM	✗	✗	✓	✗	DM Reduction
R16	Indhumathi et al. [225]	'21	IEEE	30	✓	✗	✓	Probabilistic	✗	✗	✗	✗	CVD Management
R17	Kolli et al. [226]	'19	IEEE	86,155	✓	✗	✓	LogitBoost	K5	✓	✗	✓	Coronary Artery Calcification
R18	Proposed Study	'22	✗	265	✓	✗	✓	✗	✗	✓	✓	✓	CVD risk

DS: Data size; ML: Machine learning; CVP: Cross-validation protocol; MC: Multiclass; MLB: Multi-label; GNB: Gaussian I Bayes; HD: Heart disease; CHDD: Cleveland heart disease datasets; Ensbl: Ensemble; IEEE Tran: IEEE Transaction; JRSMCD: Journal of the Royal Society of Medicine Cardiovascular disease; CT: Classifier type; ImbD: Imbalance data; JOU: Journal; SC: Sickle cells; CF: Chronic fatigue.

7.3. A Special Note on Non-Linear CVD Risk Stratification

The conventional classification CVD risk assessment systems assume the linear relationship between the covariates and the gold standard. The linear systems typically use the covariates like OBBM and LBBM or ECG signals [228–230]. With the additions of CUSIP and MedUSE, the requirement becomes more stringent on CVD calculators. In today's times, it was observed that COVID-19 can play the role of a new covariate or risk factor due to its relationship with CVD [231,232]. The risk of CVD gets accelerated in the individual with COVID-19 [233,234]. This inclusion can result in a more non-linear classification paradigm for CVD risk prediction [235]. This can improve the reliability and the accuracy of the prediction results [236]. The AI/ML approaches help in understanding the non-linear

relationship between the covariates and the ground truth. Hence there is a need for the development of non-linear classifiers in the ML/DL domain. It includes non-linear SVM classifiers [237], PCA, XGBoost [235], RF [233], generalized discriminant analysis (GDA), ELM, LDA [238]. Different non-linear methods which are applied in the CVD field are Poincare plot (PP), approximate entropy (ApEn) [235], quasi period density-prototype distance (QPD-PD) [239], fuzzy entropy [238], recurrence period density prototype distance (RPD-PD) [237], non-linear ensemble classifiers [233]. These are all out of the scope of the current study. The other application of non-linear classifiers are in the field of stroke [240] and sleep apnea [241]. The non-linearity can also be handled by using the DL approaches along with multiclass, multi-label, and ensemble-based techniques for CVD risk prediction in the future.

7.4. A Special Note on Time-to-Event for Cardiovascular Risk Prediction

This is one of the greatest assets of the machine learning system. The most important ingredient for accomplishing this solution is to ensure that we have a follow-up gold standard for the clinical data. This means one must have the gold standard (events) for the times such as 1st-year, 3rd-year, 5th-year, and 10th-year. Further, the risk factors (so-called covariates or variables) must be available for the development of the training model. Given the two pairs (covariates and the gold standard-even for that time), one can develop the machine learning model for that time-zone (1st-year, 3rd-year, 5th-year, and 10th-year). Should you intended to predict for 1st-year, 3rd-year, 5th-year, and 10th-years, it requires four kinds of machine learning models. Each time-event has to have its own machine learning model. The atherosclerosis disease which has transformed over different years and leads to the event needs to be used for the development of the training model. The only challenge with this setup is the length of time it takes to collect the event data. It is both expensive and tedious since we have to follow the patients over the 10-year period. Recently, Kakadiaris et al. [62] perused this strategy using the machine learning paradigm. The ML paradigm has the same fundamental concept of training and testing as shown in Figure 4. The left half is the training model where the gold standard will change as per the time-zone (1st-year, 3rd-year, 5th-year, and 10th-years), while the prediction will be applied for the patient for the corresponding time-zones (1st-year, 3rd-year, 5th-year, and 10th-years). It is painful to wait to accomplish this validation, since it is costly, and a large cohort is needed.

To overcome such a scenario, another way to predict the CVD risk is using the surrogate marker of carotid artery disease. Since the formation of the atherosclerotic disease in coronary artery has the same genetic make-up as the carotid artery disease, the surrogate artery can be used for the prediction of CVD or the coronary artery disease risk. Further, note that over time (1st-year, 3rd-year, 5th-year, and 10th-years), the plaque formation changes and so does the image phenotypes such as intima-media thickness, plaque burden, or plaque area/volume. Thus, one can compute the time-dependent image phenotypes which uses the ingredients which make the atherosclerotic disease. This includes rate of change of cIMT over time (age), obesity index over time (age), cholesterol change over (age), one can use this paradigm to predict the plaque burden in carotid artery-based age. This is sometimes called as vascular age of the patient. This has been shown by Khanna et al. [34]. Later, this was commercialized as AtheroEdge™ 2.0 (AtheroPoint™, Roseville, CA, USA) [36]. The CVD risk can be computed based on the intensity of the risk factors. This is called a non-ML method (also known as the statistical solution for the prediction of the 10th-year CVD risk).

7.5. A Special Note on the Advantages of Machine Learning-Based Cardiovascular Risk Stratification

Machine learning paradigm for CVD risk prediction has provided us with a way to obtain more accurate, early, and fast results. The ML systems offer following advantages against the previously published studies: (i) it handles the non-linear nature between the

covariates and ground truths (GT) [31]; (ii) ability to predict the CVD risk in granular classes, such as six different risk classes (no-risk, low-risk, mild-risk, moderate-risk, high-risk, and very-high-risk) [34,35]; (iii) ability to augment the training data using popular augmentation paradigms such as adaptive synthetic (ADASYN) and synthetic minority over-sampling technique SMOTE [227]; (iv) incorporate the cohort's knowledge during training and predicting the CVD risk; (v) flexibility of amalgamating of different types of covariates such as OBBM, LBBM, CUSIP, and MedUSE during the design of the model training; (vi) ability to interface with different types of classification techniques like multiclass, multi-label and ensemble for improving the overall performance of the system; and (vii) ability to enhance the risk factor (or covariates) such as genetic and comorbidities such as cancer. Thus, all the above-mentioned factors puts ML-based system a very strong paradigm for CVD risk stratification, unlike the conventional statistical models.

7.6. A Special Note on Deep Learning-Based Cardiovascular Risk Stratification

The Deep learning (DL) paradigm has started to emerge in the field of CVD risk prediction. The DL approach can be applied for both (a) the office-based [242,243] and (b) stress-based test settings [244–248]. DL approaches have been applied for CVD risk stratification using multiclass [249], multi-label [250], and ensemble-based paradigms [116]. Even though there are evolving CVD risk stratification techniques in the DL framework, this review does not venture deep since it is not the main focus of this review. As a result, we have not analyzed publications related to the DL paradigm. Note that, the main advantage of DL techniques is (i) automated feature selection process from the input covariates (such as OBBM, LBBM, CUSIP, and ECG signals phenotype) and (ii) prediction of more accurate and reliable results due to a large number of layers in DL network. Advanced stochastic imaging methods can be applied [251] to improve the loss function during the training paradigm. This evolving DL paradigm will flourish more in the very near future in office-based imaging and stress-based test settings.

7.7. The Future of Cardiovascular Disease Risk Stratification

The CVD risk estimation at an early stage is very much important to reduce the mortality rate due to CVD [252,253]. As it was observed that not only ML but extreme machine learning (ELM) can also be applied and further developed for CVD risk stratification [254]. Moreover, COVID-19 accelerates the atherosclerosis condition due to which fast detection of CVD in COVID-19 patients is needed [255,256]. The above circumstances are leading to an evaluation in the CVD risk stratification techniques. In the near future, cloud-based AI modalities will be very much in use for CVD risk detection. It also promotes the remote and fast prediction of the risk of CVD. It also helps in reducing prediction errors. Other non-invasive imaging techniques like carotid, femoral, arterial imaging can be used as an indirect measure of plaque build-up in these arteries. Deep learning technologies will evolve in the field of CVD risk estimation [257]. This will also include pruning of weights using evolutionary techniques such as genetic algorithms in the Deep Learning framework [147]. Devices equipped with cutting edge technologies like mobile-based AI, cloud-based AI, multiclass, multi-label, and ensemble-based systems for CVD risk prediction will be emerging in the medical imaging industry market.

8. Conclusions

This was the first review study of its kind that presented three different kinds of AI-based CVD risk stratification, namely multiclass, multi-label, and ensemble, where multiclass was most popular and multi-label was least, which happened to be our first key contribution. The second contribution was exhaustive analysis by selecting the best 265 studies using the PRISMA model for understanding the three kinds of machine learning-based systems for prediction of the CVD risk. This was based on our hypothesis that there exists a biological link between atherosclerotic disease formation and the CVD risk. The third contribution was the identification of the top four covariates, namely OBBM,

LBBM, CUSIP, and MedUSE for designing the training model using a machine learning framework. The fourth contribution was on the choice of the gold standard for an unbiased AI system design for CVD risk prediction, which leads to a robust and reliable CVD prediction system. The fifth finding and contribution required that the ML system undergo clinical and scientific validation for reliability, stability, and robustness of the system design. Lastly, we observed that with the advancement of telecommunication systems, mobile and cloud-based strategies are speedily penetrating the CVD risk stratification system designs. Low-powered edge devices like Raspberry Pi and Jetson Nano are like to be adopted in the future.

Author Contributions: Conceptualization: J.S.S., M.K.K., and N.N.K.; Data Curation: L.S., G.F., Z.R. and M.B.; Formal Analysis: J.S.S., M.K.K., S.P. and M.B.; Investigation: J.S.S., M.B., S.P., A.D.P., P.P.S., G.D.K., N.N.K., Z.R., A.M.S., S.S., G.F., J.R.L., A.M.J., M.K.K., K.I.P. and L.S.; Methodology: J.S.S., and M.B.; Project Administration: J.S.S.; Computing Resources: S.P.; Software Design and Usage: M.B.; Software Verification: J.S.S., S.S., K.I.P. and L.S.; Supervision: J.S.S., S.P., S.S. and M.K.K.; Scientific Validation: J.S.S. and M.B.; Clinical Validation and Discussions: J.S.S., M.K.K. and L.S.; Visualization: M.B.; Writing—Original Draft: J.S.S. and M.B.; Writing—Review & Editing: J.S.S., M.B., S.P., A.D.P., P.P.S., G.D.K., N.N.K., Z.R., A.M.S., S.S., G.F., J.R.L., A.M.J., M.K.K., K.I.P. and L.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Pseudo-Code for Multiclass Classification

Appendix A.1. Typical Online System for CVD Risk Stratification for Multiclass

This system shows the amalgamation of online covariates, which are then transformed by the ML-based training model using multiclass-based models. The output yields the multiclass risk marked in color (low, mild, moderate, and high risk).

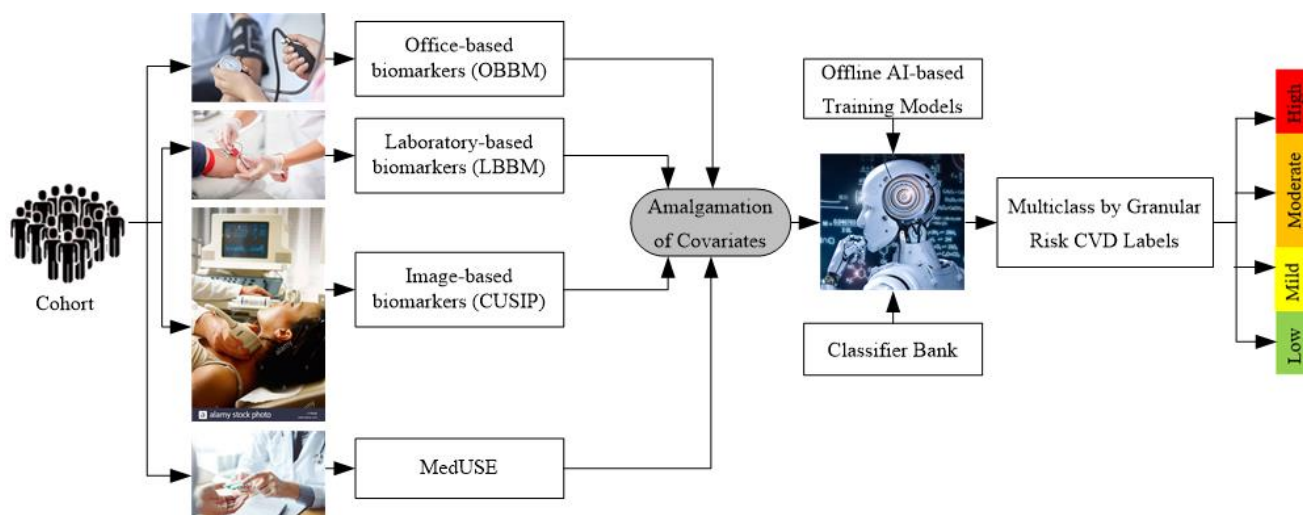


Figure A1. Typical online system for multiclass CVD risk stratification.

Appendix A.2. Pseudo-Code for Multiclass

The pseudo-code describes the process used by the multiclass algorithm for CVD risk stratification into granular risk classes. It uses the “for” loop for training and prediction of

each fold of data, which were divided into K folds. The training model is applied to the test data and the PE was predicted and stored in form of accuracy (ACC), ROC, sensitivity (Sen), specificity (Spec), F1-score, the area-under-the-curve (AUC), and precision.

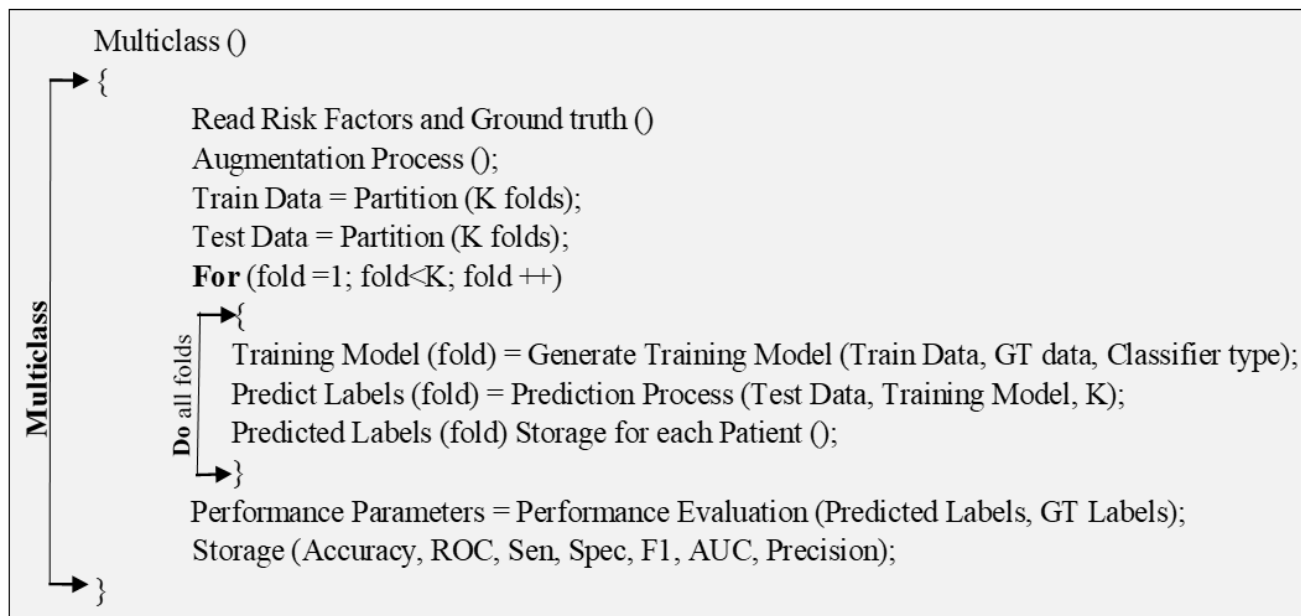


Figure A2. Pseudo-code for multiclass technique.

Appendix B. Pseudo-Code for Multi-Label Classification

Appendix B.1. Problem Transformation Methods for Multi-Label Prediction

The problem transformation method (PTM) makes the multi-label classification problem to one or more single label classification tasks. Basically, four PTM, namely BR, CC, LP, and RakEL were used as discussed below:

Binary Relevance: In the BR technique, the problems get divided into one or more single-label classification problems. The single-label classification resembles the binary prediction. An example can be described as, say M is a set of “ q ” labels with $M = \{m_1, m_2, \dots, m_q\}$, the BR technique makes “ q ” single-label binary classifiers for each label. The multi-label training sets get converted to binary datasets (“ q ”), and $E_{ij} = 1 \dots q$, where E_{ij} has all samples of the original dataset but with single positive or negative values. The dataset gets divided into “ q ” single label datasets with classifier C and the next classifier set is obtained as $C_j (E)$, $j = 1 \dots q$ by the training set E_{ij} . The label dependency was not considered by the BR classification algorithm. Thus, it shows less complexity in the computation as compared with other multi-label techniques. The process is shown in the following Figure A3 [258]. As shown in Figure A3 four examples were considered as multi-label dataset and label set M with four labels (m_1, m_2, m_3 , and m_4) which is split as four single labels that are independent.

Classifier Chain: This algorithm also works in single-label classification. This technique takes a class of classifiers where the very initial classifier is trained with the dataset, which acts as the input, following that each classifier gets trained with the whole feature space. The feature set has an original dataset with the label set used in the earlier base classifier that is in the chain. Each base classifier uses the earlier label information for training and testing models. Thus, a correlation exists in the CC algorithm. Figure A4 describes the functioning of CC [259].

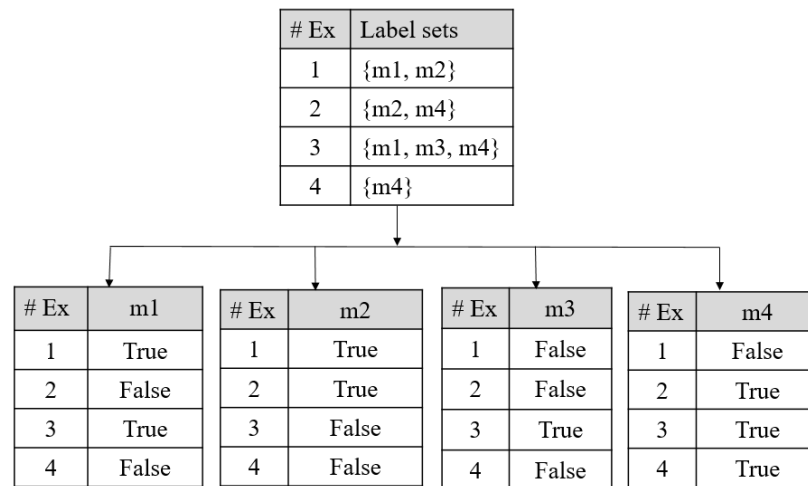


Figure A3. BR classifier, #Ex: Example.

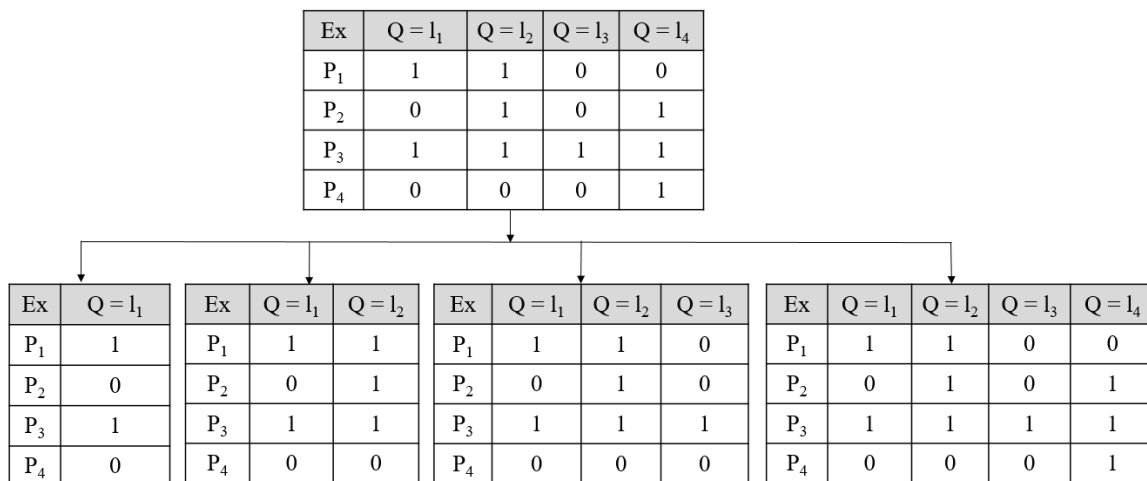


Figure A4. Classifier chain functioning; P₁–P₄ are data points under; Ex; Ex: Example.

Label Powerset: It also converts the prediction situation to a single-label multiclass prediction technique. In this technique, all possible individual group of labels is given special or unique class. Such as if three types of labels are there, then eight different types of combinations can come into the picture. LP technique has eight types of labels that get trained for prediction. This technique deals with a large number of classes that are related to small instances, and also consideration of correlation is done. The transformation was shown in Figure A5 [260]. In Figure A5 the 1st table shows the original datasets, and the 2nd table is showing the transformed datasets.

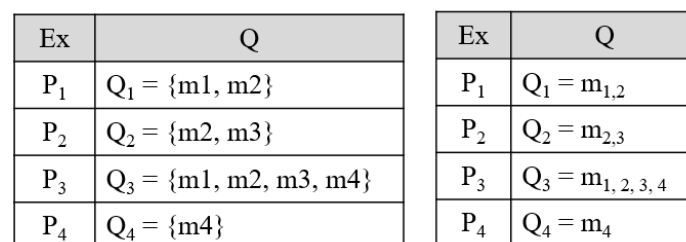


Figure A5. Label powerset transformation, Ex: Example.

Random k-label set: It is a type of combination technique used for multi-label prediction. Every combination method gets trained on a small size of the randomly selected subset of

labels by a single-label-based classifier. This process is described as if L labels in the dataset (E), the RAKEL classifier turns this data to all the possible k-label sets (L_k). Each label set is then trained for prediction. Finally, the prediction is made into positive (1) and negative (0) values in accordance with the threshold (0.5). The further implementation can be seen in [261].

Appendix B.2. Algorithm Adaptation Methods for Multi-Label Prediction

Multi-label KNN: This algorithm is basically an implementation of the KNN algorithm in multi-label datasets. The neighbors are selected from unseen training sets. Next, the labelset are found for the instance which are unseen in nature by utilizing the maximum of posteriori (MAP) principle. The full algorithm can be seen in [262].

Multi-label ARAM: It is associated with the neural network model based on resonance theory. The advantage of this algorithm is its fast learning ability. The detailed algorithm can be seen in [263].

Appendix B.3. Pseudo-Code for Multi-Label Classification Technique

Multi-label pseudo-code describes the multi-label algorithm where more than one multi-label endpoint was considered. For each multi-label endpoint, the risk class was defined. In this pseudo-code, two “for” loops are used one for multi-label and the next for multiclass prediction. Finally, the PE was determined as accuracy, sensitivity (Sen), specificity (Spec), area-under-the-curve (AUC), sample-based, and label-based metrics.

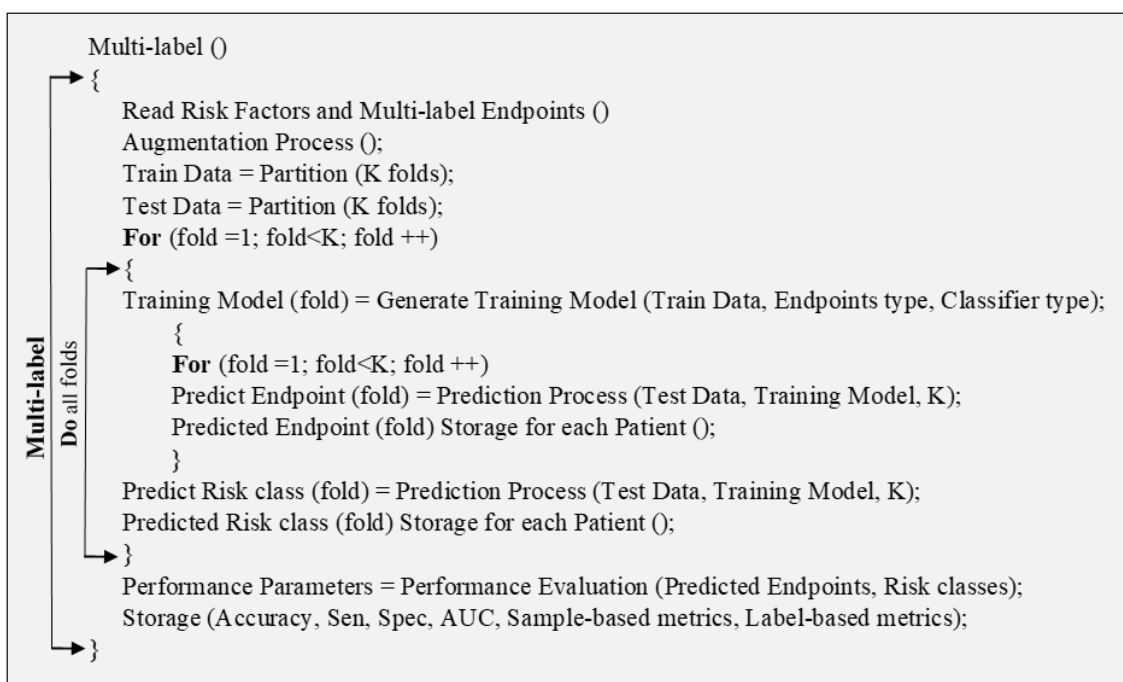


Figure A6. Pseudo-code for multi-label technique.

Appendix C. Pseudo-Code for Ensemble Classification

Pseudo-Code for Ensemble-Based Technique

Ensemble-based-CVD risk prediction uses combinations of multiple classifiers. The pseudo-code shows that the data are divided into testing and training with K folds. The prediction was done using each type of classifier for multiclass and multi-label prediction. Then each type of classifier is combined into an ensemble classifier and the final prediction was made.

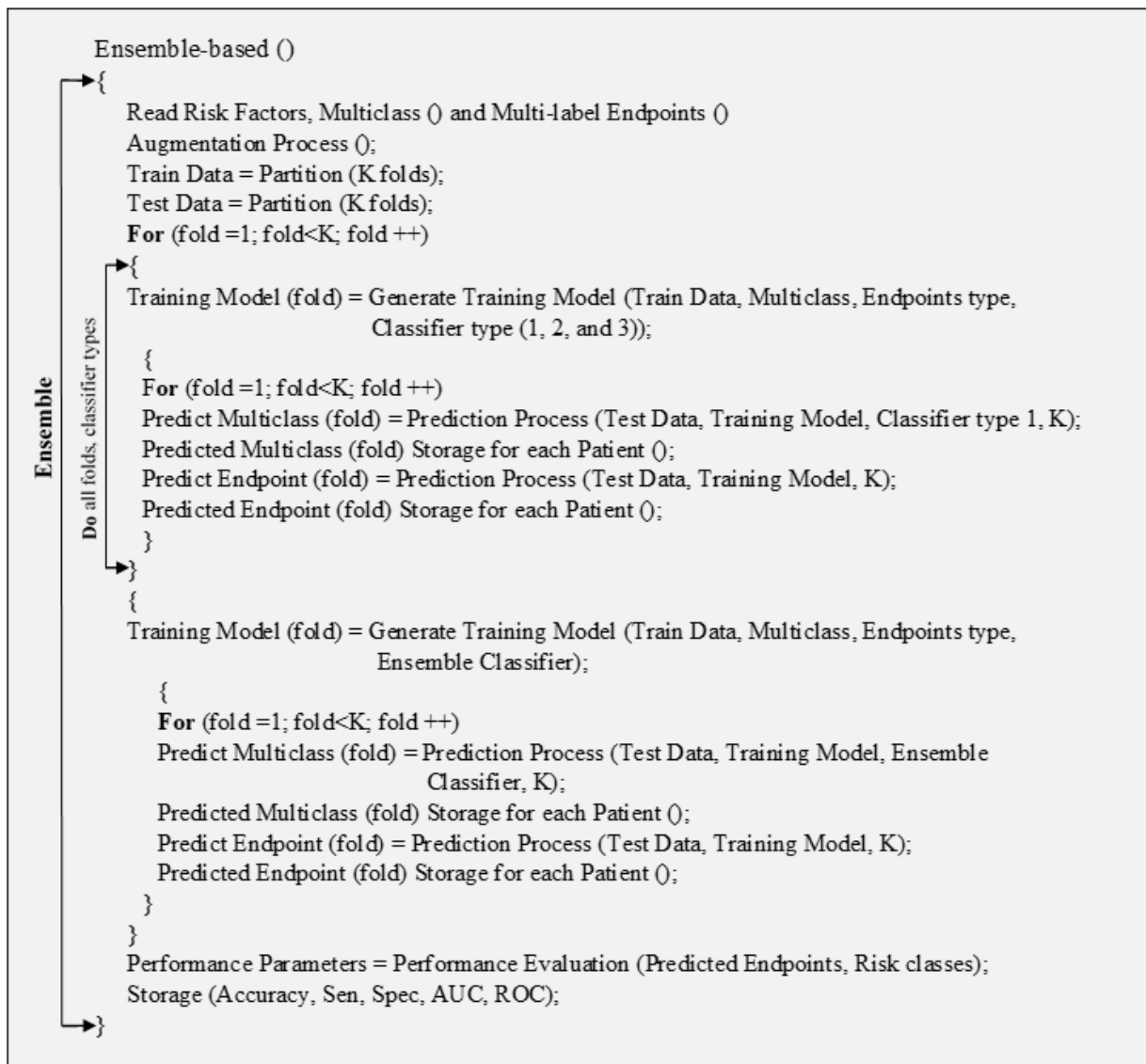


Figure A7. Pseudo-code for ensemble-based technique.

Appendix D. Comparison between 3 Paradigms

Comparison of ML-Based Multiclass, Multi-Label, and Ensemble CVD Classification

Table A1. Comparison of ML-based multiclass, multi-label, and ensemble CVD classification.

SN	Attributes	Multiclass		Multi-Label		Ensemble
-	-	Characteristics		Characteristics		Characteristics
	Total Studies	14	[69–82]	8	[83–90]	32 [80,91–121]
1	Data Size	212–66,363	[69–82]	300–46,520	[83–90]	459–823,627 [80,91–121]
2	Risk Factors	Low	[69–82]	Large	[83–90]	Moderate [80,91–121]
3	Family History	Frequent Considered	[69,71,76,77,80,82]	Seldom Considered	[83,84,90]	Considered Intermittently [80,91,96,97,99,100,102,105,106,110–112,114–120]
4	BMI	Less considered	[72,74–76,80]	Considered Moderately	[84–86]	Highly considered [46–52,80,91,93–97,99,100,102,106,107,112]

Table A1. Cont.

SN	Attributes	Multiclass	Multi-Label	Ensemble
-	-	Characteristics	Characteristics	Characteristics
5	Ethnicity	Less Considered [72,74–76,80]	Considered Moderately [84–86]	Highly Considered
6	Type of data	OBBM and LBBM [69–82]	OBBM, LBBM and Image [83–90]	OBBM and LBBM [80,91–121]
7	Hypertension	Low Usage [72,74–76,80]	High Usage [83–90]	Moderate Usage [46–52,80,91,93–97,99,100,102,106,107,112]
8	Smoking	Low Usage [72,74–76,80]	High Usage [83–90]	Moderate Usage [80,91,96,97,99,100,102,105,106,110–112,114–120]
9	Multicenter	Low Usage [72,74–76,80]	High Usage [83–90]	Moderate Usage [80,91,96,97,99,100,102,105,106,110–112,114–120]
10	MRI	Considered Moderately [71,80]	Considered Moderately [83,89]	Less Considered [80]
11	ECG	Partial Considered [72,74,75,78,79,81,82]	Strongly Considered [83,86,87,89]	Not Considered
12	CUSIP	Moderate Usage	Moderate Usage	Low Usage
13	# GT	Only 1 [69–82]	Very high (10-4) [83–90]	Average (1,2) [80,91–121]
14	# Algorithm	✗	✓ [83–90]	✗
15	Type of Algorithm	✗	-	✗
16	# Classifiers	Ranging from 1–4 [69–82]	Ranging from 1–9 [83–90]	Ranging from 1–10 [80,91–121]
SN	Attributes	Multiclass	Multi-label	Ensemble
-	-	Characteristics	Characteristics	Characteristics
17	Classifier Type	SVM, RF, CNN DT, k-NN Agatston classifier, Elastic Net, NN, NB, XGBoost SVM, ELM, OAO, OAA, DDAG, ECOC [69–82]	RF, SVM, DT, KNN, LDA, LR, XGBoost, AdaBoost, GBA, Basic RNN, GRU RNN CNN, AAM [83–90]	kNN, GaussNB, LDA, QDA, RF, MLP, CNN, LSTM, GRU, BiLSTM, BiGRU Bagging, XGBoost, Adaboost, DNN, NB, NN, RS, GAMs, Elastic Net, GBMs, DT, CART, MARS, Logistic, EB, SMO, Boosting, MLDS, AVEn, MVEn, WAVEn, HTSA [80,91–121]
18	# Classes	✓ [69–82]	✗	✗
19	Hyperparameters Used	✓ [79]	✓ [83,84,90]	✓ [92,98–100]
20	Protocol	K-10 [64–82]	K-10, K, K-5 [83–90]	K-10, k, K-5 [80,91–121]
21	# PE parameters	Ranging from 1–5 [69–82]	Ranging from 1–8 [83–90]	Ranging from 1–8 [80,91–121]
22	Precision	✓ [72,73,77,81,82]	✗	✓ [80,91–121]
23	PPV	✗	✓ [84,86]	✓ [80,91–121]
24	NPV	✗	✓ [84,86]	✓ [80,91–121]
25	FPR	✗	✓ [84,90]	✓ [80,91–121]
26	FNR	✗	✓ [84]	✓ [80,91–121]
27	Hamming Loss	✗	✓ [87]	✗
28	C-index	✗	✓ [83]	✗

Table A1. *Cont.*

SN	Attributes	Multiclass	Multi-Label	Ensemble
-	-	Characteristics	Characteristics	Characteristics
29	Statistical Analysis	✗	✓ [83–90]	✓ [80,91–121]
30	Power Analysis	✗	✓ [83,84]	✗
31	Hazard Analysis	✗	✓ [83]	✗
32	Survival Test	✗	✓ [83]	✗

SN: Serial number; SVM: Support vector machine; RF: Random forest; CNN: Convolutional neural network; DT: Decision tree, k-NN: k-Nearest neighbor; NN: Neural network; ELM: Extreme learning machine; OAO: One against one; OAA: One against all; DDAG: Decision direct acyclic graph; EOCC: Exhaustive output error correction code; LDA: Linear discriminant analysis; RNN: Recurrent neural networks; GRU: Gated recurrent unit; AAM: Algorithm adaptation methods; MARS: Multivariate adaptive regression splines; GAMS: Generalized additive models; PLR: Penalized logistic regression; GBM: Gradient boosted machines; MLP: Multilayer perceptron; CART: Classification and regression trees; SMO: Sequential minimal optimization; DNN: Deep neural network; NB: Naive Bayes; LSTM: Long short term memory network; EB: Ensemble boosting; MLDS: Multi-layer defense system; PPV: Positive predictive value; NPV: Negative predictive value; FPR: False positive rate; FNR: False negative rate; #GT: Number of ground truth.

Appendix E. Performance Evaluation Metrics

Performance Evaluation Metrics Descriptions

The PE for the multiclass and ensemble basically have accuracy (ACC), sensitivity (Sen), specificity (Spec), AUC, F1-Score which were calculated using values of true positives (TPs), false positives (FPs), false negatives (FNs), and true negatives (TNs). The formulae can be referred from Table A2. The performance evaluation for multi-label-based CVD is different as compared to multiclass and ensemble. They are label-based, instance-based performance evaluations.

In the label-based techniques, the PE parameters are checked for each label by the values of TPs, FPs, FNs, and TNs. All the labels have their own values. S, these are calculated by averaging methods (i) macro-averaging and (ii) micro-averaging [181]. The performance metrics say β is calculated by the values of TPs, FPs, FNs, and TNs, the macro-averaging techniques, macro-averaging (β_{macro}) for all labels (L) is given by averaging β for each label “ p ”, as shown in Equation (A1).

$$\beta_{macro} = \frac{1}{L} \left(\sum_{p=1}^L \beta (TP_p, FP_p, TN_p, FN_p) \right) \tag{A1}$$

In the same manner, for the micro-averaging techniques, the PE metrics are computed for each individual label and finally obtaining the micro-average (β_{micro}) by using the Equation (A2).

$$\beta_{micro} = \beta \left(\sum_{p=1}^L TP_p, \sum_{p=1}^L FP_p, \sum_{p=1}^L TN_p, \sum_{p=1}^L FN_p \right) \tag{A2}$$

For instance-based performance evaluation, the parameters are calculated for individual instances, then the average value is computed and final the performance metric is performed. The final metric has a hamming loss, precision, recall, F1-score, Jaccard similarity coefficient score, and accuracy.

The multi-label dataset is supposed to be $|E|$ with multi-label examples (p_i, Q_i), $i = 1 \dots |E|$, and $Q_i \subseteq L$, L is a set of all multiple labels. C is a multi-label classifier and $M_i = C(p_i)$ be the set of labels predicted by C . $|E|$ indicates the features of the set E , while $|Q_i \cap M_i|$ indicates the feature of the intersection of true labels and the predicted labels. $|M_i|$ indicates the features of predicted labels, and $|Q_i|$ indicates the features of the true labels.

Hamming loss shows the number of times when the label pair is misclassified. The lower value of Humming loss presents the better performance of the multi-label classifier. Jaccard score presents the ratio of the size of the intersection between predicted and the ground truth labels. Precision is the proportion of correct predictions out of all predictions. Likewise, recall is the ratio of correct predicted labels to the actual labels. F1-score is the combination of precision and recall Table A3.

Table A2. Performance evaluation metrics used in CVD risk assessment.

SN	Label-Based Performance Metrics	Mathematical Expression
1	Sensitivity (Sen), %	$Sen = \left[\frac{TP}{(TP+FN)} \right] \times 100$
2	Specificity (Spec), %	$Spec = \left[\frac{TN}{(TN+FP)} \right] \times 100$
3	Positive Predictive Rate (PPR), %	$PPR = \left[\frac{TP}{(TP+FP)} \right] \times 100$
4	Negative Predictive Rate (NPR), %	$NPR = \left[\frac{TN}{(TN+FN)} \right] \times 100$
5	False Predictive Value (FPV), %	$FPV = \left[\frac{FP}{(FP+TN)} \right] \times 100$
6	False Negative Value (FNV), %	$FNV = \left[\frac{FN}{(FN+TP)} \right] \times 100$
7	False Discovery Value, %	$FDV = \left[\frac{FP}{(TP+FP)} \right] \times 100$
8	F1-Score, %	$F1 = \left[\frac{2TP}{(2TP+FP+FN)} \right] \times 100$
9	Accuracy (ACC), %	$ACC = \left[\frac{TP+TN}{(TP+FP+TN+FN)} \right] \times 100$

Table A3. Performance evaluation metrics used in CVD risk assessment.

SN	Sample-Based Performance Metrics	Mathematical Expression
1	Hamming Loss, HL	$HL (C, E) = \frac{1}{ E } \sum_{i=1}^{ E } \frac{ Q_i \Delta M_i }{ L }$
2	Jaccard Score, JS	$JS = \frac{1}{ E } \sum_{i=1}^{ E } \frac{ Q_i \cap M_i }{ Q_i \cup M_i }$
3	Precision, Pe	$Pe (C, E) = \frac{1}{ E } \sum_{i=1}^{ E } \frac{ Q_i \cap M_i }{ M_i }$
4	Recall, Re	$Re (C, E) = \frac{1}{ E } \sum_{i=1}^{ E } \frac{ Q_i \cap M_i }{ Q_i }$
5	F1-score, F1	$F1 (C, D) = \frac{1}{ E } \sum_{i=1}^{ E } \frac{2 Q_i \cap M_i }{ Q_i + M_i }$
6	Subset Accuracy, Acc _{Subset}	$Acc_{Subset} = \frac{1}{ E } \sum_{i=1}^{ E } I(M_i = Q_i)$

Appendix F. Power Analysis

Power Analysis for Multi-Label and Ensemble-Based CVD Risk Stratification

Power analysis can be done for multi-label and ensemble-based CVD systems. Its objective was to state the smallest data or sample size (s) needed to perform the multi-label, ensemble-based CVD risk classification. The parameters which are required for calculating power analysis are confidence interval, a margin error (e) as ±5%, and a sample proportion (q̂), the z-score (z*) (taken standard z-table). Therefore, the formula used is shown in Equation (A3) [264,265].

$$s = \left(\frac{z^*}{e}\right)^2 \times \hat{q}(1 - \hat{q}) \tag{A3}$$

Appendix G. CVD Risk Assessment through Mobile, E-Health, and Cloud Techniques
Characteristic of Mobile and Cloud-Based CVD Systems

Table A4. Characteristics of mobile and cloud-based CVD systems.

C0	C1	C2	C3	C4	C5	C6	C7	C8			
SN	Authors/Citations	ST	Year	Journal	DS	Diseases	FDA	SV	Comparator		
1	Buss et al. [197]	SR	2020	JMIR	7 ED	CVD, DIA	✗	✗	No (i.e., standard care), await list control, intervention		
2	Villarreal et al. [198]	SR	2020	AIF	44	CVD	✗	✗	CVD, No CVD		
3	Xiao et al. [199]	R	2017	TM	151	CVD	✗	✗	CVD, No CVD		
4	Saba et al. [200]	R	2018	IHJ	100	CVD	✗	✓	CVD, No CVD		
5	Lillo-Castellano et al. [208]	R	2015	JBHI	6848	CVD	✓	✓	CVD, No CVD		
6	Huda et al. [201]	R	2020	TENSYMP	BIHAD	CVD	✗	✓	Normal ECG, Abnormal ECG		
7	Sakellarios et al. [209]	R	2018	EMBC	236	CAD	✗	✓	No CAD, OCAD, Non-OCAD		
8	Singh et al. [202]	R	2019	IEEEc	2	CVDa	✗	✓	Arrhythmia, CVD		
9	Spanakis et al. [203]	R	2020	EMBC	✗	CHF	✗	✓	CHF, No CHF		
10	Paredes et al. [204]	R	2018	BIBM	1600	MI, CVD	✗	✓	Acute MI, No MI		
11	Freyer et al. [205]	R	2021	AJH	✗	AF	✗	✓	AF, No AF		
12	Giansanti et al. [206]	S	2021	mHealth	✗	CVD	✗	✗	Use of AI, non-use of AI		
13	Park et al. [207]	R	2014	IEEEa	✗	Arrhythmia	✗	✗	Arrhythmia, CVD		
SN	Authors/Citations	Non ML/ML	Cloud	Mob	Sea	DE	Analysis	# O	OT	# C	Classifier
1	Buss et al. [197]	Non-ML	✗	✓	✓	✓	✓	2	Dia, CVD	3	✗
2	Villarreal et al. [198]	Non-ML	✓	✓	✓	✓	✓	1	CVD	2	✗
3	Xiao et al. [199]	Non-ML	✗	✓	✓	✓	✓	1	CVD	2	✗
4	Saba et al. [200]	Non-ML	✓	✓	✓	✓	✓	1	CVD	2	✗
5	Lillo-Castellano et al. [208]	ML	✓	✗	✓	✓	✓	1	CVD	2	k-NN
6	Huda et al. [201]	ML, DL	✓	✓	✓	✓	✓	1	Arrhythmia	2	SVM, CNN
7	Sakellarios et al. [209]	ML	✓	✗	✓	✓	✓	1	CVD	3	SVM

Table A4. Cont.

	C0	C1	C2	C3	C4	C5	C6	C7	C8			
SN	Authors/Citations	ST	Year	Journal	DS	Diseases	FDA	SV	Comparator			
8	Singh et al. [202]	DL	✓	✓	✓	✓	✓	1	CVDa	2	CNN	
9	Spanakis et al. [203]	IoT	✓	✓	✓	✓	✓	1	CHF	2	✗	
10	Paredes et al. [204]	CI	✗	✓	✓	✓	✓	2	CVD, MI	2	Bayesian	
11	Freyer et al. [205]	Non-ML	✓	✓	✓	✓	✓	1	AF	2	✗	
12	Giansanti et al. [206]	AI	✓	✓	✓	✓	✓	1	CVD	2	✗	
13	Park et al. [207]	ML	✗	✓	✓	✓	✓	1	Arrhythmia	2	DT, RF	
	C0	C19	C20	C21	C22	C23	C24	C25	C26	C27	C28	C29
SN	Authors/Citations	CV	Protocol	# PE	SEN	SPEC	Acc	Pre	F1 S	PV	SS	ROC
1	Buss et al. [197]	✗	✗	0	✗	✗	✗	✗	✗	✗	✗	✗
2	Villarreal et al. [198]	✗	✗	0	✗	✗	✗	✗	✗	✗	✗	✗
3	Xiao et al. [199]	✗	✗	1	✗	✗	✗	✗	✗	✗	2.87	✗
4	Saba et al. [200]	✓	✗	1	✗	✗	✗	✗	✗	✗	✗	1
5	Lillo-Castellano et al. [208]	✓	K	1	✗	✗	90	✗	✗	✗	✗	✗
6	Huda et al. [201]	✓	✗	1	✗	✗	96	✗	✗	✗	✗	✗
7	Sakellarios et al. [209]	✓	✗	3	44	98.7	85.1	✗	✗	✗	✗	✗
8	Singh et al. [202]	✗	✗	1	✗	✗	97	✗	✗	✗	✗	✗
9	Spanakis et al. [203]	✗	✗	1	✗	✗	1	✗	✗	✗	✗	✗
10	Paredes et al. [204]	✓	✗	0	✗	✗	✗	✗	✗	✗	✗	✗
11	Freyer et al. [205]	✓	✗	1	✗	✗	1	✗	✗	✗	✗	✗
12	Giansanti et al. [206]	✓	✗	0	✗	✗	✗	✗	✗	✗	✗	✗
13	Park et al. [207]	✓	✗	3	1	1	1	✗	✗	✗	✗	✗

SN: Serial number; CV: Cross validation; SEN: Sensitivity; SPEC: Specificity; Acc: Accuracy; Pre: Precision; F1 S: F1 Score; PV: *p*-value; SS: Silberg score. DE: Data extraction; OT: Outcome types; C: Comparators; O: Outcomes; CI: Computational intelligence; CHF: Congestive heart failure; CVDa: CVD Auscultation; Dia: Diabetes; MI: Myocardial infarction; Mob: Mobile; Sea: Scientific validation; # O: Number of outcomes; # C: Number of classes. DS: Data size; BIHAD: MIT-BIH Arrhythmia Database; IEEEc: IEEE connect; AF: Atrial fibrillation; R: Research; SR: Systemic review; ST: Study type; IHJ: Indian Heart Journal; AIF: AI Foundation; TM: Telemedicine; IEEEa: IEEE-ACAINA; SV: Scientific validation; OCAD: Obstructive CAD; NonOCAD: Non-obstructive CAD.

Appendix H. Miscellaneous Figures

Appendix H.1. Anatomical Link between the Carotid Artery and Aortic Arch and Typical Neural Network

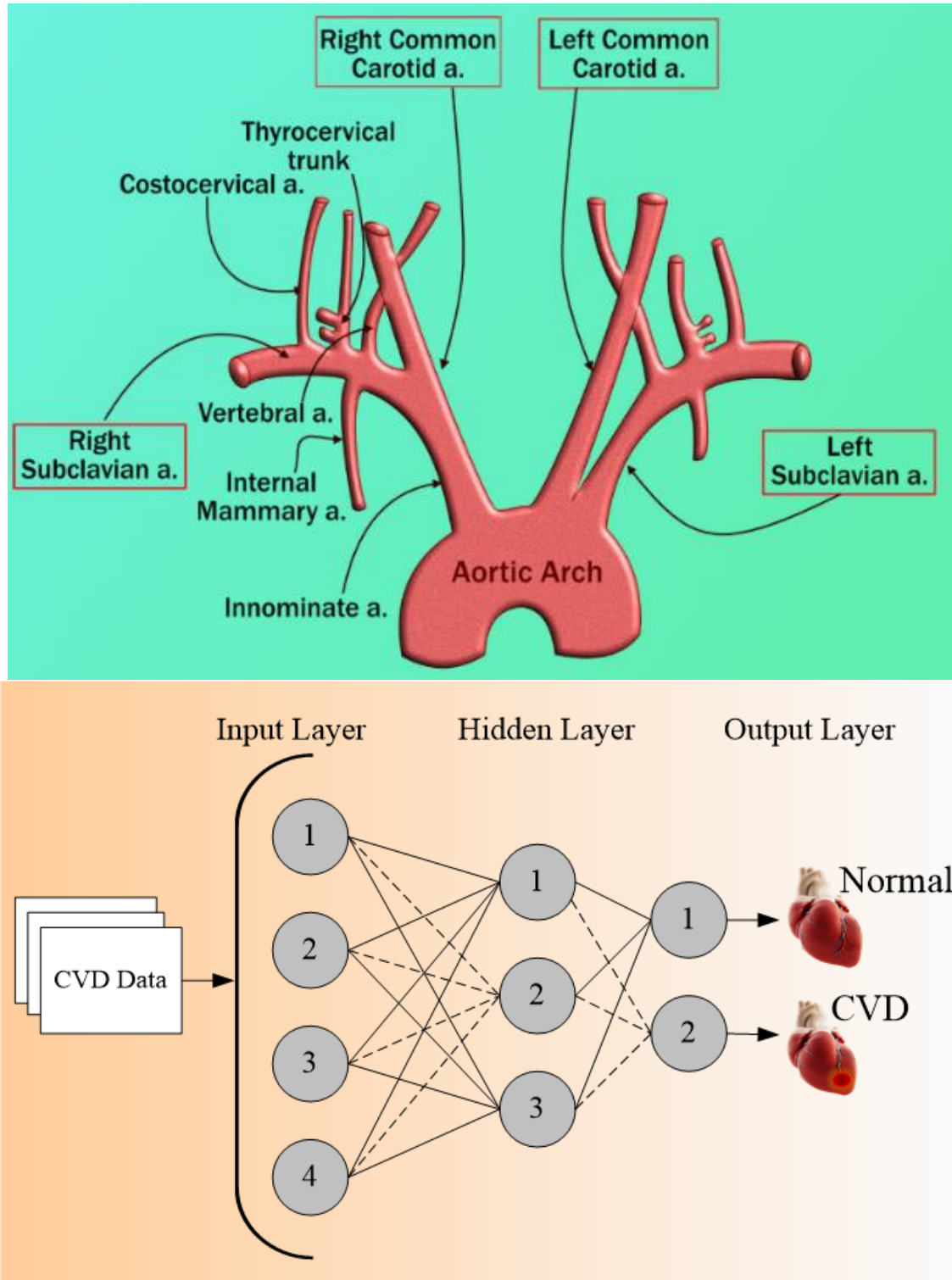


Figure A8. (Top) Anatomical link between the carotid artery and aortic arch. (Bottom) Typical neural network for CVD risk stratification.

Table A5. Acronym.

SN	Abb *	Definition	SN	Abb *	Definition
1	ACC	American college of cardiology	42	IPN	Intraplaque neovascularization
2	AD	Alzheimer's	43	KNN	K-nearest neighbor
3	AHA	American heart association	44	LBBM	Laboratory-based biomarker
4	AI	Artificial intelligence	45	LP	Label Powerset
5	ANOVA	Analysis of variance	46	LSTM	Long short-term memory network
6	APG	Acceleration Plethysmogram	47	LVD	Large vessel disease
7	ASCVD	Atherosclerotic cardiovascular disease	48	MCI	Mild cognitive impairment
8	AUC	Area-under-the-curve	49	MedUSE	Medication use
9	BCVD	Binary CVD	50	MI	Myocardial Infarction
10	BMI	Body mass index	51	ML	Machine learning
11	BR	Binary recursive	52	MLARM	Multi-label adaptive resonance asso & map
12	CAC	Coronary artery calcification	53	MLkNN	Multi-label k nearest neighbor
13	RetiCAC	Deep learning Retinal CAC score	54	MPH	Maximum plaque height
14	CAD	Coronary artery disease	55	MRI	Magnetic resonance imaging
15	CAS	Coronary artery syndrome	56	NPV	Negative predictive value
16	CC	Classifier chain	57	Non-ML	Non-machine learning
17	CCVRC	Conventional cardiovascular risk cal #	58	OBBM	Office-based biomarker
18	CHD	Coronary Heart Disease	59	PCA	principal component analysis
19	CHD	Chronic Heart Conditions	60	PCE	Pooled cohort equation
20	cIMT	Carotid intima-media thickness	61	PE	Performance evaluation matrices
21	CKD	Chronic kidney disease	62	PMCI	Progressive MCI
22	CT	Computed tomography	63	PPV	Positive predictive value
23	CUSIP	Carotid ultrasound image phenotype	64	PTC	Plaque tissue characterization
24	CV	Cross-validation	65	QRISK3	QResearch cardiovascular risk algorithm
25	CVD	Cardiovascular disease	66	RA	Rheumatoid arthritis
26	CVE	Cardiovascular events	67	RakEL	Random k-label set
27	DL	Deep learning	68	#RC	Risk classes
28	DM	Diabetes mellitus	69	RF	Random forest
29	DT	Decision tree	70	RoB	Risk-of-bias
30	ECG	Electrocardiogram	71	ROC	Receiver operating-characteristics
31	EEGS	Event-equivalent gold standard	72	RRS	Reynolds risk score
32	ESC	European society of cardiology	73	SCD	Sudden cardiac death
33	FH	Family history	74	SCG	Seismocardiography (SCG-Z)
34	FNR	False-negative rate	75	SCORE	Systematic coronary risk evaluation
35	FPR	False-positive rate	76	SCMI	Significant memory concern
36	FRS	Framingham risk score	77	SMOTE	Synthetic minority over-sampling tech.
37	GCG	Gyrocardiography	78	SVM	Support vector machine
38	GUI	Graphical user interface	79	TPA	Total plaque area
39	HTN	Hypertension	80	US	Ultrasound
40	IM	Image modalities	81	WHO	World health organization
41	IMTV	Intima-media thickness variability	-	-	-

SN: Serial Number; Abb *: Abbreviation; # Calculator; & Asso. Associative; Tech.: Technique.

References

1. Kaptoge, S.; Pennells, L.; de Bacquer, D.; Cooney, M.T.; Kavousi, M.; Stevens, G.; Riley, L.M.; Savin, S.; Khan, T.; Altay, S. World Health Organization cardiovascular disease risk charts: Revised models to estimate risk in 21 global regions. *Lancet Glob. Health* **2019**, *7*, e1332–e1345. [\[CrossRef\]](#)
2. Dunbar, S.B.; Khavjou, O.A.; Bakas, T.; Hunt, G.; Kirch, R.A.; Leib, A.R.; Morrison, R.S.; Poehler, D.C.; Roger, V.L.; Whitsel, L.P. Projected costs of informal caregiving for cardiovascular disease: 2015 to 2035: A policy statement from the American Heart Association. *Circulation* **2018**, *137*, e558–e577. [\[CrossRef\]](#)
3. Banchhor, S.K.; Londhe, N.D.; Araki, T.; Saba, L.; Radeva, P.; Khanna, N.N.; Suri, J.S. Calcium detection, its quantification, and grayscale morphology-based risk stratification using machine learning in multimodality big data coronary and carotid scans: A review. *Comput. Biol. Med.* **2018**, *101*, 184–198. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Viswanathan, V.; Jamthikar, A.D.; Gupta, D.; Shanu, N.; Puvvula, A.; Khanna, N.N.; Saba, L.; Omerzum, T.; Viskovic, K.; Mavrogeni, S. Low-cost preventive screening using carotid ultrasound in patients with diabetes. *Front. Biosci.* **2020**, *25*, 1132–1171.
5. Jamthikar, A.D.; Puvvula, A.; Gupta, D.; Johri, A.M.; Nambi, V.; Khanna, N.N.; Saba, L.; Mavrogeni, S.; Laird, J.R.; Pareek, G. Cardiovascular disease and stroke risk assessment in patients with chronic kidney disease using integration of estimated glomerular filtration rate, ultrasonic image phenotypes, and artificial intelligence: A narrative review. *Int. Angiol. J. Int. Union Angiol.* **2020**, *40*, 150–164. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Viswanathan, V.; Jamthikar, A.D.; Gupta, D.; Puvvula, A.; Khanna, N.N.; Saba, L.; Viskovic, K.; Mavrogeni, S.; Turk, M.; Laird, J.R. Integration of estimated glomerular filtration rate biomarker in image-based cardiovascular disease/stroke risk calculator: A south Asian-Indian diabetes cohort with moderate chronic kidney disease. *Int. Angiol.* **2020**, *39*, 290–306. [\[CrossRef\]](#)
7. Jamthikar, A.D.; Gupta, D.; Puvvula, A.; Johri, A.M.; Khanna, N.N.; Saba, L.; Mavrogeni, S.; Laird, J.R.; Pareek, G.; Miner, M. Cardiovascular risk assessment in patients with rheumatoid arthritis using carotid ultrasound B-mode imaging. *Rheumatol. Int.* **2020**, *40*, 1921–1939. [\[CrossRef\]](#)
8. Konstantonis, G.; Singh, K.V.; Sfikakis, P.P.; Jamthikar, A.D.; Kitas, G.D.; Gupta, S.K.; Saba, L.; Verrou, K.; Khanna, N.N.; Ruzsa, Z. Cardiovascular disease detection using machine learning and carotid/femoral arterial imaging frameworks in rheumatoid arthritis patients. *Rheumatol. Int.* **2022**, *42*, 215–239. [\[CrossRef\]](#)
9. Porcu, M.; Mannelli, L.; Melis, M.; Suri, J.S.; Gerosa, C.; Cerrone, G.; Defazio, G.; Faa, G.; Saba, L. Carotid plaque imaging profiling in subjects with risk factors (diabetes and hypertension). *Cardiovasc. Diagn. Ther.* **2020**, *10*, 1005–1018. [\[CrossRef\]](#)
10. Saba, L.; Micheletti, G.; Brinjikji, W.; Garofalo, P.; Montisci, R.; Balestrieri, A.; Suri, J.; DeMarco, J.; Lanzino, G.; Sanfilippo, R. Carotid intraplaque-hemorrhage volume and its association with cerebrovascular events. *Am. J. Neuroradiol.* **2019**, *40*, 1731–1737. [\[CrossRef\]](#)
11. Acharya, U.R.; Molinari, F.; Sree, S.V.; Chattopadhyay, S.; Ng, K.-H.; Suri, J.S. Automated diagnosis of epileptic EEG using entropies. *Biomed. Signal Process. Control* **2012**, *7*, 401–408. [\[CrossRef\]](#)
12. Acharya, U.R.; Sree, S.V.; Alvin, A.P.C.; Suri, J.S. Use of principal component analysis for automatic classification of epileptic EEG activities in wavelet framework. *Expert Syst. Appl.* **2012**, *39*, 9072–9078. [\[CrossRef\]](#)
13. El-Hasnony, I.M.; Elzeki, O.M.; Alshehri, A.; Salem, H. Multi-Label Active Learning-Based Machine Learning Model for Heart Disease Prediction. *Sensors* **2022**, *22*, 1184. [\[CrossRef\]](#)
14. Oresko, J.J.; Jin, Z.; Cheng, J.; Huang, S.; Sun, Y.; Duschl, H.; Cheng, A.C. A wearable smartphone-based platform for real-time cardiovascular disease detection via electrocardiogram processing. *IEEE Trans. Inf. Technol. Biomed.* **2010**, *14*, 734–740. [\[CrossRef\]](#)
15. Panhuyzen-Goedkoop, N.M.; Wellens, H.J.; Verbeek, A.L.; Jørstad, H.T.; Smeets, J.R.; Peters, R.J.G. ECG criteria for the detection of high-risk cardiovascular conditions in master athletes. *Eur. J. Prev. Cardiol.* **2020**, *7*, 1529–1538. [\[CrossRef\]](#)
16. Myers, P.D.; Scirica, B.M.; Stultz, C.M. Machine learning improves risk stratification after acute coronary syndrome. *Sci. Rep.* **2017**, *7*, 12692. [\[CrossRef\]](#)
17. Cuadrado-Godia, E.; Jamthikar, A.D.; Gupta, D.; Khanna, N.N.; Araki, T.; Maniruzzaman, M.; Saba, L.; Nicolaidis, A.; Sharma, A.; Omerzu, T. Ranking of stroke and cardiovascular risk factors for an optimal risk calculator design: Logistic regression approach. *Comput. Biol. Med.* **2019**, *108*, 182–195. [\[CrossRef\]](#)
18. Acharya, U.R.; Joseph, K.P.; Kannathal, N.; Min, L.C.; Suri, J.S. *Advances in Cardiac Signal Processing*; Springer: Berlin/Heidelberg, Germany, 2007.
19. Giri, D.; Acharya, U.R.; Martis, R.J.; Sree, S.V.; Lim, T.-C.; VI, T.A.; Suri, J.S. Automated diagnosis of coronary artery disease affected patients using LDA, PCA, ICA and discrete wavelet transform. *Knowl.-Based Syst.* **2013**, *37*, 274–282. [\[CrossRef\]](#)
20. Rajendra Acharya, U.; Joseph, K.P.; Kannathal, N.; Lim, C.M.; Suri, J.S. Heart rate variability: A review. *Med. Biol. Eng. Comput.* **2006**, *44*, 1031–1051. [\[CrossRef\]](#)
21. Hippisley-Cox, J.; Coupland, C.; Brindle, P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: Prospective cohort study. *BMJ* **2017**, *357*, j2099. [\[CrossRef\]](#)
22. D’Agostino Sr, R.B.; 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*, 743–753. [\[CrossRef\]](#)
23. Conroy, R.M.; Pyörälä, K.; Fitzgerald, A.E.; Sans, S.; Menotti, A.; de Backer, G.; de Bacquer, D.; Ducimetiere, P.; Jousilahti, P.; Keil, U. Estimation of ten-year risk of fatal cardiovascular disease in Europe: The SCORE project. *Eur. Heart J.* **2003**, *24*, 987–1003. [\[CrossRef\]](#)

24. Ridker, P.M.; Buring, J.E.; Rifai, N.; Cook, N.R. Development and validation of improved algorithms for the assessment of global cardiovascular risk in women: The Reynolds Risk Score. *JAMA* **2007**, *297*, 611–619. [[CrossRef](#)]
25. Goff, D.C.; Lloyd-Jones, D.M.; Bennett, G.; Coady, S.; D'agostino, R.B.; Gibbons, R.; Greenland, P.; Lackland, D.T.; Levy, D.; O'donnell, C.J. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: A report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *J. Am. Coll. Cardiol.* **2014**, *63 Pt B*, 2935–2959. [[CrossRef](#)]
26. Damman, P.; van't Hof, A.; Berg, J.T.; Jukema, J.; Appelman, Y.; Liem, A.; de Winter, R. 2015 ESC guidelines for the management of acute coronary syndromes in patients presenting without persistent ST-segment elevation: Comments from the Dutch ACS working group. *Neth. Heart J.* **2017**, *25*, 181–185. [[CrossRef](#)]
27. Members, T.F.; Montalescot, G.; Sechtem, U.; Achenbach, S.; Andreotti, F.; Arden, C.; Budaj, A.; Bugiardini, R.; Crea, F.; Cuisset, T. 2013 ESC guidelines on the management of stable coronary artery disease: The Task Force on the management of stable coronary artery disease of the European Society of Cardiology. *Eur. Heart J.* **2013**, *34*, 2949–3003.
28. Knuuti, J.; Wijns, W.; Saraste, A.; Capodanno, D.; Barbato, E.; Funck-Brentano, C.; Prescott, E.; Storey, R.F.; Deaton, C.; Cuisset, T. 2019 ESC Guidelines for the diagnosis and management of chronic coronary syndromes: The Task Force for the diagnosis and management of chronic coronary syndromes of the European Society of Cardiology (ESC). *Eur. Heart J.* **2020**, *41*, 407–477. [[CrossRef](#)]
29. Anderson, T.J.; Grégoire, J.; Pearson, G.J.; Barry, A.R.; Couture, P.; Dawes, M.; Francis, G.A.; Genest, J., Jr.; Grover, S.; Gupta, M. 2016 Canadian Cardiovascular Society guidelines for the management of dyslipidemia for the prevention of cardiovascular disease in the adult. *Can. J. Cardiol.* **2016**, *32*, 1263–1282. [[CrossRef](#)]
30. Anderson, T.J.; Grégoire, J.; Hegele, R.A.; Couture, P.; Mancini, G.J.; McPherson, R.; Francis, G.A.; Poirier, P.; Lau, D.C.; Grover, S. 2012 update of the Canadian Cardiovascular Society guidelines for the diagnosis and treatment of dyslipidemia for the prevention of cardiovascular disease in the adult. *Can. J. Cardiol.* **2013**, *29*, 151–167. [[CrossRef](#)]
31. Goldstein, B.A.; Navar, A.M.; Carter, R.E. Moving beyond regression techniques in cardiovascular risk prediction: Applying machine learning to address analytic challenges. *Eur. Heart J.* **2017**, *38*, 1805–1814. [[CrossRef](#)]
32. Deyama, J.; Nakamura, T.; Takishima, I.; Fujioka, D.; Kawabata, K.-I.; Obata, J.-E.; Watanabe, K.; Watanabe, Y.; Saito, Y.; Mishina, H. Contrast-enhanced ultrasound imaging of carotid plaque neovascularization is useful for identifying high-risk patients with coronary artery disease. *Circ. J.* **2013**, *77*, 1499–1507. [[CrossRef](#)] [[PubMed](#)]
33. Colledanchise, K.N.; Mantella, L.E.; Bullen, M.; Héту, M.-F.; Abunassar, J.G.; Johri, A.M. Combined femoral and carotid plaque burden identifies obstructive coronary artery disease in women. *J. Am. Soc. Echocardiogr.* **2020**, *33*, 90–100. [[CrossRef](#)] [[PubMed](#)]
34. Khanna, N.N.; Jamthikar, A.D.; Araki, T.; Gupta, D.; Piga, M.; Saba, L.; Carcassi, C.; Nicolaides, A.; Laird, J.R.; Suri, H.S. Nonlinear model for the carotid artery disease 10-year risk prediction by fusing conventional cardiovascular factors to carotid ultrasound image phenotypes: A Japanese diabetes cohort study. *Echocardiography* **2019**, *36*, 345–361. [[CrossRef](#)] [[PubMed](#)]
35. Jamthikar, A.; Gupta, D.; Saba, L.; Khanna, N.N.; Viskovic, K.; Mavrogeni, S.; Laird, J.R.; Sattar, N.; Johri, A.M.; Pareek, G. Artificial intelligence framework for predictive cardiovascular and stroke risk assessment models: A narrative review of integrated approaches using carotid ultrasound. *Comput. Biol. Med.* **2020**, *126*, 104043. [[CrossRef](#)] [[PubMed](#)]
36. Jamthikar, A.D.; Gupta, D.; Johri, A.M.; Mantella, L.E.; Saba, L.; Kolluri, R.; Sharma, A.M.; Viswanathan, V.; Nicolaides, A.; Suri, J.S. Low-cost office-based cardiovascular risk stratification using machine learning and focused carotid ultrasound in an Asian-Indian cohort. *J. Med. Syst.* **2020**, *44*, 208. [[CrossRef](#)] [[PubMed](#)]
37. Saba, L.; Agarwal, N.; Cau, R.; Gerosa, C.; Sanfilippo, R.; Porcu, M.; Montisci, R.; Cerrone, G.; Qi, Y.; Balestrieri, A. Review of Imaging biomarkers for the vulnerable carotid plaque. *JVS Vasc. Sci.* **2021**, *2*, 149–158. [[CrossRef](#)] [[PubMed](#)]
38. Saba, L.; Suri, J.S. *Multi-Detector CT Imaging: Principles, Head, Neck, and Vascular Systems*; CRC Press: Boca Raton, FL, USA, 2013; Volume 1.
39. Sanches, J.M.; Laine, A.F.; Suri, J.S. *Ultrasound Imaging*; Springer: Berlin/Heidelberg, Germany, 2012.
40. Londhe, N.D.; Suri, J.S. Superharmonic imaging for medical ultrasound: A review. *J. Med. Syst.* **2016**, *40*, 279. [[CrossRef](#)]
41. Sudeep, P.; Palanisamy, P.; Rajan, J.; Baradaran, H.; Saba, L.; Gupta, A.; Suri, J.S. Speckle reduction in medical ultrasound images using an unbiased non-local means method. *Biomed. Signal Process. Control* **2016**, *28*, 1–8. [[CrossRef](#)]
42. Khalifa, F.; Beache, G.M.; Gimel'farb, G.; Suri, J.S.; El-Baz, A.S. State-of-the-art medical image registration methodologies: A survey. In *Multi Modality State-of-the-Art Medical Image Segmentation and Registration Methodologies*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 235–280.
43. Roumeliotis, S.; Liakopoulos, V.; Roumeliotis, A.; Stamou, A.; Panagoutsos, S.; D'Arrigo, G.; Tripepi, G. Prognostic Factors of Fatal and Nonfatal Cardiovascular Events in Patients with Type 2 Diabetes: The Role of Renal Function Biomarkers. *Clin. Diabetes* **2021**, *39*, 188–196. [[CrossRef](#)]
44. van den Munckhof, I.C.; Jones, H.; Hopman, M.T.; de Graaf, J.; Nyakayiru, J.; van Dijk, B.; Eijsvogels, T.M.; Thijssen, D.H. Relation between age and carotid artery intima-medial thickness: A systematic review. *Clin. Cardiol.* **2018**, *41*, 698–704. [[CrossRef](#)]
45. Ho, S.S.Y. Current status of carotid ultrasound in atherosclerosis. *Quant. Imaging Med. Surg.* **2016**, *6*, 285–296. [[CrossRef](#)]
46. Touboul, P.-J.; Hennerici, M.; Meairs, S.; Adams, H.; Amarenco, P.; Bornstein, N.; Csiba, L.; Desvarieux, M.; Ebrahim, S.; Hernandez, R.H. Mannheim carotid intima-media thickness and plaque consensus (2004–2006–2011). *Cerebrovasc. Dis.* **2012**, *34*, 290–296. [[CrossRef](#)]

47. Stein, J.; Korcarz, C.; Post, W. Use of carotid ultrasound to identify subclinical vascular disease and evaluate cardiovascular disease risk: Summary and discussion of the American Society of Echocardiography consensus statement. *Prev. Cardiol.* **2009**, *12*, 34–38. [[CrossRef](#)]
48. Ikeda, N.; Araki, T.; Sugi, K.; Nakamura, M.; Deidda, M.; Molinari, F.; Meiburger, K.M.; Acharya, U.R.; Saba, L.; Bassareo, P.P. Ankle-brachial index and its link to automated carotid ultrasound measurement of intima-media thickness variability in 500 Japanese coronary artery disease patients. *Curr. Atheroscler. Rep.* **2014**, *16*, 393. [[CrossRef](#)]
49. Naqvi, T.Z.; Lee, M.-S. Carotid intima-media thickness and plaque in cardiovascular risk assessment. *JACC Cardiovasc. Imaging* **2014**, *7*, 1025–1038. [[CrossRef](#)]
50. Santos-Neto, P.J.; Sena-Santos, E.H.; Meireles, D.P.; Bittencourt, M.S.; Santos, I.S.; Bensenor, I.M.; Lotufo, P.A. Association of carotid plaques and common carotid intima-media thickness with modifiable cardiovascular risk factors. *J. Stroke Cerebrovasc. Dis.* **2021**, *30*, 105671. [[CrossRef](#)]
51. Gooty, V.D.; Sinaiko, A.R.; Ryder, J.R.; Dengel, D.R.; Jacobs, D.R., Jr.; Steinberger, J. Association between carotid intima media thickness, age, and cardiovascular risk factors in children and adolescents. *Metab. Syndr. Relat. Disord.* **2018**, *16*, 122–126. [[CrossRef](#)]
52. Johri, A.M.; Mantella, L.E.; Jamthikar, A.D.; Saba, L.; Laird, J.R.; Suri, J.S. Role of artificial intelligence in cardiovascular risk prediction and outcomes: Comparison of machine-learning and conventional statistical approaches for the analysis of carotid ultrasound features and intra-plaque neovascularization. *Int. J. Cardiovasc. Imaging* **2021**, *37*, 3145–3156. [[CrossRef](#)]
53. Johri, A.M.; Behl, P.; Héту, M.F.; Haqqi, M.; Ewart, P.; Day, A.G.; Parfrey, B.; Matangi, M.F. Carotid ultrasound maximum plaque height—A sensitive imaging biomarker for the assessment of significant coronary artery disease. *Echocardiography* **2016**, *33*, 281–289. [[CrossRef](#)]
54. Mantella, L.E.; Colledanchise, K.; Bullen, M.; Héту, M.-F.; Day, A.G.; McLellan, C.S.; Johri, A.M. Handheld versus conventional vascular ultrasound for assessing carotid artery plaque. *Int. J. Cardiol.* **2019**, *278*, 295–299. [[CrossRef](#)]
55. Saba, L.; Ikeda, N.; Deidda, M.; Araki, T.; Molinari, F.; Meiburger, K.M.; Acharya, U.R.; Nagashima, Y.; Mercurio, G.; Nakano, M. Association of automated carotid IMT measurement and HbA1c in Japanese patients with coronary artery disease. *Diabetes Res. Clin. Pract.* **2013**, *100*, 348–353. [[CrossRef](#)]
56. Jain, P.K.; Sharma, N.; Saba, L.; Paraskevas, K.I.; Kalra, M.K.; Johri, A.; Nicolaides, A.N.; Suri, J.S. Automated deep learning-based paradigm for high-risk plaque detection in B-mode common carotid ultrasound scans: An asymptomatic Japanese cohort study. *Int. Angiol. J. Int. Union Angiol.* **2021**, *41*, 9–23. [[CrossRef](#)]
57. Mitchell, C.; Korcarz, C.E.; Gepner, A.D.; Kaufman, J.D.; Post, W.; Tracy, R.; Gassett, A.J.; Ma, N.; McClelland, R.L.; Stein, J.H. Ultrasound carotid plaque features, cardiovascular disease risk factors and events: The Multi-Ethnic Study of Atherosclerosis. *Atherosclerosis* **2018**, *276*, 195–202. [[CrossRef](#)]
58. Biswas, M.; Kuppili, V.; Saba, L.; Edla, D.R.; Suri, H.S.; Cuadrado-Godia, E.; Laird, J.R.; Marinho, R.T.; Sanches, J.M.; Nicolaides, A. State-of-the-art review on deep learning in medical imaging. *Front. Biosci.* **2019**, *24*, 392–426.
59. Saba, L.; Jain, P.K.; Suri, H.S.; Ikeda, N.; Araki, T.; Singh, B.K.; Nicolaides, A.; Shafique, S.; Gupta, A.; Laird, J.R. Plaque tissue morphology-based stroke risk stratification using carotid ultrasound: A polling-based PCA learning paradigm. *J. Med. Syst.* **2017**, *41*, 98. [[CrossRef](#)]
60. Motwani, M.; Dey, D.; Berman, D.S.; Germano, G.; Achenbach, S.; Al-Mallah, M.H.; Andreini, D.; Budoff, M.J.; Cademartiri, F.; Callister, T.Q. Machine learning for prediction of all-cause mortality in patients with suspected coronary artery disease: A 5-year multicentre prospective registry analysis. *Eur. Heart J.* **2017**, *38*, 500–507. [[CrossRef](#)]
61. Alaa, A.M.; Bolton, T.; di Angelantonio, E.; Rudd, J.H.; Van der Schaar, M. Cardiovascular disease risk prediction using automated machine learning: A prospective study of 423,604 UK Biobank participants. *PLoS ONE* **2019**, *14*, e0213653.
62. Kakadiaris, I.A.; Vrigkas, M.; Yen, A.A.; Kuznetsova, T.; Budoff, M.; Naghavi, M. Machine learning outperforms ACC/AHA CVD risk calculator in MESA. *J. Am. Heart Assoc.* **2018**, *7*, e009476. [[CrossRef](#)]
63. Cawley, G.C.; Talbot, N.L. On over-fitting in model selection and subsequent selection bias in performance evaluation. *J. Mach. Learn. Res.* **2010**, *11*, 2079–2107.
64. Alalawi, H.H.; Manal, S.A. Detection of Cardiovascular Disease using Machine Learning Classification Models. *Int. J. Eng. Res. Technol. ISSN* **2021**, *10*, 2278–0181.
65. Chauhan, Y.J. Cardiovascular Disease Prediction using Classification Algorithms of Machine Learning. *Int. J. Sci. Res. ISSN* **2018**, 2319–7064.
66. Choi, E.; Schuetz, A.; Stewart, W.F.; Sun, J. Using recurrent neural network models for early detection of heart failure onset. *J. Am. Med. Inform. Assoc.* **2017**, *24*, 361–370. [[CrossRef](#)]
67. Nayan, N.A.; Hamid, H.A.; Suboh, M.Z.; Abdullah, N.; Jaafar, R.; Yusof, N.A.M.; Hamid, M.A.; Zubiri, N.F.; Arifin, A.S.K.; Daud, S.M.A. Cardiovascular Disease Prediction from Electrocardiogram by using Machine Learning Method: A Snapshot from the Subjects of the Malaysian Cohort. *Int. J. Online Biomed. Eng.* **2020**, *16*, 2626–8493.
68. Pasanisi, S.; Paiano, R. A hybrid information mining approach for knowledge discovery in cardiovascular disease (CVD). *Information* **2018**, *9*, 90. [[CrossRef](#)]
69. Sánchez-Cabo, F.; Rossello, X.; Fuster, V.; Benito, F.; Manzano, J.P.; Silla, J.C.; Fernández-Alvira, J.M.; Oliva, B.; Fernández-Friera, L.; López-Melgar, B. Machine learning improves cardiovascular risk definition for young, asymptomatic individuals. *J. Am. Coll. Cardiol.* **2020**, *76*, 1674–1685. [[CrossRef](#)]

70. Buddi, S.; Taylor, T.; Borges, C.; Nelson, R. SVM multi-classification of T2D/CVD patients using biomarker features. In Proceedings of the 2011 10th International Conference on Machine Learning and Applications and Workshops, Honolulu, HI, USA, 18–21 December 2011; pp. 338–341.
71. Chao, H.; Shan, H.; Homayounieh, F.; Singh, R.; Khera, R.D.; Guo, H.; Su, T.; Wang, G.; Kalra, M.K.; Yan, P. Deep learning predicts cardiovascular disease risks from lung cancer screening low dose computed tomography. *Nat. Commun.* **2021**, *12*, 2963. [[CrossRef](#)]
72. Devi, R.; Tyagi, H.K.; Kumar, D. A novel multi-class approach for early-stage prediction of sudden cardiac death. *Biocybern. Biomed. Eng.* **2019**, *39*, 586–598. [[CrossRef](#)]
73. Emaus, M.J.; Išgum, I.; van Velzen, S.G.; van den Bongard, H.D.; Gernaat, S.A.; Lessmann, N.; Sattler, M.G.; Teske, A.J.; Penninkhof, J.; Meijer, H. Bragatston study protocol: A multicentre cohort study on automated quantification of cardiovascular calcifications on radiotherapy planning CT scans for cardiovascular risk prediction in patients with breast cancer. *BMJ Open* **2019**, *9*, e028752. [[CrossRef](#)]
74. Hedman, Å.K.; Hage, C.; Sharma, A.; Brosnan, M.J.; Buckbinder, L.; Gan, L.-M.; Shah, S.J.; Linde, C.M.; Donal, E.; Daubert, J.-C. Identification of novel pheno-groups in heart failure with preserved ejection fraction using machine learning. *Heart* **2020**, *106*, 342–349. [[CrossRef](#)]
75. Hussein, A.F.; Hashim, S.J.; Rokhani, F.Z.; Wan Adnan, W.A. An Automated High-Accuracy Detection Scheme for Myocardial Ischemia Based on Multi-Lead Long-Interval ECG and Choi-Williams Time-Frequency Analysis Incorporating a Multi-Class SVM Classifier. *Sensors* **2021**, *21*, 2311. [[CrossRef](#)]
76. Jamthikar, A.D.; Gupta, D.; Mantella, L.E.; Saba, L.; Laird, J.R.; Johri, A.M.; Suri, J.S. Multiclass machine learning vs. conventional calculators for stroke/CVD risk assessment using carotid plaque predictors with coronary angiography scores as gold standard: A 500 participants study. *Int. J. Cardiovasc. Imaging* **2021**, *37*, 1171–1187. [[CrossRef](#)]
77. Khan, M.U.; Ali, S.Z.-e.-Z.; Ishtiaq, A.; Habib, K.; Gul, T.; Samer, A. Classification of Multi-Class Cardiovascular Disorders using Ensemble Classifier and Impulsive Domain Analysis. In Proceedings of the 2021 Mohammad Ali Jinnah University International Conference on Computing (MAJICC), Karachi, Pakistan, 15–17 July 2021; pp. 1–8.
78. Krupa, B.N.; Bharathi, K.; Gaonkar, M.; Karun, S.; Nath, S.; Ali, M. Multiclass Classification of APG Signals using ELM for CVD Risk Identification: A Real-Time Application. In Proceedings of the 16th International Conference on Biomedical Engineering, Singapore, 7–10 December 2016; Springer: Singapore, 2017; pp. 32–37.
79. Lui, H.W.; Chow, K.L. Multiclass classification of myocardial infarction with convolutional and recurrent neural networks for portable ECG devices. *Inform. Med. Unlocked* **2018**, *13*, 26–33. [[CrossRef](#)]
80. Nakanishi, R.; Slomka, P.J.; Rios, R.; Betancur, J.; Blaha, M.J.; Nasir, K.; Miedema, M.D.; Rumberger, J.A.; Gransar, H.; Shaw, L.J. Machine learning adds to clinical and CAC assessments in predicting 10-year CHD and CVD deaths. *Cardiovasc. Imaging* **2021**, *14*, 615–625. [[CrossRef](#)]
81. Ni, J.; Jiang, Y.; Zhai, S.; Chen, Y.; Li, S.; Amei, A.; Tran, D.-M.T.; Zhai, L.; Kuang, Y. Multi-class Cardiovascular Disease Detection and Classification from 12-Lead ECG Signals Using an Inception Residual Network. In Proceedings of the 2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC), Madrid, Spain, 12–16 July 2021; pp. 1532–1537.
82. Wiharto, W.; Kusnanto, H.; Herianto, H. Performance analysis of multiclass support vector machine classification for diagnosis of coronary heart diseases. *arXiv* **2015**, arXiv:1511.02352. [[CrossRef](#)]
83. Ambale-Venkatesh, B.; Yang, X.; Wu, C.O.; Liu, K.; Hundley, W.G.; McClelland, R.; Gomes, A.S.; Folsom, A.R.; Shea, S.; Guallar, E. Cardiovascular event prediction by machine learning: The multi-ethnic study of atherosclerosis. *Circ. Res.* **2017**, *121*, 1092–1101. [[CrossRef](#)]
84. Jamthikar, A.; Gupta, D.; Johri, A.M.; Mantella, L.E.; Saba, L.; Suri, J.S. A machine learning framework for risk prediction of multi-label cardiovascular events based on focused carotid plaque B-Mode ultrasound: A Canadian study. *Comput. Biol. Med.* **2021**, *140*, 105102. [[CrossRef](#)]
85. Kumar, P.; Sharma, R.; Misra, S.; Kumar, A.; Nath, M.; Nair, P.; Vibha, D.; Srivastava, A.K.; Prasad, K. CIMT as a risk factor for stroke subtype: A systematic review. *Eur. J. Clin. Investig.* **2020**, *50*, e13348. [[CrossRef](#)]
86. Mehrang, S.; Lahdenoja, O.; Kaisti, M.; Tadi, M.J.; Hurnanen, T.; Airola, A.; Knuutila, T.; Jaakkola, J.; Jaakkola, S.; Vasankari, T. Classification of Atrial Fibrillation and Acute Decompensated Heart Failure Using Smartphone Mechanocardiography: A Multilabel Learning Approach. *IEEE Sens. J.* **2020**, *20*, 7957–7968. [[CrossRef](#)]
87. Mohamed, M.; Farah, M.-C.; Fahed, A. Multi-label classification and evidential approach in diseases diagnoses using physiological signals. In Proceedings of the 2020 IEEE 5th Middle East and Africa Conference on Biomedical Engineering (MECBME), Amman, Jordan, 27–29 October 2020.
88. Nigam, P. *Applying Deep Learning to ICD-9 Multi-Label Classification from Medical Records*; Technical Report; Stanford University: Stanford, CA, USA, 2016.
89. Zamzmi, G.; Hsu, L.-Y.; Li, W.; Sachdev, V.; Antani, S. Harnessing machine intelligence in automatic echocardiogram analysis: Current status, limitations, and future directions. *IEEE Rev. Biomed. Eng.* **2020**, *14*, 181–203. [[CrossRef](#)]
90. Zeng, X.; Hu, Y.; Shu, L.; Li, J.; Duan, H.; Shu, Q.; Li, H. Explainable machine-learning predictions for complications after pediatric congenital heart surgery. *Sci. Rep.* **2021**, *11*, 17244. [[CrossRef](#)]
91. Abdar, M.; Książek, W.; Acharya, U.R.; Tan, R.-S.; Makarenkov, V.; Pławiak, P. A new machine learning technique for an accurate diagnosis of coronary artery disease. *Comput. Methods Programs Biomed.* **2019**, *179*, 104992. [[CrossRef](#)] [[PubMed](#)]

92. Baccouche, A.; Garcia-Zapirain, B.; Castillo Olea, C.; Elmaghraby, A. Ensemble deep learning models for heart disease classification: A case study from Mexico. *Information* **2020**, *11*, 207. [[CrossRef](#)]
93. Chu, H.; Chen, L.; Yang, X.; Qiu, X.; Qiao, Z.; Song, X.; Zhao, E.; Zhou, J.; Zhang, W.; Mehmood, A. Roles of anxiety and depression in predicting cardiovascular disease among patients with type 2 diabetes mellitus: A machine learning approach. *Front. Psychol.* **2021**, *12*, 645418. [[CrossRef](#)] [[PubMed](#)]
94. Cai, C.; Tafti, A.P.; Ngufor, C.; Zhang, P.; Xiao, P.; Dai, M.; Liu, H.; Noseworthy, P.; Chen, M.; Friedman, P.A. Using ensemble of ensemble machine learning methods to predict outcomes of cardiac resynchronization. *J. Cardiovasc. Electrophysiol.* **2021**, *32*, 2504–2514. [[CrossRef](#)]
95. Esfahani, H.A.; Ghazanfari, M. Cardiovascular disease detection using a new ensemble classifier. In Proceedings of the 2017 IEEE 4th International Conference on Knowledge-Based Engineering and Innovation (KBEI), Tehran, Iran, 22 December 2017; pp. 1011–1014.
96. Gibson, W.J.; Nafee, T.; Travis, R.; Yee, M.; Kerneis, M.; Ohman, M.; Gibson, C.M. Machine learning versus traditional risk stratification methods in acute coronary syndrome: A pooled randomized clinical trial analysis. *J. Thromb.* **2020**, *49*, 1–9. [[CrossRef](#)]
97. Gao, X.-Y.; Amin Ali, A.; Shaban Hassan, H.; Anwar, E.M. Improving the Accuracy for Analyzing Heart Diseases Prediction Based on the Ensemble Method. *Complexity* **2021**, *2021*, 6663455. [[CrossRef](#)]
98. Gao, L.; Ding, Y. Disease prediction via Bayesian hyperparameter optimization and ensemble learning. *BMC Res. Notes* **2020**, *13*, 205. [[CrossRef](#)]
99. Ghosh, P.; Azam, S.; Jonkman, M.; Karim, A.; Shamrat, F.J.M.; Ignatious, E.; Shultana, S.; Beeravolu, A.R.; De Boer, F. Efficient Prediction of Cardiovascular Disease Using Machine Learning Algorithms with Relief and LASSO Feature Selection Techniques. *IEEE Access* **2021**, *9*, 19304–19326. [[CrossRef](#)]
100. Hosni, M.; Carrillo de Gea, J.M.; Idri, A.; El Bajta, M.; Fernandez Aleman, J.L.; García-Mateos, G.; Abnane, I. A systematic mapping study for ensemble classification methods in cardiovascular disease. *Artif. Intell. Rev.* **2021**, *54*, 2827–2861. [[CrossRef](#)]
101. Mustafa, J.; Awan, A.A.; Khalid, M.S.; Nisar, S. Ensemble approach for developing a smart heart disease prediction system using classification algorithms. *Res. Rep. Clin. Cardiol.* **2018**, *9*, 33.
102. Jamthikar, A.D.; Gupta, D.; Mantella, L.E.; Saba, L.; Johri, A.M.; Suri, J.S. Ensemble Machine Learning and its Validation for Prediction of Coronary Artery Disease and Acute Coronary Syndrome using Focused Carotid Ultrasound. *IEEE Trans. Instrum. Meas.* **2021**, *43*, 2503810. [[CrossRef](#)]
103. Prakash, V.J.; Karthikeyan, N. Enhanced Evolutionary Feature Selection and Ensemble Method for Cardiovascular Disease Prediction. *Interdiscip. Sci. Comput. Life Sci.* **2021**, *13*, 389–412. [[CrossRef](#)]
104. Liu, N.; Li, X.; Qi, E.; Xu, M.; Li, L.; Gao, B. A novel Ensemble Learning Paradigm for Medical Diagnosis with Imbalanced Data. *IEEE Access* **2020**, *8*, 171263–171280. [[CrossRef](#)]
105. Miao, K.H.; Miao, J.H.; Miao, G.J. Diagnosing coronary heart disease using ensemble machine learning. *Int. J. Adv. Comput. Sci. Appl.* **2016**, *7*, 1–12.
106. Mienye, I.D.; Sun, Y.; Wang, Z. An improved ensemble learning approach for the prediction of heart disease risk. *Inform. Med. Unlocked* **2020**, *20*, 100402. [[CrossRef](#)]
107. Negassa, A.; Ahmed, S.; Zolty, R.; Patel, S.R. Prediction Model Using Machine Learning for Mortality in Patients with Heart Failure. *Am. J. Cardiol.* **2021**, *153*, 86–93. [[CrossRef](#)]
108. Pławiak, P.; Acharya, U.R. Novel deep genetic ensemble of classifiers for arrhythmia detection using ECG signals. *Neural Comput.* **2020**, *32*, 11137–11161. [[CrossRef](#)]
109. Reddy, K.V.V.; Elamvazuthi, I.; Aziz, A.A.; Paramasivam, S.; Chua, H.N.; Pranavanand, S. Heart Disease Risk Prediction Using Machine Learning Classifiers with Attribute Evaluators. *Appl. Sci.* **2021**, *11*, 8352. [[CrossRef](#)]
110. Rousset, A.; Dellamonica, D.; Menuet, R.; Lira Pineda, A.; Sabatine, M.S.; Giugliano, R.P.; Trichelair, P.; Zaslavskiy, M.; Ricci, L. Can machine learning bring cardiovascular risk assessment to the next level? A methodological study using FOURIER trial data. *Eur. Heart J. Digit. Health* **2021**, *093*, 93. [[CrossRef](#)]
111. Sherly, S.I. An Ensemble Based heart Disease Prediction using Gradient Boosting Decision Tree. *Turk. J. Comput. Math. Educ.* **2021**, *12*, 3648–3660.
112. Sherazi, S.W.A.; Bae, J.-W.; Lee, J.Y. A soft voting ensemble classifier for early prediction and diagnosis of occurrences of major adverse cardiovascular events for STEMI and NSTEMI during 2-year follow-up in patients with acute coronary syndrome. *PLoS ONE* **2021**, *16*, e0249338. [[CrossRef](#)]
113. Tan, C.; Chen, H.; Xia, C. The prediction of cardiovascular disease based on trace element contents in hair and a classifier of boosting decision stumps. *Biol. Trace Elem. Res.* **2009**, *129*, 9–19. [[CrossRef](#)]
114. Uddin, M.N.; Halder, R.K. An Ensemble Method Based Multilayer Dynamic System to Predict Cardiovascular Disease Using Machine Learning Approach. *Inform. Med. Unlocked* **2021**, *24*, 100584. [[CrossRef](#)]
115. Velusamy, D.; Ramasamy, K. Ensemble of heterogeneous classifiers for diagnosis and prediction of coronary artery disease with reduced feature subset. *Comput. Methods Programs Biomed.* **2021**, *198*, 105770. [[CrossRef](#)]
116. Wankhede, J.; Sambandam, P.; Kumar, M. Effective prediction of heart disease using hybrid ensemble deep learning and tunicate swarm algorithm. *J. Biomol. Struct. Dyn.* **2021**, *128*, 1–12. [[CrossRef](#)]
117. Yadav, D.C.; Pal, S. Analysis of Heart Disease Using Parallel and Sequential ensemble Methods with Feature Selection Techniques: Heart Disease Prediction. *Int. J. Big Data Anal. Healthc.* **2021**, *6*, 40–56. [[CrossRef](#)]

118. Ye, C.; Fu, T.; Hao, S.; Zhang, Y.; Wang, O.; Jin, B.; Xia, M.; Liu, M.; Zhou, X.; Wu, Q. Prediction of incident hypertension within the next year: Prospective study using statewide electronic health records and machine learning. *J. Med. Internet Res.* **2018**, *20*, e22. [[CrossRef](#)]
119. Yekkala, I.; Dixit, S.; Jabbar, M. Prediction of heart disease using ensemble learning and Particle Swarm Optimization. In Proceedings of the 2017 International Conference on Smart Technologies for Smart Nation (SmartTechCon), Bengaluru, India, 17–19 August 2017; pp. 691–698.
120. Zarkogianni, K.; Athanasiou, M.; Thanopoulou, A.C.; Nikita, K.S. Comparison of machine learning approaches toward assessing the risk of developing cardiovascular disease as a long-term diabetes complication. *IEEE J. Biomed. Health Inform.* **2017**, *22*, 1637–1647. [[CrossRef](#)]
121. Zhenya, Q.; Zhang, Z. A hybrid cost-sensitive ensemble for heart disease prediction. *BMC Med. Inform. Decis. Mak.* **2021**, *21*, 73. [[CrossRef](#)] [[PubMed](#)]
122. 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 (KDD'16), San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.
123. Chen, T.; He, T.; Benesty, M.; Khotilovich, V.; Tang, Y.; Cho, H.; Chen, K. Xgboost: Extreme gradient boosting. *R Package Version 0.4-2* **2015**, *1*, 1–4.
124. Hansson, G.K.; Libby, P.; Tabas, I. Inflammation and plaque vulnerability. *J. Intern. Med.* **2015**, *278*, 483–493. [[CrossRef](#)] [[PubMed](#)]
125. Li, H.; Sun, K.; Zhao, R.; Hu, J.; Hao, Z.; Wang, F.; Lu, Y.; Liu, F.; Zhang, Y. Inflammatory biomarkers of coronary heart disease. *Front. Biosci.* **2017**, *22*, 504–515.
126. Libby, P.; Ridker, P.M.; Hansson, G.K.; Leducq Transatlantic Network on Atherothrombosis. Inflammation in atherosclerosis: From pathophysiology to practice. *J. Am. Coll. Cardiol.* **2009**, *54*, 2129–2138. [[CrossRef](#)] [[PubMed](#)]
127. Ross, R. Cell biology of atherosclerosis. *Annu. Rev. Physiol.* **1995**, *57*, 791–804. [[CrossRef](#)] [[PubMed](#)]
128. Tabas, I.; Williams, K.J.; Borén, J. Subendothelial lipoprotein retention as the initiating process in atherosclerosis: Update and therapeutic implications. *Circulation* **2007**, *116*, 1832–1844. [[CrossRef](#)]
129. Virmani, R.; Burke, A.P.; Kolodgie, F.D.; Farb, A. Pathology of the thin-cap fibroatheroma: A type of vulnerable plaque. *J. Interv. Cardiol.* **2003**, *16*, 267–272. [[CrossRef](#)]
130. Burke, A.P.; Kolodgie, F.D.; Farb, A.; Weber, D.; Virmani, R. Morphological predictors of arterial remodeling in coronary atherosclerosis. *Circulation* **2002**, *105*, 297–303. [[CrossRef](#)]
131. Patel, A.K.; Suri, H.S.; Singh, J.; Kumar, D.; Shafique, S.; Nicolaides, A.; Jain, S.K.; Saba, L.; Gupta, A.; Laird, J.R. A review on atherosclerotic biology, wall stiffness, physics of elasticity, and its ultrasound-based measurement. *Curr. Atheroscler. Rep.* **2016**, *18*, 83. [[CrossRef](#)]
132. Arroyo, L.H.; Lee, R.T. Mechanisms of plaque rupture: Mechanical and biologic interactions. *Cardiovasc. Res.* **1999**, *41*, 369–375. [[CrossRef](#)]
133. Teng, Z.; Zhang, Y.; Huang, Y.; Feng, J.; Yuan, J.; Lu, Q.; Sutcliffe, M.P.; Brown, A.J.; Jing, Z.; Gillard, J.H. Material properties of components in human carotid atherosclerotic plaques: A uniaxial extension study. *Acta Biomater.* **2014**, *10*, 5055–5063. [[CrossRef](#)]
134. Kumar, P.R.; Priya, M. Classification of atherosclerotic and non-atherosclerotic individuals using multiclass support vector machine. *Technol. Health Care* **2014**, *22*, 583–595. [[CrossRef](#)]
135. Herr, J.E.; Héту, M.-F.; Li, T.Y.; Ewart, P.; Johri, A.M. Presence of calcium-like tissue composition in carotid plaque is indicative of significant coronary artery disease in high-risk patients. *J. Am. Soc. Echocardiogr.* **2019**, *32*, 633–642. [[CrossRef](#)]
136. Jeong, B.; Cho, H.; Kim, J.; Kwon, S.K.; Hong, S.; Lee, C.; Kim, T.; Park, M.S.; Hong, S.; Heo, T.-Y. Comparison between statistical models and machine learning methods on classification for highly imbalanced multiclass kidney data. *Diagnostics* **2020**, *10*, 415. [[CrossRef](#)]
137. Rim, T.H.; Lee, C.J.; Tham, Y.-C.; Cheung, N.; Yu, M.; Lee, G.; Kim, Y.; Ting, D.S.; Chong, C.C.Y.; Choi, Y.S. Deep-learning-based cardiovascular risk stratification using coronary artery calcium scores predicted from retinal photographs. *Lancet Digit. Health* **2021**, *3*, e306–e316. [[CrossRef](#)]
138. Tandel, G.S.; Balestrieri, A.; Jujaray, T.; Khanna, N.N.; Saba, L.; Suri, J.S. Multiclass magnetic resonance imaging brain tumor classification using artificial intelligence paradigm. *Comput. Biol. Med.* **2020**, *122*, 103804. [[CrossRef](#)]
139. Mercan, C.; Aksoy, S.; Mercan, E.; Shapiro, L.G.; Weaver, D.L.; Elmore, J.G. Multi-instance multi-label learning for multi-class classification of whole slide breast histopathology images. *IEEE Trans. Med. Imaging* **2017**, *37*, 316–325. [[CrossRef](#)]
140. Unnikrishnan, P.; Kumar, D.K.; Poosapadi Arjunan, S.; Kumar, H.; Mitchell, P.; Kawasaki, R. Development of health parameter model for risk prediction of CVD using SVM. *Comput. Math. Methods Med.* **2016**, *2016*, 3016245. [[CrossRef](#)]
141. Cheng, B.; Liu, M.; Zhang, D.; Shen, D. Robust multi-label transfer feature learning for early diagnosis of Alzheimer's disease. *Brain Imaging Behav.* **2019**, *13*, 138–153. [[CrossRef](#)]
142. Nikhar, S.; Karandikar, A. Prediction of heart disease using machine learning algorithms. *Int. J. Adv. Eng. Manag. Sci.* **2016**, *2*, 239484.
143. Rosengren, A.; Hawken, S.; Ôunpuu, S.; Sliwa, K.; Zubaid, M.; Almahmeed, W.A.; Blackett, K.N.; Sitthi-Amorn, C.; Sato, H.; Yusuf, S. Association of psychosocial risk factors with risk of acute myocardial infarction in 11,119 cases and 13,648 controls from 52 countries (the INTERHEART study): Case-control study. *Lancet* **2004**, *364*, 953–962. [[CrossRef](#)]

144. Yusuf, S.; Hawken, S.; Ôunpuu, S.; Dans, T.; Avezum, A.; Lanas, F.; McQueen, M.; Budaj, A.; Pais, P.; Varigos, J. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): Case-control study. *Lancet* **2004**, *364*, 937–952. [[CrossRef](#)]
145. Shrivastava, V.K.; Londhe, N.D.; Sonawane, R.S.; Suri, J.S. Reliable and accurate psoriasis disease classification in dermatology images using comprehensive feature space in machine learning paradigm. *Expert Syst. Appl.* **2015**, *42*, 6184–6195. [[CrossRef](#)]
146. Araki, T.; Ikeda, N.; Shukla, D.; Jain, P.K.; Londhe, N.D.; Shrivastava, V.K.; Banchhor, S.K.; Saba, L.; Nicolaidis, A.; Shafique, S.; et al. PCA-based polling strategy in machine learning framework for coronary artery disease risk assessment in intravascular ultrasound: A link between carotid and coronary grayscale plaque morphology. *Comput. Methods Programs Biomed.* **2016**, *128*, 137–158. [[CrossRef](#)]
147. Acharya, U.R.; Mookiah, M.R.K.; Sree, S.V.; Yanti, R.; Martis, R.; Saba, L.; Molinari, F.; Guerriero, S.; Suri, J.S. Evolutionary algorithm-based classifier parameter tuning for automatic ovarian cancer tissue characterization and classification. *Ultraschall Der Med.-Eur. J. Ultrasound* **2014**, *35*, 237–245.
148. Olier, I.; Ortega-Martorell, S.; Pieroni, M.; Lip, G.Y. How machine learning is impacting research in atrial fibrillation: Implications for risk prediction and future management. *Cardiovasc. Res.* **2021**, *117*, 1700–1717. [[CrossRef](#)]
149. Angelaki, E.; Marketou, M.E.; Barmparis, G.D.; Patrianakos, A.; Vardas, P.E.; Parthenakis, F.; Tsironis, G.P. Detection of abnormal left ventricular geometry in patients without cardiovascular disease through machine learning: An ECG-based approach. *J. Clin. Hypertens.* **2021**, *23*, 935–945. [[CrossRef](#)]
150. Shen, Y.; Yang, Y.; Parish, S.; Chen, Z.; Clarke, R.; Clifton, D.A. Risk prediction for cardiovascular disease using ECG data in the China Kadoorie Biobank. In Proceedings of the 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, FL, USA, 16–20 August 2016.
151. Chang, A.; Cadaret, L.M.; Liu, K. Machine learning in electrocardiography and echocardiography: Technological advances in clinical cardiology. *Curr. Cardiol. Rep.* **2020**, *22*, 161. [[CrossRef](#)]
152. Bos, J.M.; Attia, Z.I.; Albert, D.E.; Noseworthy, P.A.; Friedman, P.A.; Ackerman, M.J. Use of artificial intelligence and deep neural networks in evaluation of patients with electrocardiographically concealed long QT syndrome from the surface 12-lead electrocardiogram. *JAMA Cardiol.* **2021**, *6*, 532–538. [[CrossRef](#)] [[PubMed](#)]
153. Paragliola, G.; Coronato, A. An hybrid ECG-based deep network for the early identification of high-risk to major cardiovascular events for hypertension patients. *J. Biomed. Inform.* **2021**, *113*, 103648. [[CrossRef](#)] [[PubMed](#)]
154. Wang, S.; Li, R.; Wang, X.; Shen, S.; Zhou, B.; Wang, Z. Multiscale Residual Network Based on Channel Spatial Attention Mechanism for Multilabel ECG Classification. *J. Healthc. Eng.* **2021**, *2021*, 6630643. [[CrossRef](#)] [[PubMed](#)]
155. Xu, J.; Xu, Z.-X.; Lu, P.; Guo, R.; Yan, H.-X.; Xu, W.-J.; Wang, Y.-Q.; Xia, C.-M. Classifying syndromes in Chinese medicine using multi-label learning algorithm with relevant features for each label. *Chin. J. Integr. Med.* **2016**, *22*, 867–871. [[CrossRef](#)]
156. Longato, E.; Fadini, G.P.; Sparacino, G.; Avogaro, A.; Tramontan, L.; Di Camillo, B. A deep learning approach to predict diabetes' cardiovascular complications from administrative claims. *IEEE J. Biomed. Health Inform.* **2021**, *25*, 3608–3617. [[CrossRef](#)]
157. Fang, J.; Xu, Y.; Zhao, Y.; Yan, Y.; Liu, J.; Liu, J. Weighing features of lung and heart regions for thoracic disease classification. *BMC Med. Imaging* **2021**, *21*, 99. [[CrossRef](#)]
158. Cen, L.-P.; Ji, J.; Lin, J.-W.; Ju, S.-T.; Lin, H.-J.; Li, T.-P.; Wang, Y.; Yang, J.-F.; Liu, Y.-F.; Tan, S. Automatic detection of 39 fundus diseases and conditions in retinal photographs using deep neural networks. *Nat. Commun.* **2021**, *12*, 4828. [[CrossRef](#)]
159. Deng, F.; Zhou, H.; Lin, Y.; Heim, J.A.; Shen, L.; Li, Y.; Zhang, L. Predict multicategory causes of death in lung cancer patients using clinicopathologic factors. *Comput. Biol. Med.* **2021**, *129*, 104161. [[CrossRef](#)]
160. Wagenaar, D.J.; Chen, J.A. Nuclear imaging of vulnerable plaque: Contrast improvements through multi-labeling of nanoparticles. In Proceedings of the IEEE Nuclear Science Symposium Conference Record, Fajardo, PR, USA, 23–29 October 2005; p. 5.
161. Jie, M.; Hong, Z. Image classification algorithm based on LTS-HD multi instance multi label RBF. In Proceedings of the 2017 12th IEEE Conference on Industrial Electronics and Applications (ICIEA), Siem Reap, Cambodia, 18–20 June 2017; pp. 190–194.
162. Wu, Q.; Tan, M.; Song, H.; Chen, J.; Ng, M.K. ML-Forest: A multi-label tree ensemble method for multi-label classification. *IEEE Trans. Knowl. Data Eng.* **2016**, *28*, 2665–2680. [[CrossRef](#)]
163. Li, G.-Z.; He, Z.; Shao, F.-F.; Ou, A.-H.; Lin, X.-Z. Patient classification of hypertension in Traditional Chinese Medicine using multi-label learning techniques. *BMC Med. Genom.* **2015**, *8*, S4. [[CrossRef](#)]
164. Sun, Z.; Wang, C.; Zhao, Y.; Yan, C. Multi-label ECG signal classification based on ensemble classifier. *IEEE Access* **2020**, *8*, 117986–117996. [[CrossRef](#)]
165. Rezaei Ravari, M.; Eftekhari, M.; Saberi Movahed, F. ML-CK-ELM: An efficient multi-layer extreme learning machine using combined kernels for multi-label classification. *Sci. Iran.* **2020**, *27*, 3005–3018. [[CrossRef](#)]
166. Rezaei-Ravari, M.; Eftekhari, M.; Saberi-Movahed, F. Regularizing extreme learning machine by dual locally linear embedding manifold learning for training multi-label neural network classifiers. *Eng. Appl. Artif. Intell.* **2021**, *97*, 104062. [[CrossRef](#)]
167. Baghel, N.; Dutta, M.K.; Burget, R.J.C.M.; Biomedicine, P.i. Automatic diagnosis of multiple cardiac diseases from PCG signals using convolutional neural network. *Comput. Methods Programs Biomed.* **2020**, *197*, 105750. [[CrossRef](#)]
168. Cuadrado-Godia, E.; Dwivedi, P.; Sharma, S.; Santiago, A.O.; Gonzalez, J.R.; Balcells, M.; Laird, J.; Turk, M.; Suri, H.S.; Nicolaidis, A.; et al. Cerebral small vessel disease: A review focusing on pathophysiology, biomarkers, and machine learning strategies. *J. Stroke* **2018**, *20*, 302–320. [[CrossRef](#)]

169. Huang, W.; Ying, T.W.; Chin, W.L.C.; Baskaran, L.; Marcus, O.E.H.; Yeo, K.K.; Kiong, N.S. Application of ensemble machine learning algorithms on lifestyle factors and wearables for cardiovascular risk prediction. *Sci. Rep.* **2022**, *12*, 1033. [CrossRef]
170. Brownless, J. Nested Cross-Validation for Machine Learning with Python. 2020. Available online: <https://machinelearningmastery.com/nested-cross-validation-for-machine-learning-with-python/> (accessed on 20 February 2022).
171. Pintelas, P.; Livieris, I.E. Special issue on ensemble learning and applications. *Algorithms* **2020**, *13*, 140. [CrossRef]
172. Lo, H.-Y.; Lin, S.-D.; Wang, H.-M. Generalized k-labelsets ensemble for multi-label and cost-sensitive classification. *IEEE Trans. Knowl. Data Eng.* **2013**, *26*, 1679–1691.
173. Wang, P.; Ge, R.; Xiao, X.; Zhou, M.; Zhou, F. hMuLab: A biomedical hybrid MULTi-LABEL classifier based on multiple linear regression. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2016**, *14*, 1173–1180. [CrossRef]
174. Guo, L.; Jin, B.; Yu, R.; Yao, C.; Sun, C.; Huang, D. Multi-label classification methods for green computing and application for mobile medical recommendations. *IEEE Access* **2016**, *4*, 3201–3209. [CrossRef]
175. Farooq, A.; Anwar, S.; Awais, M.; Rehman, S. A deep CNN based multi-class classification of Alzheimer’s disease using MRI. In Proceedings of the 2017 IEEE International Conference on Imaging Systems and Techniques (IST), Beijing, China, 18–20 October 2017; pp. 1–6.
176. Zuluaga, M.A.; Cardoso, M.J.; Ourselin, S. Automatic right ventricle segmentation using multi-label fusion in cardiac MRI. *arXiv* **2020**, arXiv:2004.02317.
177. Al Hinai, G.; Jammoul, S.; Vajih, Z.; Afilalo, J. Deep learning analysis of resting electrocardiograms for the detection of myocardial dysfunction, hypertrophy, and ischaemia: A systematic review. *Eur. Heart J.-Digit. Health* **2021**, *2*, 416–423. [CrossRef]
178. Mahinrad, S.; Ferguson, I.; Macfarlane, P.W.; Clark, E.N.; Stott, D.J.; Ford, I.; Mooijaart, S.P.; Trompet, S.; Van Heemst, D.; Jukema, J.W. Spatial QRS-T angle and cognitive decline in older subjects. *J. Alzheimer’s Dis.* **2019**, *67*, 279–289. [CrossRef]
179. Narayan, S.M.; Wang, P.J.; Daubert, J.P. New concepts in sudden cardiac arrest to address an intractable epidemic: JACC state-of-the-art review. *J. Am. Coll. Cardiol.* **2019**, *73*, 70–88. [CrossRef] [PubMed]
180. Puvar, P.; Patel, N.; Shah, A.; Solanki, R.; Rana, D. Heart Disease Detection using Ensemble Learning Approach. *Int. Res. J. Eng. Technol. (IRJET)* **2021**, *8*, 2395-0072.
181. Gibaja, E.; Ventura, S. A tutorial on multilabel learning. *ACM Comput. Surv.* **2015**, *47*, 1–38. [CrossRef]
182. Krstinić, D.; Braović, M.; Šerić, L.; Božić-Štulić, D. Multi-label classifier performance evaluation with confusion matrix. *Comput. Sci. Inf. Technol.* **2020**, *10*, 1–14.
183. Tsoumakas, G.; Katakis, I. Multi-label classification: An overview. *Int. J. Data Warehous. Min.* **2007**, *3*, 1–13. [CrossRef]
184. Spolaôr, N.; Cherman, E.A.; Monard, M.C.; Lee, H.D. A comparison of multi-label feature selection methods using the problem transformation approach. *Electron. Notes Theor. Comput. Sci.* **2013**, *292*, 135–151. [CrossRef]
185. Xia, C.; Vonder, M.; Sidorenkov, G.; Ma, R.; Oudkerk, M.; van der Harst, P.; De Deyn, P.P.; Vliegthart, R. Coronary Artery Calcium and Cognitive Function in Dutch Adults: Cross-Sectional Results of the Population-Based ImaLife Study. *J. Am. Heart Assoc.* **2021**, *10*, e018172. [CrossRef]
186. Ribeiro, A.L.P.; Paixao, G.M.; Gomes, P.R.; Ribeiro, M.H.; Ribeiro, A.H.; Canazart, J.A.; Oliveira, D.M.; Ferreira, M.P.; Lima, E.M.; de Moraes, J.L. Tele-electrocardiography and bigdata: The CODE (Clinical Outcomes in Digital Electrocardiography) study. *J. Electrocardiol.* **2019**, *57*, S75–S78. [CrossRef]
187. Castelyn, G.; Laranjo, L.; Schreier, G.; Gallego, B. Predictive performance and impact of algorithms in remote monitoring of chronic conditions: A systematic review and meta-analysis. *Int. J. Med. Inform.* **2021**, *156*, 104620. [CrossRef]
188. Chugh, S.S.; Reinier, K.; Teodorescu, C.; Evanado, A.; Kehr, E.; Al Samara, M.; Mariani, R.; Gunson, K.; Jui, J. Epidemiology of sudden cardiac death: Clinical and research implications. *Prog. Cardiovasc. Dis.* **2008**, *51*, 213–228. [CrossRef]
189. Masarone, D.; Limongelli, G.; Ammendola, E.; Verrengia, M.; Gravino, R.; Pacileo, G. Risk stratification of sudden cardiac death in patients with heart failure: An update. *J. Clin. Med.* **2018**, *7*, 436. [CrossRef]
190. Palacios-Rubioa, J.; Marina-Breyse, M.; Quintanilla, J.G.; Gil-Perdomo, J.M.; Juárez-Fernández, M.; Garcia-Gonzalez, I.; Rial-Bastón, V.; Corcobado, M.C.; Espinosa, M.C.; Ruiz, F. Early prognostic value of an Algorithm based on spectral Variables of Ventricular fibrillation from the EKG of patients with sudden cardiac death: A multicentre observational study (AWAKE). *Arch. Cardiol. México* **2018**, *88*, 460–467. [CrossRef]
191. Filgueiras-Rama, D.; Calvo, C.J.; Salvador-Montañés, Ó.; Cadenas, R.; Ruiz-Cantador, J.; Armada, E.; Rey, J.R.; Merino, J.; Peinado, R.; Pérez-Castellano, N. Spectral analysis-based risk score enables early prediction of mortality and cerebral performance in patients undergoing therapeutic hypothermia for ventricular fibrillation and comatose status. *Int. J. Cardiol.* **2015**, *186*, 250–258. [CrossRef]
192. Hussein, A.F.; Hashim, S.J.; Aziz, A.F.A.; Rokhani, F.Z.; Adnan, W.A.W. Performance evaluation of time-frequency distributions for ECG signal analysis. *J. Med. Syst.* **2018**, *42*, 15. [CrossRef]
193. Mehrabi, N.; Morstatter, F.; Saxena, N.; Lerman, K.; Galstyan, A. A survey on bias and fairness in machine learning. *ACM Comput. Surv.* **2021**, *54*, 1–35. [CrossRef]
194. Cabrera, Á.A.; Epperson, W.; Hohman, F.; Kahng, M.; Morgenstern, J.; Chau, D.H. FairVis: Visual analytics for discovering intersectional bias in machine learning. In Proceedings of the 2019 IEEE Conference on Visual Analytics Science and Technology (VAST), Vancouver, BC, Canada, 20–25 October 2019.
195. Schelter, S.; Stoyanovich, J. Taming technical bias in machine learning pipelines. *Bull. Tech. Comm. Data Eng.* **2020**, *43*, 39–50.

196. Suri, J.; Agarwal, S.; Gupta, S.K.; Puvvula, A.; Viskovic, K.; Suri, N.; Alizad, A.; El-Baz, A.; Saba, L.; Fatemi, M. Systematic review of artificial intelligence in acute respiratory distress syndrome for COVID-19 lung patients: A biomedical imaging perspective. *IEEE J. Biomed. Health Inform.* **2021**, *25*, 4128–4139. [[CrossRef](#)]
197. Buss, V.H.; Leesong, S.; Barr, M.; Varnfield, M.; Harris, M. Primary Prevention of Cardiovascular Disease and Type 2 Diabetes Mellitus Using Mobile Health Technology: Systematic Review of the Literature. *J. Med. Internet Res.* **2020**, *22*, e21159. [[CrossRef](#)]
198. Villarreal, V.; Berbey-Alvarez, A. Evaluation of mHealth Applications Related to Cardiovascular Diseases: A Systematic Review. *Acta Inform. Med.* **2020**, *28*, 130–137. [[CrossRef](#)]
199. Xiao, Q.; Lu, S.; Wang, Y.; Sun, L.; Wu, Y.; e-Health. Current status of cardiovascular disease-related smartphone apps downloadable in China. *Telemedicine* **2017**, *23*, 219–225. [[CrossRef](#)]
200. Saba, L.; Banchhor, S.K.; Araki, T.; Viskovic, K.; Londhe, N.D.; Laird, J.R.; Suri, H.S.; Suri, J.S. Intra-and inter-operator reproducibility of automated cloud-based carotid lumen diameter ultrasound measurement. *Indian Heart J.* **2018**, *70*, 649–664. [[CrossRef](#)]
201. Huda, N.; Khan, S.; Abid, R.; Shuvo, S.B.; Labib, M.M.; Hasan, T. A Low-cost, Low-energy Wearable ECG System with Cloud-Based Arrhythmia Detection. In Proceedings of the 2020 IEEE Region 10 Symposium (TENSYP), Dhaka, Bangladesh, 5–7 June 2020; pp. 1840–1843.
202. Singh, K.K.; Singh, S.S. An Artificial Intelligence based mobile solution for early detection of valvular heart diseases. In Proceedings of the 2019 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), Bangalore, India, 26–27 July 2019; pp. 1–5.
203. Spanakis, E.G.; Psaraki, M.; Sakkalis, V. Congestive heart failure risk assessment monitoring through internet of things and mobile personal health systems. In Proceedings of the 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Honolulu, HI, USA, 18–21 July 2018; pp. 2925–2928.
204. Paredes, S.; Henriques, J.; Rocha, T.; de Carvalho, P.; Morais, J.; Santos, L.; Carvalho, R. The lookAfterRisk Project: Dynamic Cardiovascular Risk Assessment based on Remote Monitoring Solutions. In Proceedings of the 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Madrid, Spain, 3–6 December 2018; pp. 1122–1125.
205. Freyer, L.; von Stülpnagel, L.; Spielbichler, P.; Sappl, N.; Wenner, F.; Schreinlechner, M.; Krasniqi, A.; Behroz, A.; Eiffener, E.; Zens, M. Rationale and design of a digital trial using smartphones to detect subclinical atrial fibrillation in a population at risk: The eHealth-based bavarian alternative detection of Atrial Fibrillation (eBRAVE-AF) trial. *Am. Heart J.* **2021**, *241*, 26–34. [[CrossRef](#)]
206. Giansanti, D.; Monoscalco, L. A smartphone-based survey in mHealth to investigate the introduction of the artificial intelligence into cardiology. *Mhealth* **2021**, *7*, 8. [[CrossRef](#)]
207. Park, J.; Lee, K.; Kang, K. Pit-a-Pat: A smart electrocardiogram system for detecting arrhythmia. *Telemed. e-Health* **2015**, *21*, 814–821. [[CrossRef](#)]
208. Lillo-Castellano, J.; Mora-Jimenez, I.; Santiago-Mozos, R.; Chavarría-Asso, F.; Cano-González, A.; García-Alberola, A.; Rojo-Álvarez, J.L. Symmetrical compression distance for arrhythmia discrimination in cloud-based big-data services. *IEEE J. Biomed. Health Inform.* **2015**, *19*, 1253–1263. [[CrossRef](#)]
209. Sakellarios, A.; Siogkas, P.; Georga, E.; Tachos, N.; Kigka, V.; Tsompou, P.; Andrikos, I.; Karanasiou, G.S.; Rocchiccioli, S.; Correia, J. A clinical decision support platform for the risk stratification, diagnosis, and prediction of coronary artery disease evolution. In Proceedings of the 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Honolulu, HI, USA, 18–21 July 2018; pp. 4556–4559.
210. Suri, J.S.; Rangayyan, R.M. *Breast Imaging, Mammography, and Computer-Aided Diagnosis of Breast Cancer*; SPIE: Bellingham, WA, USA, 2006.
211. Boernama, A.W.D.; Setiawan, N.A.; Wahyunggoro, O. Multiclass classification of brain-computer interface motor imagery system: A systematic literature review. In Proceedings of the 2021 International Conference on Artificial Intelligence and Mechatronics Systems (AIMS), Bandung, Indonesia, 28–30 April 2021; pp. 1–6.
212. Collins, D.R.; Tompson, A.C.; Onakpoya, I.J.; Roberts, N.; Ward, A.M.; Heneghan, C.J. Global cardiovascular risk assessment in the primary prevention of cardiovascular disease in adults: Systematic review of systematic reviews. *BMJ Open* **2017**, *7*, e013650. [[CrossRef](#)]
213. Dissanayake, K.; Md Johar, M.G. Comparative Study on Heart Disease Prediction Using Feature Selection Techniques on Classification Algorithms. *Appl. Comput. Intell. Soft Comput.* **2021**, *2021*, 5581806. [[CrossRef](#)]
214. Galar, M.; Fernandez, A.; Barrenechea, E.; Bustince, H.; Herrera, F. A review on ensembles for the class imbalance problem: Bagging-, boosting-, and hybrid-based approaches. *IEEE Trans. Syst. Man Cybern. Part C* **2011**, *42*, 463–484. [[CrossRef](#)]
215. Stewart, J.; Manmathan, G.; Wilkinson, P. Primary prevention of cardiovascular disease: A review of contemporary guidance and literature. *JRSM Cardiovasc. Dis.* **2017**, *6*, 2048004016687211. [[CrossRef](#)]
216. Mathew, R.M.; Gunasundari, R. A Review on Handling Multiclass Imbalanced Data Classification In Education Domain. In Proceedings of the 2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India, 4–5 March 2021; pp. 752–755.
217. Uike, D.; Thorat, S. Implementation of Multiclass Algorithm for Sickle Cell Identification and Categorization—A Review. In Proceedings of the 2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), Bangalore, India, 5–7 March 2020; pp. 300–303.

218. Wang, H.; Liu, X.; Lv, B.; Yang, F.; Hong, Y. Reliable multi-label learning via conformal predictor and random forest for syndrome differentiation of chronic fatigue in traditional Chinese medicine. *PLoS ONE* **2014**, *9*, e99565. [[CrossRef](#)]
219. Wiharto, W.; Kusnanto, H.; Herianto, H. Intelligence system for diagnosis level of coronary heart disease with K-star algorithm. *Healthc. Inform. Res.* **2016**, *22*, 30–38. [[CrossRef](#)] [[PubMed](#)]
220. Boi, A.; Jamthikar, A.D.; Saba, L.; Gupta, D.; Sharma, A.; Loi, B.; Laird, J.R.; Khanna, N.N.; Suri, J.S. A survey on coronary atherosclerotic plaque tissue characterization in intravascular optical coherence tomography. *Curr. Atheroscler. Rep.* **2018**, *20*, 33. [[CrossRef](#)] [[PubMed](#)]
221. Bianchini, E.; Corciu, A.; Venneri, L.; Faita, F.; Giannarelli, C.; Gemignani, V.; Demi, M. Assessment of cardiovascular risk markers from ultrasound images: System reproducibility. In Proceedings of the 2008 Computers in Cardiology, Bologna, Italy, 14–17 September 2008; pp. 105–108.
222. Liu, L.; Tang, L. A Survey of Statistical Topic Model for Multi-Label Classification. In Proceedings of the 2018 26th International Conference on Geoinformatics, Kunming, China, 28–30 June 2018; pp. 1–5.
223. Charte, F. A comprehensive and didactic review on multilabel learning software tools. *IEEE Access* **2020**, *8*, 50330–50354. [[CrossRef](#)]
224. Siblinski, W.; Kuntz, P.; Meyer, F. A review on dimensionality reduction for multi-label classification. *IEEE Trans. Knowl. Data Eng.* **2019**, *33*, 839–857. [[CrossRef](#)]
225. Indhumathi, M.; Kumar, V.A. Healthcare Management of Major Cardiovascular Disease—A review. In Proceedings of the 2021 6th International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 20–22 January 2021; pp. 1230–1235.
226. Kolli, K.K.; Han, D.; Gransar, H.; Lee, J.H.; Choi, S.-Y.; Chun, E.J.; Jung, H.O.; Sung, J.; Han, H.-W.; Park, S.H. Machine learning algorithm to predict coronary artery calcification in asymptomatic healthy population. In Proceedings of the 2019 IEEE Healthcare Innovations and Point of Care Technologies, (HI-POCT), Bethesda, MD, USA, 20–22 November 2019; pp. 95–98.
227. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: Synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **2002**, *16*, 321–357. [[CrossRef](#)]
228. Acharya, U.R.; Joseph, K.P.; Kannathal, N.; Min, L.C.; Suri, J.S. Heart rate variability. In *Advances in Cardiac Signal Processing*; Springer: Berlin/Heidelberg, Germany, 2007; pp. 121–165.
229. Quang, D.; Chen, Y.; Xie, X. DANN: A deep learning approach for annotating the pathogenicity of genetic variants. *Bioinformatics* **2015**, *31*, 761–763. [[CrossRef](#)]
230. Ding, C.; Chen, P.; Jiao, J. Non-linear effects of the built environment on automobile-involved pedestrian crash frequency: A machine learning approach. *Accid. Anal. Prev.* **2018**, *112*, 116–126. [[CrossRef](#)]
231. Eder, S.J.; Nicholson, A.A.; Stefanczyk, M.M.; Pieniak, M.; Martínez-Molina, J.; Pešout, O.; Binter, J.; Smela, P.; Scharnowski, F.; Steyrl, D. Securing your relationship: Quality of intimate relationships during the COVID-19 pandemic can be predicted by attachment style. *Front. Psychol.* **2021**, 3016. [[CrossRef](#)]
232. Milicevic, O.; Salom, I.; Rodic, A.; Markovic, S.; Tumbas, M.; Zigic, D.; Djordjevic, M.; Djordjevic, M. PM2. 5 as a major predictor of COVID-19 basic reproduction number in the USA. *Environ. Res.* **2021**, *201*, 111526. [[CrossRef](#)]
233. Soltan, A.A.; Kouchaki, S.; Zhu, T.; Kiyasseh, D.; Taylor, T.; Hussain, Z.B.; Peto, T.; Brent, A.J.; Eyre, D.W.; Clifton, D.A. Rapid triage for COVID-19 using routine clinical data for patients attending hospital: Development and prospective validation of an artificial intelligence screening test. *Lancet Digit. Health* **2021**, *3*, e78–e87. [[CrossRef](#)]
234. Acharya, U.R.; Sree, S.V.; Ang, P.C.A.; Yanti, R.; Suri, J.S. Application of non-linear and wavelet based features for the automated identification of epileptic EEG signals. *Int. J. Neural Syst.* **2012**, *22*, 1250002. [[CrossRef](#)]
235. Poddar, M.G.; Kumar, V.; Sharma, Y.P. Automated diagnosis of coronary artery diseased patients by heart rate variability analysis using linear and non-linear methods. *J. Med. Eng. Technol.* **2015**, *39*, 331–341. [[CrossRef](#)]
236. Tran, J.; Sharma, D.; Gotlieb, N.; Xu, W.; Bhat, M. Application of machine learning in liver transplantation: A review. *Hepatol. Int.* **2022**, 1–14. [[CrossRef](#)]
237. Shandilya, S.; Ward, K.; Kurz, M.; Najarian, K. Non-linear dynamical signal characterization for prediction of defibrillation success through machine learning. *BMC Med. Inform. Decis. Mak.* **2012**, *12*, 116. [[CrossRef](#)]
238. Singh, R.S.; Saini, B.S.; Sunkaria, R.K. Detection of coronary artery disease by reduced features and extreme learning machine. *Clujul Med.* **2018**, *91*, 166. [[CrossRef](#)]
239. Shandilya, S.; Kurz, M.C.; Ward, K.R.; Najarian, K. Integration of Attributes from Non-Linear Characterization of Cardiovascular Time-Series for Prediction of Defibrillation Outcomes. *PLoS ONE* **2016**, *11*, e0141313.
240. Hedjazi, N.; Kharboutly, H.; Benali, A.; Dibi, Z. PCA-based selection of distinctive stability criteria and classification of post-stroke pathological postural behaviour. *Australas. Phys. Eng. Sci. Med.* **2018**, *41*, 189–199. [[CrossRef](#)]
241. Li, X.; Ling, S.H.; Su, S. A hybrid feature selection and extraction methods for sleep apnea detection using bio-signals. *Sensors* **2020**, *20*, 4323. [[CrossRef](#)]
242. Jo, Y.-Y.; Kwon, J.-m.; Jeon, K.-H.; Cho, Y.-H.; Shin, J.-H.; Lee, Y.-J.; Jung, M.-S.; Ban, J.-H.; Kim, K.-H.; Lee, S.Y. Detection and classification of arrhythmia using an explainable deep learning model. *J. Electrocardiol.* **2021**, *67*, 124–132. [[CrossRef](#)]
243. Muthulakshmi, M.; Kavitha, G. Deep CNN with LM learning based myocardial ischemia detection in cardiac magnetic resonance images. In Proceedings of the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, 23–27 July 2019; pp. 824–827.

244. Liang, Y.; Yin, S.; Tang, Q.; Zheng, Z.; Elgendi, M.; Chen, Z. Deep learning algorithm classifies heartbeat events based on electrocardiogram signals. *Front. Physiol.* **2020**, *11*, 1255. [[CrossRef](#)]
245. Tadesse, G.A.; Javed, H.; Weldemariam, K.; Liu, Y.; Liu, J.; Chen, J.; Zhu, T. DeepMI: Deep multi-lead ECG fusion for identifying myocardial infarction and its occurrence-time. *Artif. Intell. Med.* **2021**, *121*, 102192. [[CrossRef](#)]
246. Tadesse, G.A.; Zhu, T.; Liu, Y.; Zhou, Y.; Chen, J.; Tian, M.; Clifton, D. Cardiovascular disease diagnosis using cross-domain transfer learning. In Proceedings of the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, 23–27 July 2019; pp. 4262–4265.
247. Butun, E.; Yildirim, O.; Talo, M.; Tan, R.-S.; Acharya, U.R. 1D-CADCapsNet: One dimensional deep capsule networks for coronary artery disease detection using ECG signals. *Phys. Med.* **2020**, *70*, 39–48. [[CrossRef](#)]
248. Wang, T.; Lu, C.; Sun, Y.; Yang, M.; Liu, C.; Ou, C. Automatic ECG classification using continuous wavelet transform and convolutional neural network. *Entropy* **2021**, *23*, 119. [[CrossRef](#)]
249. Wang, R.; Fan, J.; Li, Y. Deep multi-scale fusion neural network for multi-class arrhythmia detection. *IEEE J. Biomed. Health Inform.* **2020**, *24*, 2461–2472. [[CrossRef](#)]
250. Liu, Y.; Li, Q.; Wang, K.; Liu, J.; He, R.; Yuan, Y.; Zhang, H. Automatic Multi-Label ECG Classification with Category Imbalance and Cost-Sensitive Thresholding. *Biosensors* **2021**, *11*, 453. [[CrossRef](#)]
251. El-Baz, A.; Gimel'farb, G.; Suri, J.S. *Stochastic Modeling for Medical Image Analysis*; CRC Press: Boca Raton, FL, USA, 2015.
252. Lloyd-Jones, D.M. Cardiovascular risk prediction: Basic concepts, current status, and future directions. *Circulation* **2010**, *121*, 1768–1777. [[CrossRef](#)]
253. Farzadfar, F. Cardiovascular disease risk prediction models: Challenges and perspectives. *Lancet Glob. Health* **2019**, *7*, e1288–e1289. [[CrossRef](#)]
254. Kuppili, V.; Biswas, M.; Sreekumar, A.; Suri, H.S.; Saba, L.; Edla, D.R.; Marinho, R.T.; Sanches, J.M.; Suri, J.S. Extreme learning machine framework for risk stratification of fatty liver disease using ultrasound tissue characterization. *J. Med. Syst.* **2017**, *41*, 152. [[CrossRef](#)]
255. Banerjee, A.; Chen, S.; Pasea, L.; Lai, A.G.; Katsoulis, M.; Denaxas, S.; Nafilyan, V.; Williams, B.; Wong, W.K.; Bakhai, A. Excess deaths in people with cardiovascular diseases during the COVID-19 pandemic. *Eur. J. Prev. Cardiol.* **2021**, *28*, 1599–1609. [[CrossRef](#)]
256. Magadam, A.; Kishore, R. Cardiovascular manifestations of COVID-19 infection. *Cells* **2020**, *9*, 2508.
257. Jain, P.K.; Sharma, N.; Saba, L.; Paraskevas, K.I.; Kalra, M.K.; Johri, A.; Laird, J.R.; Nicolaidis, A.N.; Suri, J.S. Unseen Artificial Intelligence—Deep Learning Paradigm for Segmentation of Low Atherosclerotic Plaque in Carotid Ultrasound: A Multicenter Cardiovascular Study. *Diagnostics* **2021**, *11*, 2257. [[CrossRef](#)] [[PubMed](#)]
258. Cherman, E.A.; Monard, M.C.; Metz, J. Multi-label problem transformation methods: A case study. *CLEI Electron. J.* **2011**, *14*, 4. [[CrossRef](#)]
259. Read, J.; Pfahringer, B.; Holmes, G.; Frank, E. Classifier chains for multi-label classification. *Mach. Learn.* **2011**, *85*, 333–359. [[CrossRef](#)]
260. Boutell, M.R.; Luo, J.; Shen, X.; Brown, C.M. Learning multi-label scene classification. *Pattern Recognit.* **2004**, *37*, 1757–1771. [[CrossRef](#)]
261. Tsoumakas, G.; Vlahavas, I. Random k-labelsets: An ensemble method for multilabel classification. In *Proceedings of European Conference on Machine Learning*; Springer: Berlin/Heidelberg, Germany, 2007; pp. 406–417.
262. Zhang, M.-L.; Zhou, Z.-H. ML-KNN: A lazy learning approach to multi-label learning. *Pattern Recognit.* **2007**, *40*, 2038–2048. [[CrossRef](#)]
263. Benites, F.; Sapozhnikova, E. Haram: A hierarchical aram neural network for large-scale text classification. In Proceedings of the 2015 IEEE international Conference on Data Mining Workshop (ICDMW), Atlantic City, NJ, USA, 14–17 November 2015; pp. 847–854.
264. Naing, N.N. Determination of sample size. *Malays. J. Med. Sci. MJMS* **2003**, *10*, 84.
265. Qualtrics, S.S. *Determining Sample Size: How to Ensure You Get the Correct Sample Size*; Qualtrics: Seattle, WA, USA, 2019.