

Comprehensive Evaluation of Ensemble Learning (EL) Models in Cardiovascular Diseases Management: A Meta-Synthesis Approach

Mahmoud Zahedian

Nezhad 

Ph.D student, Department of Information Technology Management, Ferdowsi University of Mashhad, Mashhad, Iran

Mohammad Mehraeen  *

Professor, Department of Information Technology Management, Ferdowsi University of Mashhad, Mashhad, Iran

Rouhollah Bagheri 

Associate Professor, Department of Information Technology Management, Ferdowsi University of Mashhad, Mashhad, Iran

Seyyed Mohammad

Tabatabae 

Assistant Professor, Department of Medical Informatics, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran

Abstract

Cardiovascular diseases (CVDs) are among the leading causes of mortality worldwide. The increasing volume of complex data generated by diagnostic tools, such as electrocardiograms (ECG), poses significant challenges for clinicians, impacting both diagnostic accuracy and treatment speed. Ensemble Learning (EL), by integrating multiple models, offers enhanced performance in CVD management; however, few studies have systematically compared its techniques or provided a comprehensive classification of EL applications in this domain. Therefore, this study aims to categorize EL models in CVD management, evaluate their performance and efficiency within each category, identify the advantages and limitations of each model, and analyze the role of feature engineering in their application. To achieve these objectives, a meta-synthesis approach was employed, examining

* Corresponding Author: mehraeen@um.ac.ir

How to Cite: xxxxxxxx

the use of EL models in combination with Machine Learning (ML) and Deep Learning (DL) techniques for CVD management. The seven-step method proposed by Sandelowski and Barroso (۲۰۰۷) was applied for the meta-synthesis. Initially, relevant articles were retrieved from the Web of Science using predefined keywords. Subsequently, articles were screened based on title, abstract, accessibility, full content, and research methodology, and the final selection was made for detailed analysis. Findings from the meta-synthesis indicate that EL applications in CVD management can be classified into four main domains: prediction (۷۰ articles), diagnosis (۳۰ articles), identification (۱۸ articles), and classification (۹ articles). The results confirm that EL models are dominant across all four domains, with their effectiveness significantly enhanced when integrated with ML and DL techniques. Among various approaches, Random Forest (RF) and gradient boosting algorithms such as XGBoost were the most frequently used and identified as the most efficient and accurate models. This study provides a structured overview, offering valuable insights for researchers in designing intelligent CVD prediction and diagnostic systems, as well as for clinicians in utilizing these systems, and establishes a framework for applying hybrid models to achieve more precise and effective CVD management.

۱. Introduction

Effective management of cardiovascular diseases (CVDs) is increasingly challenged by the complexity and increasing volume of data generated by modern diagnostic technologies. This data-rich environment complicates clinical interpretation and timely diagnosis. Therefore, artificial intelligence (AI) and machine learning (ML) have become essential tools, and ensemble learning (EL) techniques—such as boosting, bagging, and stacking—have emerged as particularly powerful for building robust predictive models by integrating diverse data sources. Despite significant advances in this field, there is still a gap in the systematic and comprehensive review of the literature on the effectiveness, strengths, and limitations of EL techniques across all areas of CVDs management. This study aims to address this gap by providing a comprehensive review and providing insights that will

enable researchers and practitioners to apply these models with greater accuracy and interpretability.

Research Questions:

۱. How can EL models used in CVD management be categorized?
۲. Which EL models achieve the highest accuracy within each CVD management category?
۳. What limitations and challenges exist in applying EL models to CVD management?
۴. How are EL models typically validated and evaluated?
۵. What preprocessing and feature engineering methods are used in datasets designed for EL-based CVD management?

۲. Literature Review

AI covers a wide range of methods that try to mimic how people think and solve problems. At the heart of it all, you've got ML and deep learning (DL). ML uses stats and algorithms to learn from organized data, while DL leans on artificial neural networks to pull out deeper patterns from tricky stuff like medical images or signals. Now, when we want to boost how well these models predict things and make them more reliable, you turn to EL. EL pulls together predictions from a bunch of different models and blends them into one answer that's usually more accurate. The real point of EL is to cut down on errors by mixing up different models—so if one misses something, another can catch it. It will often see approaches like bagging, boosting, stacking, and voting used here. Ensemble methods are valuable in healthcare, where model reliability and accuracy across diverse patient datasets are crucial.

۲. Methodology

A meta-synthesis method was employed in this research to comprehensively analyze and interpret literature results on EL and ML/DL models in CVD management. It conforms to a seven-step guideline process by Sandelowski and Barroso in ۲۰۰۷, which is most appropriate for meta-synthesis, especially for incorporating qualitative studies with their quantitative counterparts. Intensive literature searches were undertaken on Web of Science with its publication years ranging

from ۲۰۱۰-۲۰۲۴, and with flexible searching keywords such as “Cardiovascular Diseases,” “Machine Learning,” “Deep Learning.”

Starting with ۱,۳۲۳ articles, a rigorous staged process of screening was undertaken, with consideration for title, abstract, availability, and final relevance. Finally, ۱۳۴ articles met all criteria to be included in final meta-synthesis for consideration in their synthesis. Data from these articles was employed for analysis in meta-synthesis to finally group models for EL into four groups regarding management for CVDs.

۴. Results

The synthesis of ۱۳۴ studies shows that EL models in CVD management fall into four primary categories: prediction, diagnosis, classification, and identification. Prediction is the most extensively studied domain, represented by ۷۵ articles focusing on forecasting disease onset, stratifying patient risk, optimizing healthcare resources, and predicting survival outcomes. Diagnosis (۳۰ studies) and classification (۱۸ studies) are also significant areas where EL and ML/DL models are used to detect CVDs from biomedical signals such as ECG and PCG, as well as to differentiate between different types of heart diseases. Identification (۹ studies), though less common, focuses on recognizing key risk factors to support preventive interventions.

In terms of model adoption, EL methods have exceptionally high popularity in the literature. Random Forest is noticed to be adopted most in its individual capacity as an algorithm, representing ۱۱۱ occurrences. Amongst popular boosting algorithms, XGBoost leads, with Gradient Boosting not too far behind in second place, while AdaBoost follows. Amongst traditional ML algorithms, Support Vector Machines and Logistic Regression algorithm remain ever popular in their respective applications in classification problems. On the other hand, the DL landscape is diverse, with ANNs and their variants like MLPs being the most common, while Convolutional Neural Networks are pivotal for image and signal analysis, and Recurrent Neural Networks like LSTMs are typically applied to sequential data.

۵. Discussion and Conclusion

EL models, especially RF and XGBoost, play a crucial role in managing CVDs in the primary areas of prediction, diagnosis, classification, and

identification. One major upcoming trend involves the synergy between EL and DL models such as CNNs, which perform exceptionally on complex biomedical patterns inherent in classification and identification problems. Notably, these models require critical preprocessing steps in data, with techniques for noise reduction, dimensionality reduction, or correction for class imbalance, such as SMOTE, being necessary for improving model accuracy, consistency, and generalizability. Nevertheless, there are numerous challenges in clinical practice, especially due to their dependency on small-scale studies, computational complexity, or BI proprieties, signifying critical importance for model validation via rigorous testing for accuracy, sensitivity, specificity, and AUC-ROC, while SHAP analysis or XAI analysis is being increasingly called for in overcoming proprieties in model interpretability.

Keywords: Cardiovascular Diseases Management, Meta-synthesis Approach, Ensemble Learning (EL), Machine Learning (ML), Deep Learning (DL).

مجله دانش پزشکی
انتشار

ارزیابی جامع مدل‌های یادگیری گروهی (EL) در مدیریت بیماری‌های قلبی عروقی: رویکردی فراترکیب

محمود زاهدیان نژاد
دانشجوی دکتری رشته مدیریت فناوری اطلاعات، گروه مدیریت، دانشگاه فردوسی مشهد، مشهد، ایران



محمد مهراین*
استاد مدیریت فناوری اطلاعات، گروه مدیریت، دانشگاه فردوسی مشهد، مشهد، ایران



روح الله باقری
دانشیار مدیریت فناوری اطلاعات، گروه مدیریت، دانشگاه فردوسی مشهد، مشهد، ایران

سید محمد طباطبایی
استادیار انفورماتیک پزشکی، گروه انفورماتیک پزشکی، دانشکده پزشکی، دانشگاه علوم پزشکی مشهد، مشهد، ایران



چکیده

بیماری‌های قلبی-عروقی (CVDs) یکی از علل اصلی مرگ و میر در سراسر جهان محسوب می‌شوند. افزایش داده‌های پیچیده حاصل از ابزارهای تشخیصی مانند الکتروکاردیوگرام (ECG)، چالش‌های قابل توجهی را برای پزشکان ایجاد کرده که بر دقت تشخیص و سرعت درمان تأثیر می‌گذارد. یادگیری گروهی (EL) با ترکیب مدل‌های مختلف، عملکرد بهتری را در مدیریت CVD ها ارائه می‌دهد، اما تحقیقات محدودی به صورت سیستماتیک تکنیک‌های مختلف آن را مقایسه کرده‌اند و به بررسی یک طبقه‌بندی جامع از بکارگیری مدل‌های گروهی در CVD ها پرداخته‌اند. بنابراین، هدف این مطالعه، دسته‌بندی مدل‌های EL در مدیریت CVD ها، ارزیابی عملکرد و کارایی آن‌ها در هر دسته، شناسایی مزایا و محدودیت‌های هر مدل و تحلیل نقش مهندسی ویژگی در بکارگیری آن‌ها در مدیریت CVD ها است. این پژوهش برای دستیابی به اهداف تحقیق از رویکرد فراترکیب استفاده کرده که به بررسی کاربرد مدل‌های EL در ترکیب با مدل‌های یادگیری ماشین (ML) و یادگیری عمیق (DL) در مدیریت CVD ها می‌پردازد. برای بکارگیری روش فراتکیب (Meta-Synthesis) در این تحقیق از رویکرد هفت گانه سندلوسکی و باروسو (۲۰۰۷) استفاده شده است. در ابتدا با استفاده از کلیدواژه‌های مشخص شده و جست و جوی آن‌ها در پایگاه علمی وب آو ساینس، منجر به استخراج مقاله شده و سپس با بررسی مقالات از نظر عنوان، چکیده، دسترسی پذیری، تمام

* نویسنده مسئول: mehraeen@um.ac.ir

محتوا و روش شناسی تحقیق، مقاله جهت بررسی نهایی انتخاب شدند. یافته‌های این فراترکیب نشان می‌دهد که کاربرد مدل‌های یادگیری در مدیریت CVD ها به چهار حوزه‌ی اصلی تقسیم می‌شود: پیش‌بینی (۷۵ مقاله)، تشخیص (۳۰ مقاله)، شناسایی (۱۸ مقاله) و طبقه‌بندی (۹ مقاله). نتایج تأیید می‌کند که مدل‌های EL در تمام این چهار حوزه غالب هستند و کارایی آن‌ها با ادغام با تکنیک‌های ML و DL به طور قابل توجهی افزایش می‌یابد. در میان رویکردهای مختلف، مدل‌های جنگل تصادفی (RF) و الگوریتم‌های تقویت گرادیان مانند XGBoost، بیشترین فراوانی استفاده را داشته و به عنوان کارآمدترین و دقیق‌ترین مدل‌ها شناخته می‌شوند. این پژوهش با ارائه یک نمای کلی ساختاریافته، بینش‌های ارزشمندی را برای محققان در طراحی سیستم‌های هوشمند تشخیص و پیش‌بینی CVD ها و متخصصان بالینی جهت استفاده از این سیستم‌ها فراهم می‌کند و چارچوبی برای به کارگیری مدل‌های ترکیبی جهت دستیابی به مدیریت دقیق‌تر و مؤثرتر بیماری‌های قلبی-عروقی ارائه می‌دهد.

کلیدواژه‌ها: مدیریت بیماری‌های قلبی-عروقی (CVDs)، رویکرد فراترکیب، یادگیری گروهی (EL)، یادگیری ماشین (ML)، یادگیری عمیق (DL).

آماده انتشار

مقدمه

بیماری‌های قلب و عروقی^۱ (CVDs) شامل طیف گسترده‌ای از اختلالات است که قلب و عروق خونی را تحت تاثیر قرار می‌دهد، که باعث ایجاد یک بحران جهانی همراه با بالاترین میزان مرگ و میر می‌شود. CVD ها شامل چندین زیرگروه با خواص منحصر به فرد، عوامل خطر و چالش‌های تشخیصی هستند. این گروه شامل طیف گسترده‌ای از آسیب‌شناسی‌ها مانند بیماری عروق کرونر^۲ (CAD)، حمله قلبی، نارسایی قلبی، فشار خون بالا، آریتمی، بیماری شریانی محیطی و بیماری دریچه قلب است (Nabel ۲۰۰۳; Virani et al. ۲۰۲۰). بیماری‌های قلبی عروقی بیماری‌های چند عاملی هستند که تحت تأثیر عوامل خطر مختلفی از طیف وسیعی از عوامل فیزیولوژیکی، ژنتیکی و رفتاری قرار می‌گیرند (Roth et al. ۲۰۲۰; Virani et al. ۲۰۲۰).

در حالی که مدیریت و کنترل عوامل خطر مختلف می‌تواند بروز و شروع سریع بیماری‌های قلبی عروقی را کاهش دهد، شناسایی عوامل مهم، تشخیص زودهنگام و پیش‌بینی دقیق شروع بیماری برای مداخله موثر و بهبود پیامدهای بیمار بسیار مهم است. با پیشرفت‌های تکنولوژیکی در دهه‌های اخیر در ابزارهای تشخیصی، از جمله الکتروکاردیوگرام^۳ (ECG)، اسکن توموگرافی کامپیوتری (CT)، اکوکاردیوگرافی، تکنیک‌های تصویربرداری پیشرفته و سنسورهای پوشیدنی، همگی باعث تشخیص زودهنگام بیماری شده و به نظارت مداوم و بلادرنگ بیماران کمک کرده‌اند (Celermajer et al. ۲۰۱۲) (Moradi et al. ۲۰۲۳; Sun et al. ۲۰۲۱). با این حال، ساختار پیچیده سیستم قلبی عروقی، ابزارهای تشخیصی مختلف و داده‌های مرتبط با آنها و همچنین حجم عظیمی از داده‌های تولید شده از این سیستم پیچیده، چالش بزرگی را برای پزشکان جهت تفسیر این داده‌ها و در نهایت تشخیص بیماری و حتی نوع آن ایجاد کرده است (Chen et al. ۲۰۲۱). در چنین شرایطی، نیاز به روش‌های محاسباتی هوشمند که بتوانند این داده‌های پیچیده را تحلیل کرده و الگوهای پنهان در آنها را شناسایی کنند بیش از پیش احساس می‌شود. در این راستا، فناوری‌هایی مانند هوش مصنوعی^۴ (AI)، یادگیری ماشین^۵ (ML) و یادگیری

^۱. Cardiovascular diseases (CVDs)

^۲. Coronary Artery Disease (CAD)

^۳. Electrocardiogram (ECG)

^۴. Artificial Intelligence (AI)

^۵. Machine Learning (ML)

عمیق^۱ (DL) نقش برجسته ای در پردازش و تفسیر داده‌ها و در نهایت شناسایی، تشخیص و پیش بینی CVD ها ایفا کرده اند (Sidey-Gibbons and Sidey-Gibbons ۲۰۱۹). نوآوری‌های اخیر در یادگیری گروهی (EL) مانند طبقه بندی کننده‌ها و مدل‌های تقویت^۲، بگینگ^۳، پشته ای یا انباشتگی^۴ و رای گیری^۵، راه را برای مدل‌های مدیریت CVD های قوی تر با ادغام داده‌های چند گانه هموار کرده است (Mahajan et al. ۲۰۲۳; Miao, Miao, and Miao ۲۰۱۶). ادغام این تکنیک‌ها با سایر مدل‌های ML و DL، که در آن مدل‌های آموزش دیده بر روی یک گروه برای استفاده در گروه دیگر سازگار می شوند، پتانسیل غلبه بر مدیریت این بیماری‌ها را نشان داده است (Tiwari, Chugh, and Sharma ۲۰۲۲). علیرغم مزایای آن، مطالعات نقش محدودی در بررسی اثربخشی تکنیک‌های مختلف EL و مقایسه آنها در تشخیص و پیش بینی جامع CVD ها ایفا کرده اند. به عنوان مثال در تحقیقی به تازگی، گوئر و همکاران (۲۰۲۵)، مروری بر پیش بینی بیماری نارسائی قلبی (HF) با استفاده از مدل‌های یادگیری گروهی، با در نظر گرفتن متغیرهای محیطی و سلامت، انجام داده اند. این مطالعه، تحقیقات مربوط به روش‌های گروهی شناخته شده مانند AdaBoost، CatBoost، XGBoost، جنگل تصادفی، طبقه بندی کننده‌های انباشت و رای گیری را خلاصه و ترکیب کرده است. نتایج آنها نشان می دهد که از بین این روش‌ها، روش‌های انباشت و رای گیری همیشه از نظر پیش بینی خطر HF نسبت به سایرین برتر بوده‌اند و هر دو به دقت نسبتاً خوب و عملکرد پایدار در مجموعه داده‌های مختلف دست یافته‌اند. همچنین رویکردهای اخیر مانند تقویت حداقل خطا، BOO-ST و CBCEC را که با مسائل متعددی مانند عدم تعادل داده‌ها و انتخاب ویژگی سروکار دارند، مورد بحث قرار داده اند (Gour et al. ۲۰۲۵).

تاکور و همکاران (۲۰۲۳)، نقش روش‌های گروهی را در تشخیص قلب و عروق برجسته کرده اند. همچنین، یزدی و اسدی (۲۰۲۵)، در مطالعه ای روشی پیشرفته به نام بهینه سازی یادگیری گروهی پویا مبتنی بر GMM را با استفاده از یک الگوریتم تکاملی چند هدفه مبتنی بر تجزیه (MOEA/D) برای تشخیص CVD معرفی می کند. این رویکرد از نمونه گیری

^۱. Deep Learning (DL)

^۲. Boosting

^۳. Bagging

^۴. Stacking

^۵. Voting

مجدد تکاملی الهام گرفته از مدل مخلوط گاوسی (GMM) برای افزایش تنوع، عملکرد و کمیت طبقه‌بندی‌کننده‌ها استفاده می‌کند و مجموعه داده‌های بهینه را تولید می‌کند. روش پیشنهادی آنها بر روی هفت مجموعه داده بیماری قلبی آزمایش شده است که نتایج آنها نشان دهنده عملکرد عالی الگوریتم‌های گروهی است (Yazdi and Asadi ۲۰۲۵). در تحقیقی دیگر، کرتاناوونگ و همکاران (۲۰۲۰)، نقش الگوریتم‌های ML در پیش‌بینی CVD را مورد بررسی و ارزیابی قرار داده‌اند. بنابراین، فقدان تجزیه و تحلیل سیستماتیک و جامع از اثربخشی، نقاط قوت، محدودیت‌ها و کاربرد بالینی و همچنین همگرایی مدل‌های گروهی، مانع از درک کامل آن در مدیریت CVD‌ها شده است. برای رفع این شکاف، پژوهش حاضر نوآوری خود را در قالب یک تحلیل نظام‌مند از مطالعات موجود درباره‌ی مدل‌های EL در مدیریت این بیماری‌ها ارائه می‌کند.

این پژوهش به صورت عملیاتی از طریق ۱) طبقه‌بندی جدید و ساختارمند مدل‌ها و تکنیک‌های EL بر اساس حوزه‌های کاربردی مختلف در مدیریت CVD‌ها (مانند تشخیص، پیش‌بینی و طبقه‌بندی خطر)، ۲) مقایسه‌ی کمی معیارهای عملکردی مدل‌ها از نظر معیارهایی مانند دقت، حساسیت، خاصیت و AUC، ۳) تحلیل روش‌های اعتبارسنجی به کاررفته در مطالعات منتخب، ۴) شناسایی و طبقه‌بندی چالش‌ها و محدودیت‌های موجود در پیاده‌سازی مدل‌های EL، و در نهایت ۵) بررسی جامع روش‌ها و فنون مدیریت داده‌ها و مهندسی ویژگی‌ها برای به کارگیری مؤثر مدل‌های EL در داده‌های چندمنبعی CVD، به شکاف پژوهشی موجود پاسخ می‌دهد.

بنابراین، پرداختن به این شکاف، بینش‌های ارزشمندی را برای محققان و متخصصان CVD فراهم می‌کند که به دنبال استفاده از این مدل‌ها به شیوه‌ای دقیق‌تر، قابل اعتمادتر و قابل تفسیرتر در مدیریت این بیماری‌ها هستند. از این رو، این پژوهش به دنبال رفع این شکاف پژوهشی با طرح و پاسخ به سؤالات زیر است.

۱. بکارگیری مدل‌ها و تکنیک‌های EL در مدیریت CVD‌ها در چند دسته تقسیم بندی می‌شوند؟
۲. کدام مدل‌های EL بیشترین دقت و عملکرد را در هر دسته از مدیریت CVD‌ها داشته‌اند؟
۳. چگونه مدل‌ها و تکنیک‌های EL در مدیریت CVD‌ها اعتبارسنجی می‌شوند؟

4. محدودیت‌ها و چالش‌های مربوط به بکارگیری مدل‌های EL در مدیریت CVDها چیست؟

5. جهت بکارگیری مدل‌های EL در مدیریت داده‌ها CVD چه تکنیک‌ها و

روش‌های پیش پردازش داده‌ها و مهندسی ویژگی‌های استفاده می‌شود؟

برای پاسخ به سؤالات مطرح شده، این مطالعه از یک روش فراترکیب استفاده می‌کند که به طور سیستماتیک یافته‌های ادبیات موجود را ادغام و تجزیه و تحلیل می‌کند. در ادامه، مفاهیم اساسی EL و ML/DL در بخش دوم بررسی و به دنبال آن مروری بر ادبیات مربوط به کاربرد این روش‌ها در مدیریت CVDها در بخش سوم انجام می‌شود. بخش چهارم روش تحقیق را با تمرکز بر رویکرد فراترکیب و پایگاه داده وب آو ساینس توصیف می‌کند. بخش پنجم یافته‌های فراترکیب را در استفاده از مدل‌های یادگیری، به ویژه EL، در مدیریت CVDها ارائه می‌دهد. بخش ششم شامل یک بررسی جامع از مدیریت CVDها از نظر چهار دسته اصلی یعنی تشخیص، پیش‌بینی، طبقه‌بندی و شناسایی بر اساس مدل‌های یادگیری است. بخش‌های هفتم، هشتم و نهم به ترتیب معیارهای ارزیابی عملکرد مدل‌ها، محدودیت‌ها و چالش‌های استفاده از مدل‌های یادگیری گروهی و ترکیبی و اهمیت مهندسی ویژگی و پیش پردازش داده‌ها را مورد بحث قرار می‌دهند. در نهایت، مقاله با بخش دهم به پایان می‌رسد که نتیجه‌گیری کلی را ارائه می‌دهد.

پیشینه پژوهش

مبانی نظری پژوهش

هوش مصنوعی حوزه گسترده‌ای است که هدف آن ایجاد ماشین‌ها یا سیستم‌هایی است که قادر به شبیه‌سازی عملکردهای شناختی شبیه انسان مانند ادراک، استدلال و تصمیم‌گیری هستند. در زمینه CVDها، AI این امکان را فراهم کرده است تا از حجم عظیم داده‌های بالینی، تصویری و سیگنال‌های فیزیولوژیکی الگوهای پنهان استخراج شود و تشخیص‌های سریع‌تر و دقیق‌تری صورت گیرد. به عنوان یک اصطلاح چتر گسترده، هوش مصنوعی شامل ML و DL است که روش‌های تخصصی‌تری برای به دست آوردن قابلیت‌های هوش مصنوعی هستند.

یادگیری ماشین بر الگوریتم‌های آماری و داده‌محور متکی است که می‌توانند از داده‌های گذشته یاد بگیرند و در مواجهه با داده‌های جدید تصمیم‌گیری کنند (Navaei, ۲۰۲۵). Safarzadeh, and Sobhani بر اساس داده‌ها، ML و DL سیستم‌ها را قادر می‌سازد تا تشخیص، پیش‌بینی، طبقه‌بندی و تصمیمات آگاهانه بگیرند (Azmi et al. ۲۰۲۲). ML در طول زمان با در دسترس قرار گرفتن داده‌های جدید با استفاده از روش‌های مبتنی بر داده برای یافتن الگوها، پیش‌بینی‌ها و تکمیل وظایف طبقه‌بندی، سازگار می‌شود. نقطه قوت ML سازگاری و ظرفیت آن برای کار با داده‌های ساختار یافته است (Chen et al. ۲۰۲۲).

شاخه‌های اصلی ML یادگیری تحت نظارت، بدون نظارت و تقویتی هستند و با استفاده از مدل‌ها و الگوریتم‌های آماری نتایج را پردازش، پیش‌بینی و طبقه‌بندی می‌کند (Azmi et al. ۲۰۲۲). به عنوان مثال، علی و همکاران (۲۰۲۱)، به بررسی الگوریتم‌های یادگیری ماشین نظارت شده در پیش‌بینی بیماری قلبی پرداخته‌اند و از نظر عملکرد و دقت مورد مقایسه قرار داده‌اند. نتایج آنها نشان می‌دهد که الگوریتم‌های k -نزدیک‌ترین همسایه (KNN)، درخت تصمیم (DT) و جنگل‌های تصادفی (RF)، بهترین عملکرد را در میان سایر الگوریتم‌های ML نظارت شده کسب کرده‌اند (Ali et al. ۲۰۲۱). در تحقیق دیگر، کرتاناونگ و همکاران (۲۰۲۰)، بر اساس یک فراتحلیل بر روی ده‌ها مطالعه در حوزه ی بکارگیری الگوریتم‌های نظارت‌شده در پیش‌بینی CVD نشان داده‌اند که الگوریتم‌های مبتنی بر تقویت گرادیانت (مانند XGBoost)، درخت تصمیم و SVM توانمندی بالایی در پیش‌بینی و تشخیص بیماری‌های قلبی از جمله CAD و تحلیل ریسک قلبی دارند (Krittanawong et al. ۲۰۲۰). همچنین انصاری و همکاران (۲۰۲۳)، در زمینه تشخیص سیگنال ECG و ریزکلاس‌بندی آریتمی‌ها مروری بر مطالعات مبتنی بر یادگیری نظارت‌شده و شبکه‌های عصبی عمیق (به‌ویژه مدل‌های کانولوشنی و MLP) بین سال‌های ۲۰۱۷ تا ۲۰۲۳ داشته‌اند. نتایج آنها گزارش‌کننده‌ی عملکرد برتر الگوریتم‌های نظارت‌شده در کنار شبکه‌های عصبی در تشخیص و طبقه‌بندی آریتمی و سایر نشانگرهای قلبی بوده‌اند (Ansari et al. ۲۰۲۳).

از سوی دیگر، روش‌های بدون نظارت برای فشرده‌سازی، خوشه‌بندی و فنوتایپینگ بیماران کاربرد یافته‌اند (Nucifora et al. ۲۰۲۵)؛ برای مثال، سی. جنتزر و همکاران

(۲۰۲۴)، با استفاده از الگوریتم‌های خوشه‌بندی و مدل‌های یادگیری نهفته مانند autoencoders زیرگروه‌های بالینی متمایزی در نارسایی قلبی و تصویربرداری قلبی شناسایی کرده‌اند که قابلیت ریسک‌سنجی و طبقه‌بندی مجدد بیماران را نشان می‌دهد (Jentzer et al. ۲۰۲۴). همچنین چائو و همکاران (۲۰۲۵)، گزارش داده‌اند که ترکیب تکنیک‌های یادگیری ماشین نظارت‌شده (مثل XGBoost / LightGBM) با فرایندهای انتخاب ویژگی و تنظیم فرآیندها بر اساس الگوریتم‌های نظارت نشده می‌تواند دقت پیش‌بینی حوادث قلبی را به‌طور معناداری بهبود دهد (Cao et al. ۲۰۲۵). در تحقیقی دیگر پریادارشی و همکاران (۲۰۲۴)، به بررسی نقش روش‌های یادگیری بدون نظارت در پیشبرد تحلیل تصویربرداری پرداخته‌اند. آنها الگوریتم‌های بدون نظارت، از جمله تکنیک‌های خوشه‌بندی، کاهش ابعاد و مدل‌های تولیدی، را بررسی و مورد تجزیه و تحلیل قرار داده‌اند. نتایج آنها نشان می‌دهد که الگوریتم‌های خوشه‌بندی مانند K-means، خوشه‌بندی سلسله مراتبی، DBSCAN، خوشه‌بندی طیفی، C-means فازی و خوشه‌بندی تغییر میانگین، و همچنین روش‌های کاهش ابعاد مانند PCA و t-SNE، و مدل‌های تولیدی مانند VAE‌ها و GAN‌ها، بیشترین استفاده را در کشف الگوهای پنهان و افزایش دقت تشخیصی تصویربرداری قلبی داشته‌اند (Priyadarshi et al. ۲۰۲۴). این شواهد نشان می‌دهد که هر دو رده‌ی روش‌ها — نظارت‌شده و بدون نظارت — نقش عملی و مکمل در پژوهش‌های CVD ایفا می‌کنند.

میشرا و همکاران (۲۰۲۵)، در تحقیق دیگر، یک مرور عمیق بر اساس روش بررسی مرور نظام مند (SLR) در زمینه بکارگیری تکنیک‌های یادگیری تقویتی (RL) برای تشخیص و کنترل بیماری‌های بحرانی ارائه داده‌اند. نتایج آنها نشان می‌دهد که تکنیک‌های تقویت پیشرفته شامل تقویت عمیق، تقویت چند عاملی، یادگیری تقویتی سلسله مراتبی، یادگیری تقویتی معکوس و یادگیری فدرال، بیشترین کاربردها را در بهبود تشخیص و کنترل بیماری‌ها از جمله بیماری‌های قلبی عروقی و قلبی داشته‌اند (Mishra, Shrivastava, and Bhushan ۲۰۲۵). در حوزه CVD‌ها، ML در تحلیل داده‌های ساختار یافته مانند سوابق بیماران، نتایج آزمایشگاهی و ویژگی‌های بالینی بسیار کاربرد دارد. برای مثال، مدل‌های ML در شناسایی بیماران در معرض خطر سکته قلبی یا پیش‌بینی پیامدهای درمانی بر اساس پارامترهای جمعیت‌شناختی و سبک زندگی عملکرد قابل توجهی داشته‌اند.

شکل دیگری از ML به نام "یادگیری عمیق" از شبکه های عصبی مصنوعی^۱ (ANNs) استفاده می کند که می تواند نمایش های سلسله مراتبی را یاد بگیرد و روابط داده پیچیده را بیان کند. در مدیریت CVD ها، DL به ویژه در تحلیل سیگنال های ECG، تصاویر اکوکاردیوگرافی، MRI و CT نقش کلیدی دارد؛ زیرا می تواند الگوهای پیچیده ای را که برای پزشک به صورت چشمی قابل تشخیص نیستند، استخراج کند. برای نمونه، شبکه های CNN در تشخیص آریتمی ها از سیگنال ECG یا در تفکیک بافت های آسیب دیده از تصاویر قلبی دقت بالایی نشان داده اند (Ansari et al. ۲۰۲۳).

با وجود این، مدل های منفرد ML یا DL ممکن است در مواجهه با نویز داده، ناهمگنی بیماران، یا کمبود داده های آموزشی دچار افت عملکرد شوند. در چنین شرایطی، یادگیری گروهی به عنوان رویکردی ترکیبی و قدرتمند مطرح می شود. مدل های EL یکی از متناسب ترین راه حل هاست که پیش بینی کنندگان مختلف را از چندین مدل جمع آوری می کند. در مقایسه با یک طبقه بندی واحد، رویکردهای گروهی از چندین الگوریتم (ML یا DL) برای تولید پیش بینی های دقیق تر استفاده می کنند. هدف اصلی مدل های گروهی کاهش خطاهای تعمیم است. هنگامی که مدل های پایه مستقل و متنوع هستند، این روش به طور موثری خطاهای پیش بینی را کاهش می دهد (Mahajan et al. ۲۰۲۳).

نکته ی مهم آن است که EL با مدل های ترکیبی (Hybrid ML-DL) تفاوت دارد. در EL، چندین مدل جداگانه و مستقل اجرا می شوند و نتایج آن ها در مرحله ی نهایی ادغام می شود؛ در حالی که در مدل های ترکیبی، معماری های ML و DL به صورت درونی یا زنجیروار با هم ترکیب می شوند تا یک مدل واحد با ساختار چندلایه ایجاد گردد. به عنوان مثال، در یک مدل ترکیبی ممکن است ویژگی ها ابتدا با شبکه ی CNN استخراج و سپس با الگوریتم جنگل تصادفی طبقه بندی شوند. اما در EL، CNN و جنگل تصادفی هر کدام خروجی مستقل دارند و سپس نتایج آن ها با رأی گیری یا استکینگ ادغام می شود.

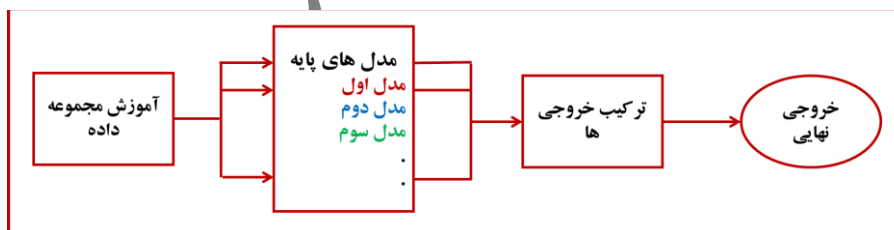
تحت یک قانون تجمیع خاص، رویکردهای گروهی چندین الگوریتم تک طبقه بندی کننده را ترکیب می کنند. هر یک از اعضای گروه به طور مستقل آموزش دیده اند تا همان کار را حل کنند و راه حلی برای مشکل کلی ارائه دهند. سپس خروجی نهایی گروه با استفاده از یک قانون تجمیع تولید می شود که خروجی های مدل های جداگانه را ترکیب می کند. همانطور که در شکل یک نشان داده شده است، این رویکرد برای پیش بینی به

^۱. Artificial Neural Networks (ANNs)

خروجی جمعی مدل های فردی متکی است. اگرچه یک مجموعه از چندین مدل پایه تشکیل شده است، اما به عنوان یک مدل واحد عمل و رفتار می کند (Hosni et al. ۲۰۲۱; Mahajan et al. ۲۰۲۳). این تکنیک به ویژه در مراقبت های بهداشتی ارزشمند است، جایی که مجموعه داده های متنوع و معیارهای تشخیصی نیاز به عملکرد مدل قابل اعتماد دارند. به عنوان مثال، این ساختار نشان می دهد که داده های آموزشی وارد مجموعه ای از مدل های پایه (مانند درخت تصمیم، شبکه عصبی یا SVM) می شوند، خروجی آن ها ترکیب شده و نتیجه ی نهایی تولید می شود. در کاربردهای قلبی عروقی، همین ساختار با داده های خاص مانند سیگنال های ECG، تصاویر CT یا داده های بالینی به کار می رود تا مدل نهایی با دقت و پایداری بالاتری عمل کند.

روش های گروهی، مانند بگینگ، بوستینگ یا تقویتی، انباشته و رای گیری، چندین مدل را برای افزایش دقت و استحکام پیش بینی ترکیب می کنند (Alotaibi ۲۰۲۵; Hosni et al. ۲۰۲۴; Rane, Choudhary, and Rane ۲۰۲۱). در کاربردهای قلبی عروقی، EL می تواند داده های چندمنبعی (مانند ECG، داده های آزمایشگاهی و سوابق پزشکی) را ادغام کرده و از ترکیب نقاط قوت مدل های مختلف برای کاهش خطاهای تشخیص و افزایش قابلیت تعمیم استفاده کند. به عنوان مثال، ترکیب شبکه های عصبی با درخت های تصمیم در چارچوب EL منجر به بهبود عملکرد در پیش بینی حملات قلبی شده است.

شکل ۱. رویکرد ساختاری یادگیری گروهی



مرور مطالعات گذشته

این بخش یک نمای کلی از مطالعات قبلی را ارائه می دهد. چندین محقق پیمایش ها و بررسی هایی را در مورد کاربردهای EL و ML در بیماری های قلبی عروقی انجام داده اند

و هدف آنها ارائه راهنمایی عملی در مورد استفاده از این تکنیک ها و مدل ها در زمینه قلب و عروق و نقشه راه برای تحقیقات آینده است. جدول I خلاصه ای از آنها را ارائه می دهد. حسنی و همکاران (۲۰۲۱)، یک رویکرد مطالعه نقشه برداری سیستماتیک ارائه می دهند که کاربرد تکنیک های طبقه بندی گروهی را در بیماری های قلبی عروقی بررسی می کند. ابتدا نویسندگان جستجوی کلیدواژه ها را در چهار کتابخانه ACM Digital، IEEE Xplore، PubMed، Library و Scopus انجام دادند و ۳۵۱ مقاله را بازیابی کردند که بر اساس ۹ سؤال تحقیق مورد تجزیه و تحلیل قرار گرفتند. این مطالعه نشان داد که در اکثر مطالعات CVD، گروه های همگن رایج ترین روش های گروهی هستند که مدل هایی مانند درختان تصمیم^۱ (DTs)، ANN ها و طبقه بندی کننده های بیزی رایج ترین تکنیک های منفرد هستند. بیشترین نقش مدل ها در حیطه تشخیصی بود و روش اعتبارسنجی متقابل^۲ k-fold غالب ترین تکنیک برای ارزیابی مدل ها بود. همچنین نتایج نشان می دهد که مطالعات قبلی به طور جامع از ابزارهای متن باز مانند Weka، Python، R Software برای پردازش مدل ها استفاده کرده اند و برخی مطالعات نیز بهینه سازی پارامترهای طبقه بندی کننده را بررسی کرده اند. این مطالعه بر عملکرد برتر روش های گروهی نسبت به طبقه بندی کننده های تکی تأکید می کند (Hosni et al. ۲۰۲۱).

کریتاناونگ و همکاران (۲۰۲۰)، با توجه به استفاده روزافزون از الگوریتم های ML در پیش بینی CVD، یک فرا تحلیل را بر اساس داده های ۱۰۳ مطالعه و ۳،۳۷۷،۳۱۸ نفر ارزیابی کردند. الگوریتم های تقویتی و مدل های سفارشی عملکرد قوی در پیش بینی بیماری CAD داشته اند. برای پیش بینی سگته مغزی، مدل ماشین های بردار پشتیبان^۳ (SVMs)، الگوریتم های تقویتی و شبکه های عصبی پیچشی^۴ (CNNs) به AUC های تجمعی بالای ۰/۹۰ درصد دست یافته اند. این مطالعه نشان داد که مدل های تقویتی و SVM عملکرد پیش بینی امیدوار کننده ای را نشان می دهند (Krittana Wong et al. ۲۰۲۰).

عزمی و همکاران (۲۰۲۲)، با ارزیابی ۴۱ مقاله، تجزیه و تحلیل مقایسه ای جامعی از تکنیک های ML برای پیش بینی CVD انجام داده اند. چالش های کلیدی در ادبیات موجود شامل محدودیت در اندازه مجموعه داده، توزیع ناسازگار داده ها، عدم همبستگی

^۱. Decision Trees (DTs)

^۲. Cross-validation

^۳. Support vector machines (SVMs)

^۴. Convolutional Neural Network (CNNs)

بالینی، روش های ناکافی انتخاب ویژگی، تنظیم ناکافی هایپرپارامتر، استفاده محدود از تکنیک های فراابتکاری و گروهی، و شکاف در ادغام بالینی است. چندین مدل ML - از جمله DT، RF، بیز ساده^۱ (NB)، رگرسیون لجستیک^۲ (LR)، ANN، K نزدیکترین همسایه^۳ (KNN) و SVM - از نظر دقت پیش بینی مورد بررسی قرار گرفته اند. در این میان، RF که یک الگوریتم گروهی است بالاترین دقت را نشان داده و پس از آن SVM قرار دارد (Azmi et al. ۲۰۲۲).

تاکور و همکاران (۲۰۲۳)، نقش روش های گروهی را در تشخیص قلب و عروق برجسته می کنند و بر توانایی آنها در افزایش قابلیت اطمینان و کاهش تشخیص های نادرست تأکید می کنند. این مطالعه مدل های ML و DL را بر اساس کاربردهای تشخیصی آنها طبقه بندی می کند و نشان می دهد که چگونه ادغام چندین مدل دقت پیش بینی را بهبود می بخشد و نرخ خطا را کاهش می دهد. این بر مزیت نسبی رویکردهای گروهی برای دستیابی به نتایج قابل اعتماد در دستیابی به دقت تشخیصی بالا و کاهش خطاها در وظایف پیش بینی تأکید می کند (Thakur et al. ۲۰۲۳).

در سال های اخیر، دو رویکرد محاسباتی نوپهور یعنی یادگیری انتقالی^۴ (TL) و یادگیری فدرال^۵ (FL)، به صورت فزاینده ای در پژوهش های مرتبط با بیماری های قلبی عروقی مورد توجه قرار گرفته اند. یادگیری انتقالی اجازه می دهد مدل های پیش آموزش دیده (مثلاً روی مجموعه داده های بزرگ تصویری یا سیگنال) برای تسریع همگرایی و بهبود عملکرد در مجموعه داده های کوچک تر یا خاص حوزه ی قلبی مورد «فاین-تیون» قرار گیرند؛ این رویکرد به ویژه در تحلیل تصاویر پزشکی و سیگنال های ECG که داده های برچسب خورده محدود دارند مفید است، و مطالعات اخیر نشان می دهند که استفاده از چارچوب های انتقالی چندمنظوره می تواند دقت تشخیصی را در پیش بینی CVD به صورت قابل توجهی افزایش دهد (Karthikeyan and others ۲۰۲۴).

از سوی دیگر، یادگیری فدرال به عنوان راه حلی برای چالش های حریم خصوصی و محدودیت اشتراک گذاری داده های بیمارستانی پدیدار شده است؛ در FL مدل ها به صورت

^۱. Naive Bayes (NB)

^۲. Logistic regression (LR)

^۳. K-Nearest Neighbors (KNN)

^۴. Transfer Learning (TL)

^۵. Federated Learning (FL)

توزیع شده بر روی داده‌های محلی در مراکز مختلف آموزش داده می‌شوند و تنها پارامترها یا وزن‌ها به صورت مرکزی ادغام می‌شوند، بدون آنکه داده‌های خام بیمار تبادل شوند. این رویکرد برای مطالعات قلب و عروق بسیار مناسب است زیرا امکان استفاده از داده‌های چندمرکزی و متنوع را بدون نقض محرمانگی فراهم می‌آورد و مطالعات کاربردی اخیر نشان داده‌اند که FL می‌تواند برای پیش‌بینی آریتمی (مثلاً AF) از سیگنال ECG با حفظ حریم خصوصی داده‌ها عملکرد رقابتی ارائه دهد (Rahman, Karmarkar, and Islam ۲۰۲۴).

در تحقیقی دیگر، چورنی و هوی لینگ (۲۰۲۵) یک مجموعه داده توزیع شده غیرمستقل و یکسان برای تشخیص فیبریلاسیون دهلیزی ایجاد کرده‌اند. در تحقیق آنها دو خط پایه (یک طبقه‌بندی‌کننده متمرکز و یک طبقه‌بندی‌کننده فدرال) آموزش داده شده‌اند و عملکرد طبقه‌بندی‌کننده‌هایی که قبل از آموزش فدرال روی داده‌های مشترک از قبل آموزش دیده‌اند، مقایسه شده‌اند. نتایج آنها نشان دهنده این است که طبقه‌بندی‌کننده متمرکز از همه مدل‌های دیگر بهتر عمل کرده و مدل فدرال عملکرد مناسبی را کسب نکرده است. یافته‌های کلی این تحقیق نشان می‌دهد که اشتراک‌گذاری جزئی داده‌ها در ایجاد مدل‌های یادگیری ماشین فدرال ممکن است عملکرد را به طور قابل توجهی بهبود بخشد (Chorney and Ling ۲۰۲۵).

در عمل، ترکیب TL و FL (مثلاً فاین-تیون کردن مدل‌های پیش‌آموزش دیده به صورت محلی در هر مرکز و ادغام پارامترها به صورت فدرال) یک جهت پژوهشی بسیار امیدوارکننده برای افزایش تعمیم‌پذیری مدل‌ها و هم‌زمان حفظ حریم خصوصی و قابلیت همکاری میان بیمارستان‌ها فراهم کرده است (Alreshidi et al. ۲۰۲۴).

روش شناسی تحقیق

تکنیک‌ها، الگوریتم‌ها و مدل‌های متعددی برای تشخیص، پیش‌بینی و طبقه‌بندی بیماری‌های قلبی عروقی به کار گرفته شده‌اند. این مدل‌ها در درجه اول بر اساس اصول EL و ML/DL هستند. از آنجا که روش فراترکیب یک رویکرد کیفی است که به طور هم‌زمان یافته‌های کمی و کیفی را بررسی و تلفیق می‌کند، می‌تواند به عنوان ابزاری مؤثر برای شناسایی و تحلیل این مدل‌ها و تکنیک‌ها عمل کند. در این مطالعه از روش فراترکیب برای استخراج و تحلیل مدل‌ها و تکنیک‌های EL در حوزه بیماری‌های قلبی عروقی استفاده

شده است. این پژوهش از نظر ماهیت کاربردی بوده و از نظر گردآوری داده ها در هر دو دسته مطالعات کیفی و کمی قرار می گیرد. برای شناسایی و استخراج تکنیک ها و مدل های مختلف EL مورد استفاده در مدیریت CVD ها، رویکرد فراترکیب هفت مرحله ای پیشنهاد شده توسط سندلوسکی و باروسو^۱ (۲۰۰۷) (Sandelowski, Barroso, and Voils) (۲۰۰۷) به عنوان چارچوب تحقیق پذیرفته شده است. مراحل این روش در شکل دوم نشان داده شده است.

شکل ۲. الگوی ۷ مرحله ای رویکرد فراترکیب سندلوسکی و باروسو (۲۰۰۷)



اولین گام در رویکرد فراترکیب شامل تعریف سؤالات پژوهشی است که هدف مطالعه پرداختن به آنها است. سؤالات پژوهش که در بخش مقدمه مطرح شده و مبنای این رویکرد است، به شرح زیر است:

۱. مدل ها و تکنیک های EL بکار گرفته شده در مدیریت بیماری های قلبی عروقی در چند دسته تقسیم بندی می شوند؟
۲. کدام مدل های EL بیشترین دقت و عملکرد را در هر دسته از مدیریت CVD ها داشته اند؟

۳. محدودیت ها و چالش های مربوط به مدل های EL در مدیریت CVD ها چیست؟

۴. چگونه مدل ها و تکنیک های EL در مدیریت CVD ها اعتبارسنجی می شوند؟

۵. برای استفاده از مدل های در مدیریت ها داده ها و ویژگی های یک مجموعه داده

با چه تکنیک ها و روش هایی پیش پردازش و مهندسی می شوند؟

در گام دوم رویکرد فراترکیب، پایگاه وب آو ساینس به عنوان منبع اصلی برای مرور نظام مند ادبیات علمی انتخاب شد. سه کلمه کلیدی اصلی که برای انجام جستجوها استفاده شدند شامل «Cardiovascular Diseases»، «Machine Learning» و «Deep Learning» از سال ۲۰۱۰ تا پایان سال ۲۰۲۴ بودند. دلیل استفاده از اصطلاحات "یادگیری

^۱. Sandelowski and Barroso (۲۰۰۷)

ماشین" و "یادگیری عمیق" به جای "یادگیری گروهی" بر اساس اهداف انجام این مطالعه بود. اصطلاحات جستجو به صورت زیر فرموله شدند:

- TS= ("Cardiovascular Diseases" AND "Machine Learning")
- TS= ("Cardiovascular Diseases" AND "Deep Learning")

در مرحله سوم، فرآیند غربالگری و انتخاب مقالات مرتبط انجام شد. انتخاب مقالات بر اساس معیارهای دقیق و از پیش تعریف شده انجام شده است که جزئیات آن در جدول یک ارائه شده است. سپس مقالات استخراج شده در پنج مرحله ارزیابی عنوان، بررسی چکیده، بررسی دسترسی، ارتباط محتوا و در نهایت ارزیابی کیفیت روش شناختی مورد ارزیابی قرار گرفتند. این فرآیند تضمین می کند که در نهایت فقط مطالعات معتبر و مرتبط در تحقیق گنجانده می شوند.

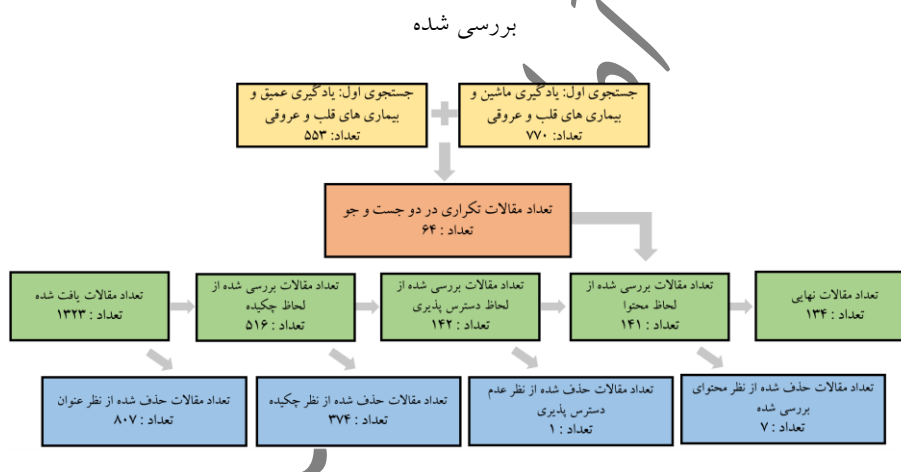
جدول ۱. معیارهای ورود مقالات

معیارهای خروج	شرایط ورود	معیارهای
مقالاتی که به زبان انگلیسی نوشته نشده اند	مقالات نوشته شده به زبان انگلیسی	زبان مقالات
مقالات منتشر شده قبل از سال ۲۰۱۰	مقالات منتشر شده بین سال های ۲۰۱۰ تا ۲۰۲۴	بازه زمانی انتشار
مطالعات غیرمرتبط با موضوع مشخص شده	پیش بینی و تشخیص بیماری های قلبی عروقی بر اساس EL	موضوع تحقیق
مطالعات خارج از محدوده مشخص شده	شناسایی مدل ها، تکنیک ها و الگوریتم های EL به کار رفته در تشخیص، پیش بینی، طبقه بندی و شناسایی بیماری های قلبی عروقی.	دامنه تحقیق

در طول فرآیند غربالگری مقاله، ۷۷۰ مقاله از پرس و جو اول و ۵۵۳ مقاله از پرس و جو دوم بازیابی شد که ۶۴ مقاله بین این دو همپوشانی داشتند. برای شناسایی مقالات تکراری، عناوین هر دو جستجو در ابتدا به طور جداگانه بررسی شدند. قبل از رفتن به مرحله دوم (بررسی چکیده)، مقالات تکراری و همپوشانی شناسایی و حذف شدند. در این مرحله ۵۱۶ مقاله بر اساس عناوین آنها مرتبط تلقی شدند. مرحله دوم که شامل بررسی چکیده مقالات بود، منجر به انتخاب ۱۴۲ مقاله شد. در مرحله سوم، در مورد دسترسی به مقالات، ۱۴۱ مقاله در دسترس و به طور کامل استخراج شد. سپس این مقالات از نظر ارتباط محتوایی مورد بررسی قرار گرفتند و در نهایت ۱۳۴ مقاله بر اساس اهداف و سؤالات پژوهش در مرحله چهارم انتخاب شدند. برای بررسی مرحله نهایی - ارزیابی کیفیت روش شناختی - جهت

افزایش دقت و اطمینان از کیفیت روش‌شناختی مقالات نهایی، ارزیابی کیفیت بر اساس دستورالعمل و چک‌لیست رسمی JBI انجام شد. این چک‌لیست شامل معیارهایی چون وضوح سؤال پژوهش، کفایت روش نمونه‌گیری، شفافیت طراحی مطالعه، تحلیل داده‌ها و کنترل سوگیری است. تنها مقالاتی که امتیاز کافی بر اساس معیارهای JBI کسب کردند در تحلیل نهایی باقی ماندند. استفاده از این ابزار، که یکی از پرکاربردترین و معتبرترین ابزارهای ارزیابی کیفیت در مطالعات مروری نظام‌مند است، موجب افزایش اطمینان به اعتبار و تکرارپذیری یافته‌ها گردید. فرآیند غربالگری پنج مرحله‌ای در شکل سوم نشان داده شده است.

شکل ۳. فرآیند غربالگری مقالات استخراج شده از پایگاه داده وب آو ساینس از نظر پنج مرحله



جدول دوم مراحل هفت‌گانه روش فراترکیب را همراه با تعریف مختصر هر مرحله و نحوه اجرای آن در این پژوهش نشان می‌دهد.

جدول ۲. مراحل هفت‌گانه فراترکیب سندلوسکی و باروسو (۲۰۰۷) و نحوه به‌کارگیری در این پژوهش

مراحل فراترکیب	توضیح مفهومی	نحوه اجرا در پژوهش حاضر (کاربرد در CVD و EL)
۱. شناسایی هدف و سؤال پژوهش	تعیین دقیق هدف و سؤالات پژوهش برای تمرکز در فرایند فراترکیب	سؤالات اصلی درباره شناسایی، طبقه‌بندی، ارزیابی و چالش‌های مدل‌های EL در مدیریت CVD ها در ابتدای پژوهش تعریف شد.

۲. جستجوی نظام‌مند مطالعات	انتخاب پایگاه‌ها و کلیدواژه‌ها برای گردآوری شواهد مرتبط	پایگاه Web of Science با کلیدواژه‌های “Cardiovascular Diseases” AND “Machine Learning” برای بازه ۲۰۱۰-۲۰۲۴ استفاده شد.
۳. غربالگری و انتخاب مطالعات	تعیین معیارهای ورود و خروج و حذف موارد تکراری یا غیرمرتبط	معیارهای ورود و خروج در جدول ۱ آورده شد. ۱۳۴ مقاله مرتبط پس از غربالگری پنج مرحله‌ای انتخاب شدند.
۴. استخراج داده‌ها و یافته‌ها	شناسایی و استخراج مفاهیم کلیدی از مطالعات منتخب	از هر مقاله نوع مدل EL، الگوریتم‌ها، داده‌های ورودی (ECG، CT، داده‌های بالینی)، هدف و نتایج عملکرد استخراج شد.
۵. مقایسه و طبقه‌بندی یافته‌ها	مقایسه مفاهیم مشترک و ایجاد طبقات معنایی	مدل‌ها و تکنیک‌های EL در چهار گروه اصلی (تشخیص، پیش‌بینی، طبقه‌بندی و شناسایی) طبقه‌بندی شدند.
۶. ترکیب و یکپارچه‌سازی مفاهیم	تلفیق یافته‌ها و تولید مفاهیم یا چارچوب‌های جدید	یافته‌ها از مطالعات کمی و کیفی ترکیب شدند تا الگوی مفهومی ارتباط میان EL، ML/DL و کاربردهای CVD استخراج شود.
۷. ارائه و تفسیر نتایج نهایی	تبیین چارچوب نهایی و ارائه پیامدهای علمی و عملی	چارچوب مفهومی حاصل، نشان‌دهنده نقش مدل‌های EL در بهبود دقت و نتایج در مدیریت CVD است.

در نتیجه، ۱۳۴ مقاله که معیارهای ارتباط محتوایی را داشتند، به عنوان انتخاب نهایی در نظر گرفته شدند. یافته‌های فرایند فراترکیب نشان داد که انواع مدل‌ها و تکنیک‌های EL در کنار رویکردهای ML و DL در مدیریت بیماری‌های قلبی عروقی به کار گرفته شده است. این مدل‌ها و تکنیک‌ها در ۱۳۴ مقاله منتخب شناسایی شدند. در مرحله‌ی تحلیل داده‌ها، کلیه‌ی اطلاعات استخراج شده از ۱۳۴ مقاله‌ی منتخب شامل نوع بیماری قلبی عروقی، منبع و نوع داده‌ها، نام و نوع مدل‌های EL، روش‌های پیش‌پردازش داده و مهندسی ویژگی، معیارهای ارزیابی عملکرد و محدودیت‌های مطالعات در یک فایل اکسل ذخیره و کدگذاری شدند.

سپس با استفاده از روش تحلیل محتوا، داده‌های کیفی، کمی و مفهومی مورد بررسی قرار گرفتند. در طی این فرآیند، الگوهای مفهومی تکرارشونده میان مقالات شناسایی و بر اساس هدف و ماهیت هر مطالعه در قالب چهار حوزه‌ی اصلی مدیریت CVD شامل

پیش‌بینی^۱، تشخیص^۲، طبقه‌بندی^۳ و شناسایی^۴ سازمان‌دهی و دسته‌بندی شدند. این دسته‌بندی در مرحله‌ی نهایی فراترکیب و بر اساس کدگذاری محوری (روش تحلیل محتوا با استفاده از شمارش فراوانی) انجام گرفت تا الگوهای غالب در کاربرد مدل‌های EL در حوزه‌ی CVD شناسایی شود. نتایج این فراترکیب و دسته‌بندی آنها در جدول سوم ارائه شده است.

یافته‌های فراترکیب

با تمرکز بر هدف اصلی مطالعه، یعنی بررسی نقش الگوریتم‌های یادگیری گروهی در مدیریت CVDها، یافته‌ها نشان دادند که مدل‌های EL مانند جنگل تصادفی، AdaBoost، Gradient Boosting، XGBoost و Stacking Ensemble در بسیاری از مطالعات به‌صورت مستقل یا در ترکیب با مدل‌های ML و DL به‌کار رفته‌اند. این مدل‌ها به دلیل قابلیت ادغام تصمیم‌گیری چند الگوریتم، کاهش واریانس، و مقابله با داده‌های نویزی و نامتوازن، در مقایسه با مدل‌های منفرد عملکرد دقیق‌تر و پایداری بیشتری ارائه داده‌اند. به طور خاص، در حوزه‌های پیش‌بینی و تشخیص بیماری‌های قلبی، مدل‌های XGBoost و جنگل تصادفی بیشترین فراوانی و بالاترین دقت را ثبت کرده‌اند. این یافته‌ها تأیید می‌کنند که رویکرد EL نه تنها به‌عنوان یک روش مکمل، بلکه به‌عنوان یکی از چارچوب‌های اصلی در طراحی سیستم‌های هوشمند تشخیص و پیش‌بینی CVD مطرح است.

جدول ۳. دسته‌بندی مدیریت بیماری‌های قلبی عروقی مبتنی بر مدل‌های یادگیری بر اساس یافته‌های روش فراترکیب.

دسته‌بندی مدیریت CVD ها	زیر مجموعه ها	مراجع
پیش‌بینی (۷۵ مقاله)	پیش‌بینی بیماری‌های قلبی عروقی با استفاده از مدل‌های EL	(Ahamed, Koli, et al. ۲۰۲۲; Ahamed, Mir, and Chishti ۲۰۲۲; Al-Yarimi ۲۰۲۲; Alfaidi et al. ۲۰۲۲; Anbumani and Kaniyaiah ۲۰۲۴; Bertsimas et al. ۲۰۲۲; Bhatt et al. ۲۰۲۳; Datta et al. ۲۰۲۲; Dinh et al. ۲۰۱۹; Dritsas and Trigka ۲۰۲۳; Dubey, Sinhal, and Sharma ۲۰۲۲; Ghosh et al. ۲۰۲۱; Guo et al. ۲۰۲۰; Hassan et al. ۲۰۲۴; Henni et al. ۲۰۲۳; Husain et al. ۲۰۲۳; Islam et al. ۲۰۲۲; Kim et al. ۲۰۲۱; Lacson et al. ۲۰۱۹; Lambay and Mohideen ۲۰۲۴; Lee et al. n.d.; Li et al. ۲۰۲۱; Lo, Fujita, and Pai ۲۰۱۶; Lv et al. ۲۰۲۱; Muhammad et al. ۲۰۲۳; Najim and Nasri ۲۰۲۴; Nissa et al.

^۱. Prediction

^۲. Diagnosis

^۳. Classification

^۴. Identification

۲۰۲۴; Noor et al. ۲۰۲۳; Ohashi et al. ۲۰۲۳; Ospina et al. ۲۰۲۳; Ott et al. ۲۰۲۴; Pan et al. ۲۰۲۲; Rahim et al. ۲۰۲۱; Sajid et al. ۲۰۲۱, ۲۰۲۲; Samaras et al. ۲۰۲۳; Saputra et al. ۲۰۲۳; Sheeba et al. ۲۰۲۲; Sherazi, Zheng, and Lee ۲۰۲۳; Sinha et al. ۲۰۲۳; Stepanyan et al. ۲۰۲۲; Subramani et al. ۲۰۲۳; Tachie et al. ۲۰۲۳; Tiwari et al. ۲۰۲۲; Tolentino et al. ۲۰۲۰; Torres et al. ۲۰۲۲; Tsai et al. ۲۰۲۲; Varale and Thakre ۲۰۲۰)		
An et al. ۲۰۲۱; Barriada et al. ۲۰۲۲; Bertsimas et al. ۲۰۲۲;) Chun et al. ۲۰۲۱; Du et al. ۲۰۲۳; Hossain, Uddin, and Khan ۲۰۲۱; Islam et al. ۲۰۲۲, ۲۰۲۳; Lee et al. ۲۰۲۴; Mondal et al. ۲۰۲۴; Oanh and Tung ۲۰۲۲; B. A. S. Oliveira et al. ۲۰۲۳; et al. ۲۰۲۲; Reddy et al. ۲۰۲۱; Sozen et al. Paul Navarrete (۲۰۲۳; Teresa Garcia-Ordas et al. ۲۰۲۳; You et al. ۲۰۲۳	افزایش پیش بینی خطر CVD	
Ahn et al. ۲۰۲۱; Daghistani et al. ۲۰۱۹; Hu et al. ۲۰۲۰;) (Karboub and Tabaa ۲۰۲۲; M. Oliveira et al. ۲۰۲۳	بهینه سازی منابع بیمارستان و مدیریت بیمار	
Chicco and Jurman ۲۰۲۰; Hassan et al. ۲۰۲۴; Moreno-) (Sanchez ۲۰۲۳; Tadege, Tegegne, and Dessie ۲۰۲۴	پیش بینی پیامدهای بقا برای بیماران CVD	
Baghdadi et al. ۲۰۲۳; Chalapathi et al. ۲۰۲۴; Dhanka,) Bhardwaj, and Maini ۲۰۲۳; Elvas et al. ۲۰۲۳; Husain et al. ۲۰۲۳; Ibarra et al. ۲۰۲۲; Jiang et al. ۲۰۲۳; Kokol et al. ۲۰۲۱; Moskaleva et al. ۲۰۲۲; Najafi et al. ۲۰۲۳; Ogunpola et al. (۲۰۲۴; Oyewola, Dada, and Misra ۲۰۲۴; Sapra et al. ۲۰۲۳	تشخیص و تشخیص زودهنگام بیماری های قلبی عروقی با استفاده از مدل های EL	تشخیص (۳۰ مقاله)
Aarthy and Iqbal ۲۰۲۴; Al-Absi et al. ۲۰۲۲; Ardeti et al.) ۲۰۲۲; Balaha et al. ۲۰۲۲; Divya, Shadrach, and Padmaja ۲۰۲۳; Latha et al. ۲۰۲۲; Lilda, Jayaparvathy, and Balaji ۲۰۲۲; Pandey et al. ۲۰۲۳; Ribeiro et al. ۲۰۲۴; Sharma et al. ۲۰۲۰; Shchetinin and Glushkova ۲۰۲۲; Simegn, Gebeyehu, (and Degu ۲۰۲۲; Yang et al. ۲۰۱۸; Zafar et al. ۲۰۲۳	تشخیص و تشخیص بیماری های قلبی عروقی با استفاده از PCG, PPG, ECG و تصویربرداری	
Boulkaboul et al. ۲۰۲۴; Hassaballah et al. ۲۰۲۳; He et al.) ۲۰۲۱; Kumar and Kumar ۲۰۲۳; Pavic et al. ۲۰۲۳; Xiang et (al. ۲۰۲۳; Yongcharoenchaiyasit et al. ۲۰۲۳	طبقه بندی CVD ها و تشخیص زودهنگام با استفاده از مدل های EL	طبقه بندی (۱۸ مقاله)

Eleyan and Alboghbaish ۲۰۲۴; Enad and Mohammed) ۲۰۲۴; Essa and Xie ۲۰۲۱; Lee et al. ۲۰۲۱; Maheshwari et al. ۲۰۲۳; Wu et al. ۲۰۲۱; Yadav et al. ۲۰۲۰; Zhanquan et al. (۲۰۲۲; Zou et al. ۲۰۲۲	طبقه بندی ECG و سیگنال های قلبی با استفاده از مدل های EL و DL	
Guarneros-Nolasco et al. ۲۰۲۱; Gutierrez-Esparza et al.) ۲۰۲۳; Hsu ۲۰۱۸; Kirboga and Kucuksille ۲۰۲۳; Kumar et al. ۲۰۲۱; Louridi, Douzi, and El Ouahidi ۲۰۲۱; Tadesse et al. (۲۰۲۱; Wu et al. ۲۰۲۲	شناسایی عوامل خطر بیماری های قلبی عروق	شناسایی (۹ مقاله)
(Desai et al. ۲۰۲۲; Li et al. ۲۰۲۰; Maqsood et al. ۲۰۲۱)		دسته بندی دیگر (۳ مقاله)

زیر بخش های زیر تجزیه و تحلیل جامعی از نحوه استفاده از مدل های EL و ML/DL برای مدیریت CVD ها ارائه می دهند. همانطور که نتایج در جدول سوم در بالا نشان می دهد، نقش این مدل ها در مدیریت بیماری های قلبی عروقی به چهار حوزه اصلی دسته بندی می شوند: پیش بینی، تشخیص، طبقه بندی و شناسایی. هر دسته از زیر مجموعه های خاصی تشکیل شده است که کاربردهای متنوع و تعداد مطالعات متمرکز بر هر جنبه را برجسته می کند.

الف) مدیریت بیماری های قلبی عروقی

پیش بینی بیماری های قلبی عروقی: پیش بینی نقش محوری در مدیریت پیشگیرانه بیماری های قلبی عروقی ایفا می کند و ۷۵ مقاله به این دسته اختصاص داده شده است. در این میان، تمرکز قابل توجهی بر پیش بینی CVD با استفاده از مدل های EL است، جایی که روش های گروهی مانند تقویت گرادیان شدید^۱ (XGBoost)، RF و تقویت کنندگی وفقی^۲ (AdaBoost) دقت پیش بینی را افزایش می دهند. افزایش پیش بینی خطر بیماری های قلبی عروقی یکی دیگر از زیرمجموعه های حیاتی است که هدف آن شناسایی افراد در معرض خطر بالا قبل از شروع شرایط شدید است. علاوه بر این، بهینه سازی منابع بیمارستانی و مدیریت بیمار از طریق مدل های پیش بینی برای بهبود ارائه مراقبت های بهداشتی، کاهش هزینه ها و تخصیص کارآمد منابع پزشکی مورد توجه قرار می گیرد. یکی دیگر از حوزه های کلیدی، پیش بینی نتایج بقا برای بیماران CVD است، جایی که مدل ها

^۱. eXtreme Gradient Boosting (XGBoost)

^۲. Adaptive Boosting (AdaBoost)

پیش آگهی بیمار را پیش بینی می کنند و به برنامه ریزی درمان شخصی کمک می کنند و در نتیجه مراقبت طولانی مدت از بیمار را بهبود می بخشد

تشخیص بیماری های قلبی عروقی: تشخیص به موقع و دقیق برای مدیریت موثر بیماری های قلبی عروقی ضروری است و ۳۰ مقاله بر این دسته متمرکز شده است. تشخیص و تشخیص زودهنگام CVD با استفاده از مدل های EL برای شناسایی علائم اولیه و اطمینان از مداخلات سریع تمرکز اصلی مقالات این حوزه بوده اند. روش های مبتنی بر سیگنال های ECG، فوتوپلتیسموگرافی^۱ (PPG)، صوتی قلب^۲ (PCG) و تصویربرداری به استاندارد برای تشخیص ناهنجاری های ساختاری و عملکردی قلب تبدیل شده اند. استفاده از مدل های EL همراه با پردازش سیگنال زیست پزشکی، دقت بالاتر در تشخیص را بدون احتمال تشخیص اشتباه تضمین کرده است.

طبقه بندی بیماری های قلبی و عروقی: طبقه بندی دقیق CVD ها برای تعیین مسیرهای درمانی مناسب بسیار مهم است، همانطور که در ۱۸ مقاله منعکس شده است. زیر طبقات شامل طبقه بندی بیماری های قلبی عروقی و تشخیص زودهنگام با استفاده از مدل های EL است که شرایط مختلف قلبی عروقی را با دقت بالا طبقه بندی می کند. علاوه بر این، کاربرد مدل های DL و EL در طبقه بندی ECG و سیگنال های قلبی در حال افزایش است و بهترین استراتژی را برای شناسایی و تجزیه و تحلیل الگوهای پیچیده سیگنال های زیست پزشکی به دست آورده است. روش های طبقه بندی برای تمایز بین بیماری های قلبی مانند آریتمی و بیماری های ایسکمیک قلبی ضروری هستند تا امکان درمان های انتخابی تر فراهم شود.

شناسایی بیماری های قلبی عروقی: شناسایی عوامل خطر مرتبط با بیماری های قلبی عروقی برای پیشگیری و استراتژی های درمان ضروری است. این دسته که در ۹ مقاله پوشش داده شده است، بر شناسایی عوامل خطر CVD با استفاده از مدل های پیشرفته ML و EL تأکید دارد. با شناسایی عوامل اصلی تعیین کننده (عوامل خطر) - مانند فشار خون بالا، دیابت و غیره - این مدل ها به پزشکان CVD در شروع استراتژی های پیشگیرانه و تجویز تغییرات سبک زندگی برای افراد در معرض خطر کمک می کنند.

^۱. Photoplethysmogram (PPG)

^۲. Phonocardiogram (PCG)

ب) مدل های EL، ML و DL در مدیریت CVD ها

این زیربخش یک نمای کلی از مدل ها، روش ها و الگوریتم هایی را که معمولاً در مطالعات مختلف برای مدیریت CVD ها استفاده شده اند، با تأکید ویژه بر روش های EL، تکنیک های کلاسیک ML و معماری های DL ارائه می دهد. جدول سوم تمام این مدل ها و تکنیک های مورد استفاده در مقالات را به همراه فراوانی آنها نشان می دهد.

مدل های یادگیری گروهی (EL): تجزیه و تحلیل یافته ها نشان می دهد

که مدل های EL یکی از پرکاربردترین رویکردها هستند. در بین روش های تقویتی، XGBoost بیشترین (۵۰ بار) و پس از آن تقویت گرادیان^۱ (GB) (۳۵ بار) و AdaBoost (۲۷ بار) قرار دارند. ماشین تقویت کننده گرادیان سبک^۲ (LightGBM) (۱۲ بار) و تقویت گرادیان دسته بندی شده^۳ (CatBoost) (۹ بار) مدل های جدیدی هستند که محبوبیت فزاینده ای دارند که نشان دهنده علاقه بیشتر به روش های تقویتی کارآمد و مقیاس پذیر است. انواع مدل های تقویتی کمتر رایج، از جمله درخت افزایش گرادیان^۴ (۵ بار) و Adaptive Boosting KNN (۱ بار)، کاربرد گسترده آن را نشان می دهد. روش های بگینگ، اگرچه کمتر از بوسستینگ رایج هستند، اما هنوز هم نقش مهمی دارند. مدل Bagging Classifier (۱۴ بار) و مدل های دیگری مانند Bagged Tree و Bagging with Logistic Regression بیشترین تکرار را داشته اند. قابل ذکر است که مدل RF (۱۱۱ بار) به طور کلی پرکاربردترین مدل گروهی است که استحکام و تطبیق پذیری آن را در مطالعات نشان می دهد. سایر تکنیک های گروهی، مانند درختان اضافی^۵ (ETs) (۱۱ بار)، پشته ای (۷ بار) و طبقه بندی کننده های رای گیری^۶ (۹ بار) نیز اهمیت ترکیب چندین مدل را برای افزایش عملکرد پیش بینی نشان می دهند.

مدل های یادگیری ماشین (ML) سنتی و کلاسیک: در میان مدل های

سنتی، SVM ها بسیار محبوب هستند و در ۹۴ مقاله ظاهر شده اند و اثربخشی آنها را در طبقه بندی داده های با ابعاد بالا منعکس می کند. در عین حال، رگرسیون لجستیک (۶۹ بار)

^۱. Gradient Boosting (GB)

^۲. Light Gradient Boosting Machine (LightGBM)

^۳. Categorical Gradient Boosting (CatBoost)

^۴. Gradient Boosting Tree

^۵. Extra Trees (ETs)

^۶. Voting Classifiers

یک مدل پیش فرض برای طبقه بندی باینری است. مدل های مبتنی بر درخت مانند DT (۵۴) بار) و RF (۱۱۱) به دلیل قابلیت تفسیر و تطبیق پذیری مورد علاقه بسیاری از مقالات بوده اند. مدل های KNN (۵۶) و بیز ساده (۴۶) نیز به دلیل سادگی بودنشان به خوبی مورد استناد قرار گرفته اند. مدل های بهینه سازی مانند طبقه بندی کننده های نزولی گرادیان تصادفی^۱ (SGDC) (۱۳) بار) و الگوریتم های ژنتیک (۱) بار) نیاز به بکارگیری مدل هایی را نشان می دهند که بتوانند مجموعه داده های بزرگ را با همگرایی سریعتر مدیریت کنند. مدل های تجزیه و تحلیل بقا که کمتر مورد استفاده مطالعات قرار گرفته اند، مانند جنگل های بقا تصادفی و رگرسیون کاکس، اهمیت داده های زمانی را در حوزه های خاص قلب و عروق نشان می دهند. با پیشرفت مدل های جدید ML مانند مدل های یادگیری ماشین کوانتومی^۲ (QML) همچون SVC کوانتومی (QSVC) (۲) بار) و شبکه های عصبی کوانتومی^۳ (QNNs) (۱) بار) تلاش های اولیه برای استفاده از محاسبات کوانتومی برای ML را نشان می دهند. در نهایت، مدل های خوشه بندی، مانند K-means (۲) بار) و خوشه بندی سلسله مراتبی (۱) بار)، در رویکردهای یادگیری بدون نظارت، و همچنین روش های طبقه بندی مبتنی بر فازی (۱) بار) نقش مدل های کلاسیک خوشه بندی و طبقه بندی را در رسیدگی به CVD ها را در بعضی از مقالات نشان می دهند.

مدل های یادگیری عمیق (DL): مدل های DL تنوع قابل توجهی را در مطالعات نشان می دهند. شبکه های عصبی مصنوعی (ANNs) به طور کلی با ۷۵ بار تکرار پرکاربردترین مدل ها و شبکه های مورد استفاده هستند و انواع آنها مانند پرسپترون چند لایه^۴ (MLP) (۲۲) بار) و شبکه های عصبی پس انتشار^۵ (BPNN) (۴) بار) کاربرد گسترده ای در مقالات داشته اند. CNN ها و معماری های آنها، مانند ResNet (۹) بار)، MobileNet (۴) بار) و InceptionV۳ (۱) بار)، کاربرد قابل توجهی در وظایف پردازش تصویر و سیگنال را ایفا کرده اند. از بین شبکه های عصبی بازگشتی^۶ (RNNs)، شبکه های حافظه طولانی کوتاه مدت^۷ (LSTM) (۱۱) بار) استفاده شده است، در حالی که انواع آن مانند BiLSTM (۶)

^۱. Stochastic gradient descent classifiers (SGDC)

^۲. Quantum machine learning (QML)

^۳. Quantum neural networks (QNNs)

^۴. Multi-Layer Perceptron (MLP)

^۵. Backpropagation

^۶. Recurrent Neural Networks (RNNs)

^۷. Long short-term memory (LSTM)

بار) و واحد بازگشتی دروازه‌ای^۱ (GRU) (۲ بار) برای داده‌های متوالی استفاده شده‌اند. قابل ذکر است که یادگیری انتقالی^۲ (TL) و معماری‌های پیشرفته، اگرچه کمتر تکرار شده‌اند اما در برخی مطالعات بخاطر اثر پذیری آنها در محیط‌های دارای منابع محدود مورد استفاده قرار گرفته‌اند.

جدول ۴. مدل‌ها و تکنیک‌های یادگیری مورد استفاده در مدیریت بیماری‌های قلبی عروقی بر اساس سه دسته اصلی همراه با فراوانی آنها.

دسته بندی کلیه مدل ها، تکنیک ها و الگوریتم های مورد استفاده	نام الگوریتم یا مدل و تعداد فراوانی آنها در مقالات
مدل های یادگیری گروهی (EL)	
مدل های تقویتی (GBMs) (۱۳۸ بار)	XGB (۵۰), Gradient Boosting (۳۵), ADABOOST (۲۷), LGBM (۱۳), Catboost (۹), Gradient Boosting Tree (GBT) (۵), Boosting (۵), Adaptive Boosting (۲), Gradient Boost Decision Tree (GBDT) (۳), Adaptive Boosting KNN (۱), Adaptive Boosting Tree (۱), Boosted generalized linear model (BoGLM) (۱), AdaBoostM ^۱ with LR (۱), AdaBoostM ^۱ with decision stump (DS) (۱), Boosted tree (۱), LogitBoost (۱), Catboost-tuned (۱), Stochastic Gradient Boosting (SGB) (۱)
مدل های بگینگ (۱۳۱ بار)	RF (۱۱۱), Bagging Classifier (BaggC) (۴), Bagged tree (BT) (۳), Bagging_QSVM (۱), Bagging with LR (۱), Bagging with REPTree (۱),
سایر مدل های گروهی (۴۹ بار)	Ensemble Classifier (۱۸), Extra Trees (۱۱), Stacking Algorithm (۷), Soft and hard voting (۵), Voting (۴), Regression Ensemble (۱), Ensemble Deep Learning (EDL) (۱), Weighted Ensemble Model (WEM) (۱),
مدل های یادگیری ماشینی سنتی و کلاسیک (ML).	
مدل های رگرسیون (۹۴ بار)	LR (۶۹), Linear regression (۹), Ridge Classifier (۲), Elastic / Elated-Net (۲), Multiple /Multivariate Linear Regression (MLR) (۲), LogisticRegressionCV (LogRegCV) (۱), least squares (LS) boost (۱), Flexible discriminant analysis (FDA) (۱), Cox proportional hazard (CPH) (۱), Generalized additive model (GAM) (۱), Lasso (۱), Generalized linear model (GLM) (۱), Binomial logistic regression (BLR) (۱), Multivariate

^۱. Gated Recurrent Unit (GRU)

^۲. Transfer Learning (TL)

logistic regression (MLR) (۱), Multivariate adaptive regression spline (MARS) (۱),	
DT (۵۴), Classification & regression trees (CART) (۳), Rotation Forest (RotF) (۲), logistic model tree (J۴۸) (۲), K-best algorithm (۱), Optimal classification trees (OCT) (۱), XGBRF (۱), Conditional inference tree (CIT) (۱), Fine tree (۱), Random tree (RT) (۱),	طبقه بندی درخت (۶۸ بار)
SVM/SVC (۸۱), Linear SVM (LSVM) (۴), Sequential minimal optimization (SMO) (۲), GaussianProcessClassifier (GauPro) (۱), OneVsRestClassifier (OvsR) (۱), Nu svc (۱), Gaussian radial kernel SVM (۱), SVM RBF / Support Vector Classifier with Radial Basis Function kernel (SVCR) (۱), Adaptive binary tree-based SVM (۱), Passive Aggressive (۱),	طبقه بندی بردار پشتیبان (SVC) (۹۴ بار)
Naïve Bayes (NB) (۳۲), GaussianNB (GauNB) (۸), Bayesian Network (۳), Bayesian generalized linear model (BGLM) (۱), Multinomial naive bayes (۱), Bernoulli naive bayes (۱),	شبکه های بیزین (۴۶ بار)
KNN (۵۶), JRip (۱), One Rule (۱), Instance-based classifier (IBk) (۱),	نزدیکترین همسایه ها (۵۹ بار)
Stochastic Gradient Descent Classifier (SGDC) (۱۳), Linear Discriminant Analysis (LDA) (۷), Passive Aggressive Classifier (PAC) (۱), Quadratic discriminant analysis (QDA) (۱), SGDClassifierMod (SGDCMod) (۱) Modified Artificial Plant Optimization (MAPO) (۱), Genetic Algorithm (۱),	مدل های بهینه سازی (۲۵ بار)
Random Survival Forest (RSF) (۳), Extra Survival Trees (۱), Survival SVM (SSVMs) (۱), Survival Tree (۱), Semi parametric (cox) survival (۱), Discrimination and Calibration of Cox regression (۱), Cox regression (۱),	مدل های تجزیه و تحلیل بقا (۹ بار)
K-means (۲), Hierarchical Self-Calibrating (HSC) (۱), Markov statistics nonparametric techniques (۱), Case-based reasoning (CBR) (۱), DBSCAN (۱), Mean Shift Clustering (۱), Hierarchical Clustering (۱), K-modes (۱), Kstar (۱) Fuzzy-based classification methods (۱),	خوشه بندی و مدل های فازی (۱۱ بار)
Quantum SVC (QSVC) (۲), Quantum NNs (QNNs) (۱), Hybrid quantum multi-layer perceptron (HWMLP) algorithm (۱),	یادگیری ماشین کوانتومی (QML) (۴ بار)
مدل های یادگیری عمیق (DL)	

CNN (۲۱), ResNet/۵۰ (۹), Mobilenet (۴), Transfer learning algorithms (۲), AlexNet (۲), TabNet (۱), Xception (۱), NASNet (۱), Capsulenet (۱), InceptionV۳ (۱), DenseNet-۱۲۱ (۱), U-Net (۱), Region Convolutional Neural Network (RCNN) (۱)	شبکه های عصبی پیچشی (CNNs) (۴۶ بار)
ANN (۲۳), MLP (۲۲), Neural Network (۱۴), BPNN (۴), Backpropagation (۲), Deep Neural Network (DNN) (۲), Extreme Learning Machine (ELM) (۲), Radial basis function (RBF) (۱), Quadratic Classifier (۱), Cascade Forward Neural Network (CFNN) (۱), Model average neural network (mANN) (۱), Deep Belief Network (DBN) (۱), Feed Forward Neural Network (FFNN) (۱), HQMLP (۱)	شبکه های عصبی مصنوعی (ANNs) (۷۵ بار)
LSTM (۱۱), BiLSTM (۶), RNN (۳), GRU (۲), Stacked LSTM (۲), Elman Neural Network (ELMAN) (۱), Vanilla LSTM (۱), Extreme boosting LSTM (۱), A-LSTM (۱),	شبکه های عصبی بازگشتی (RNNs) (۲۸ بار)
Transfer learning (TL) (۲), Boltzmann machine (۱), Deep Auto-Encoder (DAE) (۱),	سایر مدل های یادگیری عمیق (۴ بار)

تحزیه و تحلیل یافته ها و نتایج فراترکیب

این بخش نگاهی عمیق به هر چهار دسته مدیریت CVD با استفاده از مدل ها و تکنیک های EL ارائه می دهد.

تشخیص زود هنگام بیماری های قلبی عروقی با استفاده از مدل های یادگیری گروهی

اهداف تشخیصی در حوزه بیماری های قلبی عروقی بر بهره گیری از منابع داده چند دامنه مانند سیگنال های ECG، تصویربرداری، متابولومیک و داده های بالینی تأکید دارند تا دقت تشخیصی بیماری های مختلف از جمله آریتمی ها، انفارکتوس میوکارد، CAD و سکتة مغزی را فراهم کنند. بررسی مطالعات در این زمینه نشان دهنده پیشرفت چشمگیر در استفاده از مدل های EL و ML/DL برای تشخیص CVD ها است. مدل ها و تکنیک های گروهی نظیر RF، XGBoost، GB، AdaBoost، CatBoost و Bagging به دلیل توانایی

یکپارچه‌سازی خود، همواره عملکرد بهتری نسبت به مدل‌های منفرد داشته‌اند؛ این روش‌ها برای مدیریت مجموعه داده‌های نامتعادل، بهینه‌سازی ویژگی‌ها و تعمیم مدل بسیار مناسب بوده‌اند. RF به دلیل عملکرد بالا در داده‌های بالینی ساختاریافته و سیگنال‌های ECG با ابعاد بالا برجسته است، در حالی که XGBoost در شناسایی روابط غیرخطی و مقابله با داده‌های نویزدار و نامتعادل برتری دارد. علاوه بر این، مدل‌های ترکیبی که ML و DL را با مدل‌های EL ترکیب می‌کنند، تعادل بهینه‌ای بین قابلیت تفسیر و دقت تشخیصی ارائه داده‌اند.

مدل‌های یادگیری عمیق، به ویژه CNN ها، LSTM ها، GRU ها، ANN ها و شبکه‌های عصبی عمیق^۱ (DNNs)، به سرعت به روشی مطلوب برای مدیریت داده‌های بزرگ و پیچیده تبدیل شده‌اند و در وظایفی مانند طبقه‌بندی سیگنال ECG، تحلیل صدای قلب و تشخیص مبتنی بر تصویر، به دقت‌های قابل توجهی دست یافته‌اند. تکنیک‌های پیش‌پردازش و بهینه‌سازی داده‌ها و ویژگی‌ها، از جمله روش‌های انتخاب ویژگی مانند MAPO و GradCAM، و همچنین روش‌های رسیدگی به عدم تعادل کلاس‌ها با استفاده از تکنیک‌های نمونه‌گیری مانند تکنیک بیش‌نمونه‌گیری اقلیت مصنوعی^۲ (SMOTE)، نقش حیاتی در افزایش عملکرد مدل‌ها ایفا کرده‌اند. بهینه‌سازی فرآیندها^۳ و تکنیک‌های فیلترینگ مانند کاهش نویز در سیگنال‌های ECG و PPG نیز به بهبود دقت کمک کرده‌اند. یافته‌ها نشان می‌دهند که ادغام مدل‌های سنتی ML و تکنیک‌های پیشرفته DL با روش‌های EL و استخراج ویژگی‌های مهم، تشخیص CVD ها را به طور چشمگیری بهبود بخشیده است و در کاربردهای خاص به دقت‌های بیش از ۹۰ درصد دست یافته‌اند؛ حتی برخی از مدل‌های CNN که با تبدیل موجک^۴ (WT) ترکیب شده‌اند، دقت تا ۹۹٫۵۱ درصد را گزارش کرده‌اند.

– الگوریتم‌ها و تکنیک‌های کلیدی مورد استفاده برای تشخیص CVD:

برای تشخیص، مطالعات بررسی شده در این دسته، الگوریتم‌ها و تکنیک‌های کلیدی را به کار برده‌اند. در میان آنها، RF، انواع تقویت‌گرادیان (مانند XGBoost)، و SVM به عنوان پرکاربردترین‌ها برجسته شده‌اند. همچنین سایر مدل‌های سنتی ML شامل رگرسیون

^۱. Deep Neural Networks (DNNs)

^۲. Synthetic Minority Oversampling Technique (SMOTE)

^۳. Hyperparameters

^۴. Wavelet transforms (WTs)

لجستیک، K-نزدیکترین همسایه، درخت تصمیم و بی‌یز ساده به دلیل سادگی و کارایی خود در تشخیص CVD ها رایج بوده‌اند. با این حال، عملکرد آن‌ها اغلب در مواجهه با داده‌های پیچیده و با ابعاد بالا ضعیف گزارش شده است. به عنوان مثال، SVM، اگرچه در مطالعات تشخیص آریتمی‌ها استفاده شده، اما به طور کلی ضعیف‌تر از الگوریتم‌های پیچیده‌تر عمل کرده است. RF، به دلیل توانایی‌اش در تجمیع چندین درخت تصمیم و افزایش دقت، به طور گسترده در مقالات استفاده شده (Zafar et al. ۲۰۲۳) (Elvas et al. ۲۰۲۳; Moskaleva et al. ۲۰۲۲; Simegn et al. ۲۰۲۲; Yang et al. ۲۰۱۸) (Antoniades et al. ۲۰۲۲) و در برخی مطالعات، مانند (Antoniades et al. ۲۰۲۲) (Simegn et al. ۲۰۲۲)، (Husain et al. ۲۰۲۳) (Aarthy and Iqbal ۲۰۲۴) (Zafar et al. ۲۰۲۳). به عنوان بهترین الگوریتم شناخته شده است.

تکنیک‌های یادگیری گروهی، از جمله Bagging، RF، AdaBoost و XGBoost، به طور مداوم با رویکردهای ML و DL ترکیب شده‌اند تا عملکرد را بهبود بخشند. این روش‌ها در مدیریت داده‌های نامتعادل و نویزدار مؤثر بوده و تشخیص‌ها و پیش‌بینی‌های پایدارتری نسبت به مدل‌های منفرد ارائه داده‌اند. برای مثال، RF و XGBoost اغلب در ترکیب با تکنیک‌هایی مانند SMOTE یا روش‌های انتخاب ویژگی استفاده می‌شوند، و مطالعاتی نظیر (Dhanka et al. ۲۰۲۳) و (Ardeti et al. ۲۰۲۲) به کاربرد XGBoost در ترکیب با سایر طبقه‌بندی‌کننده‌ها اشاره کرده‌اند. همچنین، الگوریتم‌هایی مانند CatBoost، AdaBoost و LightGBM نیز به طور مکرر در مقالات ظاهر شده‌اند که کارایی آن‌ها را در بهبود دقت تشخیصی نشان می‌دهد؛ به عنوان مثال، ترکیب XGBoost و AdaBoost برای افزایش دقت در مدل‌های غیرخطی در (Balaha et al. ۲۰۲۲). (Kokol et al. ۲۰۲۱) مشاهده شده است. در حالی که روش‌های DL مانند CNN، LSTM، ANN و DNN کمتر از تکنیک‌های کلاسیک ML رایج بوده‌اند، اما برای پردازش داده‌های پیچیده مانند سیگنال‌های ECG و PPG بسیار ضروری بوده‌اند (Sapra et al. ۲۰۲۳)، (Eleyan and Alboghbaish ۲۰۲۴) و (Elvas et al. ۲۰۲۳). CNN ها به طور گسترده در تحلیل ECG استفاده شده و مدل‌های ترکیبی مانند AlexNet-۵۰ و ResNet-۵۰ نتایج امیدوارکننده‌ای در تحلیل سیگنال‌های ECG و صدای قلب نشان

داده‌اند. جدول چهارم خلاصه‌ای از این یافته‌ها و مسائل مورد بحث در بالا در مورد تشخیص CVD ها با استفاده از مدل‌های مختلف یادگیری را نشان می‌دهد.

- بهترین الگوریتم با بالاترین عملکرد

یافته‌های مطالعات این دسته نشان می‌دهند که یادگیری عمیق در ترکیب با یادگیری گروهی (EDL) بهترین نتایج را ارائه می‌دهد و به دقت ۱۰۰٪ دست یافته است، در حالی که RF و XGBoost از پرکاربردترین و قابل اعتمادترین الگوریتم‌ها در مطالعات متعدد بوده‌اند. در این مطالعات، عملکرد معیار دقت بین ۷۳ تا ۱۰۰ درصد بسته به عواملی مانند پیچیدگی داده‌ها، کیفیت پیش پردازش و روش‌های بهینه‌سازی متغیر بوده است. الگوریتم RF چهار بار در مطالعات برای عملکرد بالای خودش با دقت بین ۸۲٪ تا ۱۰۰٪ ذکر شده است. به طور مشابه، XGBoost کاربرد بالایی را نشان داده و به دقت بین ۹۱٫۸۵٪ تا ۱۰۰٪ دست یافته است. به عنوان مثال، XGBoost و RF با هم به دقت ۱۰۰٪ در مطالعه (Simegn et al. ۲۰۲۲) برای تشخیص بیماری‌های قلبی کسب کرده است. XGBoost در مقالات متعددی دارای دقت متفاوت است به عنوان مثال در مطالعه ی (Dhanka et al. ۲۰۲۳) به دقت ۹۱٫۸۵٪ و در مطالعه ی (Ogunpola et al. ۲۰۲۴) پس از بهینه‌سازی به دقت ۹۸٫۵۰ درصد رسیده است. سایر مدل‌های گروهی نیز عملکرد خوبی از خود نشان داده‌اند، -CRDHHO- EL (یک مدل ترکیبی EL) در مطالعه ی (Divya et al. ۲۰۲۳) دقت ۹۹٫۴۳٪ را به دست آورده است. همچنین، طبقه‌بندی کننده درختان اضافی (ETC) دارای دقت ۱۰۰٪ در مطالعه ی (Balaha et al. ۲۰۲۲) بوده است.

از سوی دیگر، مدل‌های DL، به ویژه EDL، نمرات کاملی را به دست آورده‌اند، در حالی که روش‌های سنتی ML مانند LR و SVM به طور کلی دقت کمتری (حدود ۷۰-۸۰٪) داشته‌اند. برخی از مدل‌های DL که در مقالات عملکرد بالایی کسب کرده‌اند عبارتند از: مدل CNN در [۵۵] با دقت ۹۹٫۴۰٪ در تشخیص آریتمی؛ مدل AlexNet۵۰ در [۸۰] با دقت ۹۸٫۹۶٪؛ مدل ترکیبی LASSO-CNN با دقت ۹۹٫۵۱٪ در [۱۰۲] جهت تشخیص بیماری‌های قلبی؛ مدل DNN با دقت ۹۸٫۳٪ جهت مدیریت سیگنال‌های ECG؛ مدل CatBoost با دقت ۹۰٫۹۴٪؛ شبکه‌های عصبی ساده در دو مطالعه با دقت ۹۶ درصد؛ و مدل CapsulNet با دقت ۹۶٫۷ درصد در (Latha et al. ۲۰۲۲).

^۱. Extra Trees Classifier (ETC)

پیش بینی و پیش آگهی بیماری های قلبی عروقی با استفاده از مدل های یادگیری گروهی

مطالعات بررسی شده در زمینه پیش بینی بیماری های قلبی عروقی بر افزایش قدرت تشخیصی جهت پیش بینی، پیش آگهی و پیامدهای درمان تمرکز دارند. این تحقیقات به طور عمده بر توسعه مدل های پیش بینی خطر بسیار دقیق برای بیماری های قلبی و فشار خون بالا، شناسایی عوامل خطر حیاتی، و بهبود قابلیت تفسیر مدل های EL برای متخصصان مراقبت های بهداشتی تأکید می کنند. بیماری های شایعی مانند CVD، بیماری عروق کرونر، نارسایی قلبی و فشار خون بالا به عنوان پیش بینی کننده های کلیدی مورد مطالعه قرار گرفته اند. ویژگی های تحلیل شده عمدتاً شامل سن، کلسترول و فشار خون هستند. اکثر مطالعات به چالش هایی نظیر مجموعه داده های نامتعادل، مقادیر از دست رفته، برازش بیش از حد و فرآیندهای محاسباتی فشرده پرداخته اند تا استحکام و کارایی مدل ها را افزایش دهند. برخی تحقیقات بر پیش بینی بقای بیمار (مانند بقای نارسایی قلبی)، مدیریت منابع بیمارستانی، و پیش بینی مرگ و میر در بیمارستان تمرکز کرده اند. همچنین، کاربردپذیری مدل های EL و تکنیک های ترکیبی برای افزایش دقت و کاهش هشدارهای کاذب بررسی شده است، و تمرکز فزاینده ای بر استفاده از هوش مصنوعی قابل توضیح برای افزایش شفافیت در پیش بینی خطر CVD مشاهده می شود. هدف کلی این مطالعات، افزایش پیش بینی بیماری، پیش بینی خطر و قدرت تشخیصی برای پیش بینی و تحلیل بقا با استفاده از تکنیک های پیشرفته EL و ML/DL است.

بهینه سازی مدل های یادگیری از طریق انتخاب ویژگی و تنظیم فرامترها در اکثر جنبه ها و اهداف مطالعات پیش بینی مشاهده می شود. رویکردهای رایج در این دسته شامل مدل های گروهی (مانند RF، GBMs، Stacking و چارچوب های ترکیبی)، روش های یادگیری عمیق (مانند CNN، LSTM، GRU، DNN) و سایر الگوریتم های کلاسیک ML (مانند SVM، DT، LR، KNN، و NB) هستند. مدل های XGBoost، GB، LightBoost، RF و سایر طبقه بندی کننده های EL به طور گسترده برای بهینه سازی نتایج پیش بینی در CVD ها استفاده شده اند. توسعه استراتژی های ترکیبی، مانند ترکیب ML و

DL با تکنیک‌های EL، نشان‌دهنده نیاز به مدل‌های قوی‌تر برای مقابله با پیچیدگی داده‌های پزشکی است. یافته‌های کلی نشان می‌دهد که الگوریتم‌هایی مانند XGBoost، RF و LightGBM اغلب از نظر دقت از دیگر مدل‌ها بهتر عمل کرده‌اند و انتخاب ویژگی و روش‌های EL در بهبود عملکرد مدل نقش اساسی داشته‌اند. مطالعات همچنین بر رفع محدودیت‌ها در روش‌های سنتی مانند چالش‌های پیش‌پردازش داده‌ها، عدم تعادل طبقه‌ای، بهینه‌سازی پارامترهای مدل و دستیابی به دقت پیش‌بینی و تشخیصی بالاتر متمرکز بوده‌اند.

- الگوریتم‌های تکراری در پیش‌بینی CVD ها:

XGBoost و RF الگوریتم‌هایی هستند که اغلب در مطالعات ذکر شده‌اند و چندین بار در وظایف پیش‌بینی CVD ظاهر شده‌اند. این الگوریتم‌ها به دلیل استحکام، توانایی مدیریت متغیرهای طبقه‌بندی شده و پیوسته و دقت پیش‌بینی بالا به خوبی شناخته شده‌اند. RF تقریباً در هر مطالعه‌ای به دلیل مدیریت مجموعه داده‌های پیچیده دیده شده است. XGBoost، نیز پرکاربردترین الگوریتم در چندین مقاله بوده که به طور برجسته در مطالعاتی مانند (Lee et al. ۲۰۲۰; Hu et al. ۲۰۲۲; Datta et al. ۲۰۲۳; Bhatt et al. ۲۰۲۳) (You et al. ۲۰۲۳; ۲۰۲۴) نتایج خوبی را کسب کرده است. XGBoost به طور مداوم عملکرد قوی را در وظایف مختلف، از جمله توانبخشی CVD (Torres et al. ۲۰۲۳)، پیش‌بینی ترخیص از بیمارستان (Ahn et al. ۲۰۲۱) و پیش‌بینی فشار خون بالا (Islam et al. ۲۰۲۲, ۲۰۲۳) نشان داده است. مدل GB اغلب با RF جفت شده تا دقت را افزایش دهد، در حالی که AdaBoost اغلب در تنظیمات گروهی برای بهبود عملکرد پیش‌بینی استفاده شده است. CatBoost، در مطالعاتی مانند (Pan ۲۰۲۴; Baseer et al. ۲۰۲۳) (Samaras et al. ۲۰۲۲; et al. ۲۰۲۲)، در مطالعاتی مانند (Ohashi et al. ۲۰۲۳) و (Tachie et al. ۲۰۲۳)، نیز به کار گرفته شده‌اند. علاوه بر این، مدل‌های سفارشی و ترکیبی، مانند RFRF-ILM (Guo et al. ۲۰۲۰) و ConvSGLV (Rustam et al. ۲۰۲۲)، به طور مکرر برای افزایش بیشتر دقت و استحکام پیشنهاد شده است. سایر الگوریتم‌های رایج شامل LR، SVM، DT، NB و KNN هستند. در این میان، الگوریتم LR یک معیار سنتی و در عین حال پرکاربرد برای مقایسه است، در حالی که الگوریتم SVM به ویژه برای مجموعه داده‌های کوچکتر موثر بوده است.

- مدل ها و الگوریتم ها با بالاترین عملکرد در پیش بینی CVD ها:

مدل های EL، به ویژه آنهایی که مبتنی بر تقویت گرادیان و RF هستند، بالاترین سطح از عملکرد را در پیش بینی بیماری های قلبی-عروقی نشان داده اند. در میان الگوریتم های تقویتی، XGBoost در مطالعات متعدد به بالاترین دقت دست یافته است که شامل معیارهایی چون $R^2=92\%$ (Torres et al. ۲۰۲۳)، $AUROC=0.865$ (Ahn et al. ۲۰۲۱)، $AUC=0.99$ (Lv et al. ۲۰۲۱) و دقت 96% (Mondal et al. ۲۰۲۴) و 88.81% (Islam et al. ۲۰۲۳) می شود. سایر مدل های مبتنی بر GBM مانند AdaBoost نیز دقت 95% را ثبت کرده اند (Nissa et al. ۲۰۲۴). RF نیز در ترکیب با روش های دیگر به نتایج چشمگیری رسیده است؛ از جمله ترکیب آن با EG-HFS (دقت 96.23%) (Lambay and Mohideen ۲۰۲۴)، روش Relief در مدل RFBM (دقت 99.05%) (Ghosh et al. ۲۰۲۱) مدل های خطی بهبودیافته در RFRF-ILM (دقت 96.6%) (Guo et al. ۲۰۲۰) و همراه با درخت تصمیم (دقت 96.75%) (Husain et al. ۲۰۲۳). علاوه بر این، روش های پشته ای نیز کارایی بالایی از خود نشان داده اند، مانند مدل PaRSEL با دقت 97% (Noor et al. ۲۰۲۳) و سایر مدل های Stacked Ensemble با دقت هایی نظیر 92.34% (Tiwari et al. ۲۰۲۲) و 87.8% (Dritsas and Trigka ۲۰۲۳). روش های گروهی دیگر مانند دسته بندی های با رأی گیری نرم^۱ (SVEC) با $AUC=99.97\%$ (Lertampaiporn et al. ۲۰۲۲) و ماشین یادگیری افراطی (ELM) با دقت تا 99.06% (Najim and Nasri ۲۰۲۴) نیز عملکرد بی نقصی را به ثبت رسانده اند.

رویکردهای ترکیبی سفارشی، مدل های DL و برخی الگوریتم های سنتی نیز نتایج برجسته ای را به دست آورده اند. مدل های ترکیبی پیشنهادی مانند GRU-ODE-Bayes با $AUROC=0.812$ (Lee et al. ۲۰۲۴)، BMDA-MKELM با دقت 99.5% (Sheeba et al. ۲۰۲۲)، IACPSO-LR با دقت 98% (Dubey et al. ۲۰۲۲)، MaLCaDD با دقت 99.1% (Rahim et al. ۲۰۲۱) و MMFO-MLP با دقت 96.89% (Ahmadian et al. ۲۰۲۲) عملکردی فراتر از مدل های پایه از خود نشان دادند. در حوزه یادگیری عمیق، LSTM به $AUROC$ برابر 0.98 (Datta et al. ۲۰۲۲)، DNN به دقت 90.2% (Stepanyan et al. ۲۰۲۲) و ترکیب CNN با مدل های گروهی

^۱. RF bagging machine (RFBM)

^۲. Soft-voting ensemble classifiers (SVEC)

(ConvSGLV) به دقت ۹۳٪ (Rustam et al. ۲۰۲۲) دست یافتند. همچنین ادغام TabNet و CatBoost دقت ۹۸٪ را به همراه داشت (Baseer et al. ۲۰۲۴). جالب توجه است که برخی مدل‌های سنتی‌تر نیز در شرایط خاصی درخشیده‌اند؛ برای مثال، KNN به دقت ۹۲٫۱٪ (Sozen et al. ۲۰۲۳) و رگرسیون خطی تعمیم‌یافته به دقت ۹۸٪ (Karboub ۲۰۲۲ and Tabaa) دست یافته‌اند که نشان‌دهنده اهمیت انتخاب الگوریتم متناسب با داده و مسئله است.

طبقه‌بندی CVD ها بر اساس مدل‌های یادگیری گروهی

طبقه‌بندی بیماری‌های قلبی-عروقی نقشی محوری در مدیریت آنها ایفا می‌کند، زیرا به طور مستقیم بر تشخیص، پیش‌بینی و برنامه‌ریزی درمانی تأثیر می‌گذارد. طبقه‌بندی صحیح، امکان شناسایی بیماری‌های مشخصی مانند آریتمی و بیماری عروق کرونر، درجه‌بندی شدت خطر برای تنظیم درمان و ارائه برنامه‌های پیشگیرانه سفارشی را فراهم می‌آورد، در حالی که طبقه‌بندی نادرست می‌تواند به تصمیمات پزشکی نامناسب یا تأخیر در درمان منجر شود. مطالعات در این حوزه بر چهار دسته اصلی متمرکز هستند: طبقه‌بندی نوع بیماری، تفکیک ریتم طبیعی قلب از غیرطبیعی، پیش‌بینی خطر مَرگ ناگهانی قلبی و تشخیص چندکلاس بیماری‌هایی نظیر نارسایی قلبی، تنگی آئورت و زوال عقل.

هدف مشترک در مطالعات تحلیل شده‌ی این دسته، طبقه‌بندی هر چه بهتر بیماری‌ها/بیماران و دستیابی به تشخیص زودهنگام و دقیق CVD ها از طریق تلفیق مدل‌های نوآورانه EL با ML/DL است. این مقالات با بهره‌گیری از ویژگی‌های پیشرفته استخراج شده از سیگنال‌های ECG و اصوات قلب، از روش‌های متنوعی مانند CNN ها، LSTM ها، مدل‌های تقویت گرادیان (مانند XGBoost) و حتی چارچوب‌های یادگیری ماشین کوانتومی (QML) برای بهبود دقت تشخیصی در کاربردهای مختلفی همچون طبقه‌بندی آریتمی‌ها، تحلیل نقایص دریچه قلب، و پیش‌بینی خطر مَرگ ناگهانی قلبی در بیماران خاص استفاده می‌کنند. این رویکرد یکپارچه، حرکت به سمت استفاده از محاسبات پیشرفته برای مدیریت شخصی‌سازی‌شده‌ی بیماری‌های قلبی را نشان می‌دهد.

– الگوریتم ها و تکنیک های پر کاربرد و تکراری با عملکرد بالا در طبقه بندی CVD ها

در حوزه طبقه بندی CVD ها، مدل های EL رویکردی غالب محسوب می شوند که در میان آن ها، الگوریتم RF به دلیل قابلیت تعمیم بالا، بیشترین کاربرد را داشته و غالباً با مدل های دیگر نظیر CNN یا SVM ترکیب شده اند. در کنار الگوریتم های سنتی ML مانند SVM و KNN، مدل های DL از جمله CNN، LSTM و ResNet، به ویژه در استخراج ویژگی و طبقه بندی سیگنال های ECG، نقش کلیدی دارند. در میان آنها SVM یکی از الگوریتم های پر کاربرد است که اغلب با استفاده از تکنیک های فرابتنکاری برای به حداکثر رساندن دقت طبقه بندی تنظیم و ترکیب شده است. موفقیت بالای مدل های ترکیبی مانند CNN-LSTM و EXGB-LSTM در طبقه بندی های پیچیده و چندکلاسه در مطالعاتی مانند (Essa and Xie ۲۰۲۱، Eleyan and Alboghbaish ۲۰۲۴) و (Abdullah et al. ۲۰۲۳) به اثبات رسیده است. علاوه بر این، رویکردهای نوینی مانند یادگیری ماشین کوانتومی (QML) (Maheshwari et al. ۲۰۲۳؛ Enad and Mohammed ۲۰۲۴)، از جمله SVM کوانتومی (QSVM)، شبکه های عصبی کوانتومی (QNNs) و Bagging-QSVM، و یادگیری انتقال (Xiang et al. ۲۰۲۳) نیز برای افزایش دقت و پایداری مدل ها مورد استفاده قرار گرفته اند.

عملکرد این مدل ها در طیف وسیعی از دقت، بین ۶۷/۳۴٪ تا ۹۹/۹۲٪، قرار می گیرد که اکثر چارچوب های یادگیری عمیق و ترکیبی به دقتی بالاتر از ۹۰٪ دست یافته اند. مدل D-ResNet با دقت ۹۸/۹۲٪ (Boukhaboul et al. ۲۰۲۴) و مدل ترکیبی FFT+CNN-LSTM با دقت ۹۷/۶٪ (Eleyan and Alboghbaish ۲۰۲۴) نتایج برجسته ای کسب کرده اند. الگوریتم RF به تنهایی در یک مطالعه به دقت استثنایی ۹۹/۹۲٪ (Hassaballah et al. ۲۰۲۳) و در ترکیب با XGB به حساسیت و ویژگی ۹۹٪ (Pavic et al. ۲۰۲۳) رسیده است. سایر مدل ها نیز نتایج قابل توجهی داشته اند؛ برای مثال، مدل K-PCA با EXGB-LSTM به دقت ۹۴/۳٪ (Abdullah et al. ۲۰۲۳)، مدل XGB به دقت ۹۰/۴۶٪ (Lee et al. ۲۰۲۱) و الگوریتم کوانتومی QSVM به دقت ۹۰/۱۶٪ (Enad and Mohammed ۲۰۲۴) کسب کرده اند.

شناسایی CVD ها با استفاده از مدل های یادگیری گروهی

مطالعات تحلیل شده در این دسته بر استفاده از مدل های EL همراه با رویکردهای ML و DL برای شناسایی بیماری های قلبی عروقی و شرایط مرتبط با آن تمرکز دارند. هدف اصلی این تحقیقات، شناسایی عوامل خطر کلیدی و پیش بینی کننده های CVD برای افزایش دقت تشخیص، درمان و پیشگیری از این بیماری هاست. شناسایی دقیق و سریع عواملی نظیر فشار خون بالا، دیابت، کلسترول بالا، سابقه بیماری قلبی، سبک زندگی و سابقه خانوادگی، که تشخیص CVD ها را دشوار می کنند، از اهمیت بالایی برخوردار است. این مطالعات همچنین بر مدیریت شخصی CVD ها، توضیح پذیری مدل ها و ادغام فناوری هایی مانند اینترنت اشیا تأکید دارند. نتایج کلی بر سه هدف اصلی متمرکز است: شناسایی عوامل و ویژگی های خطر حیاتی، بهبود عملکرد تشخیصی با استفاده از مدل های EL و ML/DL، و بهینه سازی کارایی و هزینه های مراقبت های بهداشتی. در این میان، EL به عنوان متداول ترین روش، با ترکیب نقاط قوت الگوریتم های متعدد، عملکرد برتر را به نمایش می گذارد. رویکردهای ترکیبی مانند Stacking نیز قدرت پیش بینی را بیشتر کرده اند.

مطالعات از طیف گسترده ای از روش های گروهی مانند RF، Gradient Boosting، XGBoost، AdaBoost و الگوریتم های Bagging در کنار مدل های ML/DL و روش های یادگیری انتقال استفاده کرده اند. مدل های RF و XGBoost نیز به طور مکرر با هم ظاهر و ترکیب شده اند که نشان دهنده اثربخشی آنها در مقابله با مجموعه داده های پیچیده CVD است. برای مقابله با چالش هایی مانند عدم تعادل داده ها و مقادیر از دست رفته، تکنیک های پیش پردازش مانند SMOTE، انتساب چند گانه توسط معادلات زنجیره ای^۱ (MICE) و الگوریتم های انتخاب ویژگی از جمله الگوریتم حداکثر ارتباط حداقل افزونگی^۲ (mRMR) به طور گسترده ای به کار گرفته شده اند. قابلیت تفسیر مدل از طریق تجزیه و تحلیل توضیحات افزودنی شاپلی^۳ (SHAP) و شناسایی عوامل حیاتی نیز در اکثر مطالعات این دسته کاربرد دارد. ویژگی هایی مانند شاخص توده بدنی، سن، سطح کلسترول، فشار خون سیستولیک، سیگار کشیدن و سطح گلوکز در بسیاری از مقالات به عنوان عوامل تأثیرگذار در پیش بینی بیماری های قلبی عروقی مطرح شده اند. اعتبارسنجی

^۱. Multiple Imputation by Chained Equation (MICE)

^۲. Minimum Redundancy Maximum Relevance (mRMR)

^۳. SHapley Additive exPlanations (SHAP)

داده‌ها بر روش‌های قوی مانند اعتبارسنجی متقابل k -fold، روش‌های بوت استرپ و بهینه‌سازی جستجوی شبکه‌ای متکی هستند که مدل‌های قابل اعتماد و قابل تعمیم را تضمین می‌کنند. علی‌رغم پیشرفت‌ها، مطالعات با محدودیت‌هایی مانند حجم نمونه کوچک، عدم تعادل داده‌ها و تعمیم‌پذیری محدود به دلیل استفاده از مجموعه داده‌های تک بیمارستانی یا منطقه خاص مواجه‌اند.

- پرکاربردترین و بهترین الگوریتم‌ها در شناسایی CVD ها:

در این دسته نیز مانند سه دسته دیگر، روش‌های EL به عنوان یک دسته غالب، به طور گسترده در مطالعات مختلف مورد استفاده قرار گرفته‌اند. الگوریتم‌هایی نظیر RF، GB، XGBoost، AdaBoost، Bagging، Extra Trees و LightGBM اغلب در تحقیقات متعدد به کار رفته‌اند. این الگوریتم‌ها غالباً به منظور بهره‌گیری از نقاط قوت هر یک، به صورت ترکیبی مورد استفاده قرار می‌گیرند؛ نمونه‌هایی از این ترکیبات شامل RF + XGB، GB + AdaBoost + LGBM می‌باشند. در این میان، RF تقریباً در تمامی مقالات، مورد بررسی قرار گرفته و به عنوان یکی از مورد بحث‌ترین و تطبیق‌پذیرترین الگوریتم‌ها شناخته می‌شود که کارایی خود را در وظایفی نظیر پیش‌بینی بقا، تشخیص جنسیت و طبقه‌بندی گروه سنی اثبات کرده است. به طور مشابه، XGBoost و GB نیز از اهمیت بالایی برخوردار بوده و به کرات مورد بحث قرار گرفته‌اند. تحلیل مطالعات این دسته نشان می‌دهد که مدل‌های RF، AdaBoost، XGBoost معمولاً با تکنیک SMOTE ترکیب شده تا با چالش عدم تعادل داده‌ها مقابله کنند. همچنین، برای تفسیر نتایج مدل‌های پیچیده، به ویژه در ترکیب با الگوریتم‌هایی نظیر XGBoost و RF، از روش SHAP به وفور استفاده شده است.

علاوه بر الگوریتم‌های EL، مدل‌های ML شامل درخت تصمیم، ماشین بردار پشتیبان، K نزدیک‌ترین همسایگان، رگرسیون لجستیک، بی‌بیز گوسی (GNB) و رگرسیون کاکس^۱، اغلب به صورت جداگانه یا گروهی در تحلیل‌های مقایسه‌ای ظاهر شده‌اند. رویکردهای DL، هرچند کمتر از ML سنتی مورد بحث قرار گرفته‌اند، اما در برخی مطالعات گنجانده شده‌اند. تکنیک‌هایی نظیر شبکه‌های عصبی پیچشی، شبکه‌های عصبی بازگشتی و یادگیری انتقالی غالباً با روش‌های گروهی ترکیب شده‌اند و رویکردهای ترکیبی

^۱. Cox Regression

را برای وظایف پیچیده استخراج ویژگی و پیش‌بینی برجسته کرده‌اند. در زمینه شناسایی CVD، روش‌های EL به ویژه پشته‌ای به طور مداوم عملکرد بهتری نسبت به الگوریتم‌های مستقل از خود نشان داده‌اند که این امر بر اثربخشی ترکیب مدل‌ها تأکید می‌کند. به طور خاص، روش‌های EL (مانند Stacking، RF، XGBoost) به طور پیوسته به دقت و مقدار AUC بالاتری دست یافته‌اند به عنوان مثال، در مطالعه‌ای توسط (Louridi et al. ۲۰۲۱)، الگوریتم پشته‌ای بالاترین دقت (۹۵٫۸۳٪) را کسب کرد که آن را در میان الگوریتم‌های مورد اشاره به مؤثرترین روش تبدیل نمود. نتایج حاصل از RF نیز ۹۴٪ - ۹۶٪ برای پیش‌بینی بقا، تشخیص جنسیت و طبقه‌بندی گروه سنی) کارایی بسیار مؤثر این الگوریتم را در زمینه‌های خاص برجسته می‌کند (Kumar et al. ۲۰۲۱). در پژوهشی دیگر (Tadesse et al. ۲۰۲۱)، مدل DeepMI با بهره‌گیری از استراتژی‌های تلفیق داده به AUROC ۹۶٫۷٪ دست یافته است. ارزیابی معیار دقت در مطالعات تحلیل شده نیز از ۴۹٫۵٪ (با الگوریتم mRMR (Hsu ۲۰۱۸) تا ۹۵٫۸۳٪ (با الگوریتم انباشته) (Louridi et al. ۲۰۲۱) متغیر است.

مقالات دیگر

از بین مطالعات بررسی شده، سه مقاله در چهار دسته فوق قرار نمی‌گیرند و بر استفاده از مدل‌های EL و ML/DL در پایش بیماری‌های قلبی عروقی و فشار خون تمرکز دارند. در مقاله (Desai et al. ۲۰۲۲) یک سیستم رایانش ابری مبتنی بر ML به نام (HealthCloud) برای پایش سلامت بیماران قلبی معرفی شده است که از داده‌های مجموعه‌ی UCI استفاده می‌کند و مدل‌هایی مانند LR، شبکه‌های عصبی، SVM، درختان تقویت گرادیان و KNN را بررسی می‌کند. مقاله (Li et al. ۲۰۲۰) به برآورد غیرتهاجمی فشار خون ناشی از سیگنال‌های PPG، ECG و فشار خون شریانی^۱ (ABP) در مجموعه داده MIMIC II با استفاده از مدل BiLSTM پرداخته است. این مطالعه دقت بالای آن را در مقایسه با مدل‌هایی مانند رگرسیون خطی چندگانه^۲ (MLR) و روش‌های گروهی رگرسیون (مانند بگینگ، تقویت و پشته‌ای) نشان داده است. در نهایت، مقاله (Maqsood et al. ۲۰۲۱) روش‌های استخراج ویژگی از سیگنال‌های PPG برای پیش‌بینی فشار خون را با استفاده از

^۱. Arterial blood pressure (ABP)

^۲. Multiple linear regression (MLR)

مجموعه داده های MIMIC-II و PPG-BP تجزیه و تحلیل کرده است و نشان می دهد که ویژگی های حوزه زمان در ترکیب با مدل های DL مانند BiLSTM و GRU بهترین عملکرد را دارند. هر سه مقاله بر استفاده از مدل های مختلف یادگیری برای بهبود نظارت و پیش بینی بیماری های قلبی عروقی و نظارت بر فشار خون و بررسی چالش هایی مانند دقت پیش بینی، همگام سازی حسگر و پیچیدگی پردازش داده ها متمرکز بودند.

معیارهای ارزیابی عملکرد مدل های یادگیری در مدیریت CVD ها

تجزیه و تحلیل مطالعات نشان می دهند که معیارهای ارزیابی مدل برای ارزیابی مدل های ML/DL و EL برای مدیریت CVD ها چقدر مهم هستند. این معیارها که از تکنیک های مختلف آماری و ریاضی حاصل می شوند، دقت، قابلیت اطمینان و تعمیم پذیری مدل ها را تضمین می کنند. طیف وسیعی از معیارهای آماری به دست آمده از ماتریس درهم ریختگی^۱ و سایر روش های کمی عمدتاً در ارزیابی مدل های یادگیری استفاده شده اند. در مدیریت بیماری های قلبی عروقی، ماتریس در ارزیابی میزان تمایز مدل بین بیمارانی که این بیماری را دارند و بیمارانی که این بیماری را ندارند بسیار مهم است (Husain et al. ۲۰۲۳; Lambay and Mohideen ۲۰۲۴; Lilda et al. ۲۰۲۲; Nissa et al. ۲۰۲۴). شکل ۴ ساختار ماتریس درهم ریختگی را بر حسب چهار پیامد اصلی و به دنبال آن معیارهای ارزیابی مهم ارائه می دهد.

شکل ۴. چارچوب ماتریس درهم ریختگی و معیارهای ارزیابی حاصل از آن.

^۱. Confusion Matrix

کلاس پیش بینی شده

		کلاس واقعی	
	مثبت	منفی	
مثبت	مثبت واقعی (TP)	منفی کاذب (FN) خطای نوع دوم	معیار حساسیت (TPR) $\frac{TP}{TP + FN}$
منفی	مثبت کاذب (FP) خطای نوع اول	منفی واقعی (TN)	معیار خاصیت (FPR) $\frac{TN}{TN + FP}$
	مقدار پیش بینی شده مثبت (دقت) (PPV) $\frac{TP}{TP + FP}$	مقدار پیش بینی شده منفی (NPV) $\frac{TN}{TN + FN}$	معیار صحت $\frac{TP + TN}{TP + FP + TN + FN}$

این معیارها توانایی پیش بینی، عملکرد و سازگاری مدل ها را در وظایفی مانند طبقه بندی باینری و چند کلاسه تخمین می زنند. عملکرد طبقه بندی معمولاً با معیارهای قابل توجهی مانند صحت^۱، حساسیت (یادآوری)^۲، خاصیت^۳، دقت^۴، نمره $F1^5$ و مساحت زیر منحنی مشخصه عملکرد سیستم^۶ (AUC-ROC) اندازه گیری می شود. برای پرداختن به مسائلی مانند طبقه بندی چند طبقه ای و عدم تعادل طبقاتی، مطالعات معیارهای پیچیده دیگری مانند از دست دادن همینگ، دقت متوازن و ضریب همبستگی متیوز^۷ (MCC) را بررسی کرده اند. همچنین برای پیش بینی پیامدهای پیوسته در برخی از مدل های خاص، از معیارهای رگرسیون یا خطای خاص و همبستگی مانند R^2 (R-squared)، میانگین خطای مطلق^۸ (MAE)، جذر میانگین مربعات خطا^۹ (RMSE)، میانگین مربعات خطا^{۱۰} (MSE) و غیره استفاده شده است. با ارائه بینش جامع در مورد عملکرد مدل در بسیاری از مجموعه داده ها، تکنیک های اعتبارسنجی متقابل مانند اعتبارسنجی متقابل k-fold اعتبار این ارزیابی

^۱. Accuracy

^۲. Sensitivity (Recall)

^۳. Specificity

^۴. Precision

^۵. $F1$ - Score

^۶. The area under the ROC curve (AUC-ROC)

^۷. Matthews correlation coefficient (MCC)

^۸. Mean absolute error (MAE)

^۹. Root mean square error (RMSE)

^{۱۰}. Mean square error (MSE)

ها را بهبود بخشیده اند. تعدادی از این معیارها از ماتریس درهم ریختگی مشتق شده اند که پیش بینی ها را به مثبت کاذب^۱ (FP)، مثبت واقعی^۲ (TP)، منفی کاذب^۳ (FN) و منفی واقعی^۴ (TN) دسته بندی می کند. بعضی از مطالعات از معیارهای خاص مدل، معیارهای اعتبارسنجی و آماری استفاده کرده اند که برای مشخصات مجموعه داده مناسب هستند، مانند پرداختن به عدم تعادل مجموعه داده یا ارائه بینش عمیق تر و دقیق تر از طریق معیارهایی مانند نمرات بریر، کاپای کوهن، مقدار احتمال^۵، شاخص یودن^۶ (YI)، قابلیت تفسیر و غیره. هدف اصلی مدیریت CVD ها بهینه سازی مدل ها و تضمین طبقه بندی، حساسیت و صحت دقیق است. جدول پنجم طبقه بندی جامعی از معیارهای ارزیابی عملکرد مورد استفاده در مدل های EL و ML/DL برای مدیریت CVD ها را ارائه می دهد. این معیارها را در دسته های مهمی گروه بندی می کند، مانند معیارهایی برای ارزیابی عمومی بر اساس ماتریس های درهم ریختگی، معیارهای مبتنی بر خطای رگرسیون، معیارهای همبستگی، معیارهای خاص یک مدل، روش های اعتبارسنجی، معیارهای آماری و شاخص های عملکرد محاسباتی. هر دسته شامل طیف وسیعی از معیارها است که برای ارزیابی دقت، قابلیت تفسیر و کارایی محاسباتی مدل ضروری است.

جدول ۵. طبقه بندی جامعی از معیارهای ارزیابی عملکرد مورد استفاده برای مدل های EL و ML/DL در مدیریت CVD ها، همراه با فراوانی آنها در مطالعات بررسی شده.

معیارهای ارزیابی و فراوانی آنها	دسته بندی معیارهای ارزیابی
- Confusion Matrix (۴۴)، - Accuracy (۱۱۴)، Balanced Accuracy (۱)، Mean Accuracy (MA) (۱)، Average accuracy of each state (MPA) (۱)، Accuracy of each state (CPA) (۱)، Average Accuracy (۱)، Overall accuracy (PA) (۱)، Multi-value accuracy (MACC) (۱)، Prediction accuracy of TAE (۱)، - Specificity (۵۱)، Average specificity (ASp) (۱)، - Precision (۸۲)، Macro-precision (MP) (۱)،	معیارهای ارزیابی عمومی (معیارهای ماتریس درهم ریختگی)

^۱. False Positive (FP)

^۲. True Positive (TP)

^۳. False Negative (FN)

^۴. True Negative (TN)

^۵. P-Value

^۶. Youden Index (YI)

- Sensitivity (۵۵), Recall (۶۵), Positive Predictive Value (PPV) (۹), Macro-recall (MR) (۱), Average sensitivity (ASe) (۱), - F-measure / F1-Score (۸۶), Macro-F1 score (MF1) (۲), Macro-Averaged F1-Score (۱), Weighted F1-Score (۱), - AUC-ROC (۶۶), AUC-PR (۸), PRC Area (۱), - Matthews Correlation Coefficient (MCC) (۱۰), Fallout (۱), Gmean (۱), Critical Success Index (CSI) (۱), The Dice similarity coefficient (DSC) (۲), Jaccard score (۱), Detection Rate (DR) (۱), Detection Prevalence (DP) (۱), - False Positive Rate (FPR) (۱۸), True Positive Rate (TPR) (۱۴), Negative predictive values (NPVs) (NPR) (۶), True Negative Rate (TNR) (۴), False Negative Rate (FNR) (۱), False discovery rate (FDR) (۱), False omission rate (FOR) (۷), Average True Positive Predictive (ATPP) (۱),	
جذر میانگین مربعات خطا (RMSE) (۱۰)، میانگین خطای مطلق (MAE) (۴)، میانگین مربعات خطا (MSE) (۱)، میانگین مربعات خطای نرمال شده (NMSE) (۱)، میانگین درصد مطلق خطای (MAPE) (۱)، SEE (۱)، خطای مطلق نسبی (RAE) (۱)، Average squared error (ASE) (۱)، نرخ خطا (۷)، تلفات ورود به سیستم (۲)، ضرر همینگ (۱)، نرخ ضرر (۱)،	معیارهای رگرسیون (معیارهای مبتنی بر خطا)
ضریب تعیین (R^۲) (۵)، مربع R (R^۲)، اسپیرمن (۱)، ضریب همبستگی (۱)، Hosmer-Lemeshow-۲ (۱)،	معیارهای همبستگی
قابلیت تفسیر (I) (۱)، شاخص تفسیرپذیری-دقت (IAI) (۱)، شاخص تفسیرپذیری-سازگاری (ICI) (۱)، شاخص سازگاری (C-index) (۱)، شاخص-IPCW C (۱)، Prescription Effectiveness (PE) (۱)، Intersection Over Union (IoU) (۱)، Prescription Robustness (PR) (۱)، احتمال آستانه (۲)، حداکثر اختلاف میانگین (MMD) (۱)، Expected Reciprocal Rank (ERR) (۱)،	معیارهای خاص مدل
فواصل اطمینان بوت استرپ برای اعتبارسنجی خارجی (۱)، اعتبارسنجی متقابل برای اعتبارسنجی داخلی (به عنوان مثال، hold-out، k-fold) (۵۴)، میانگین اعتبارسنجی متقابل (۱)، میانگین دقت K-fold (۱)، شاخص Youden (YI) (۳)، فاصله اطمینان (CI) (۷)،	معیارها و تکنیک های اعتبارسنجی
نمره بریر (BS) (۳)، میانگین نمره بریر (۱)، نمودارهای کالیبراسیون (۱)، آماره کاپا کوهن یا نمره کاپا (KS) (۹)، P-Value (۶)، انحراف معیار (SD) (۴)، آزمون Mann U Whitney (۱)، نمودار وابستگی جزئی (PDP) (۲)، نسبت پایداری (۱)، نمودار کاپلان مایر (۱)،	معیارهای آماری
زمان کل (۵)، QoS (۱)، تأخیر (۱)، استفاده از حافظه (۱)، گزارش طبقه بندی (۱)، نرخ طبقه بندی نادرست (۱)،	معیارهای محاسباتی

محدودیت ها و چالش های مدیریت بیماری های قلبی عروقی با استفاده از مدل های EL و ترکیبی

مطالعاتی که بر استفاده از تکنیک های EL، ML و DL در مدیریت بیماری های قلبی عروقی تمرکز دارند، با چالش ها و محدودیت های مکرری مواجه هستند. این محدودیت ها عمدتاً شامل مسائل مربوط به کیفیت و حجم داده ها، تعمیم پذیری، عدم تعادل کلاس ها، و نیازهای محاسباتی است. یکی از مشکلات برجسته، اتکا به مجموعه داده های کوچک یا تک مرکزی است که به دقت متوسط، قابلیت تعمیم پذیری پایین، و تنوع عملکردی محدود منجر می شود. همچنین، عدم تعادل کلاس در مجموعه داده ها می تواند عملکرد مدل را در کلاس های اقلیت کاهش دهد، در حالی که داده های از دست رفته (مانند سیگنال های ECG ناقص) بر دقت و طبقه بندی تأثیر منفی می گذارند. علاوه بر این، استفاده از مدل های خطی برای شناسایی الگوهای پیچیده و غیرخطی در داده های قلبی عروقی ناکافی است، و فقدان اعتبارسنجی مدل ها در مجموعه داده های دنیای واقعی، قابلیت تعمیم پذیری آن ها را به زمینه های بالینی مختلف زیر سوال می برد.

محدودیت های دیگر شامل پیچیدگی محاسباتی و نیازهای منابع برای مدل های پیچیده، به ویژه مدل های DL مانند CNN، است که استقرار آن ها را در کاربردهای بلندرنگ محدود می کند. همچنین، پیچیدگی روش های گروهی می تواند قابلیت تفسیر مدل ها را کاهش دهد و از سودمندی آن ها برای تصمیم گیری بالینی بکاهد. مشکلات در مهندسی و انتخاب ویژگی نیز از جمله چالش های مهم است؛ وجود نویز در داده های با ابعاد بالا کارایی مدل را کاهش می دهد و عدم وجود روش های قوی برای انتخاب ویژگی های مرتبط، دقت پیش بینی ها را محدود می کند. در نهایت، مطالعات با محدودیت های عملکردی مانند حساسیت پایین در شناسایی بی نظمی های کمتر رایج یا تمرکز صرف بر طبقه بندی عمومی صداهای قلب مواجه هستند، که دامنه کاربرد آن ها را محدود می سازد. موانع اجرایی و عدم پوشش کافی پایش بلندرنگ و تجزیه و تحلیل مستمر در تحقیقات نیز مانع استقرار عملی این مدل ها می شوند. شکاف در آزمایش و اعتبارسنجی مکرر، از جمله فقدان اعتبارسنجی خارجی، قابلیت اطمینان نتایج بسیاری از مطالعات را کاهش می دهد. برای غلبه

بر این چالش‌ها، نیاز به مجموعه داده‌های بزرگ‌تر و متنوع‌تر، روش‌های بهبود یافته انتخاب ویژگی، اعتبارسنجی خارجی، و چارچوب‌های محاسباتی کارآمدتر وجود دارد. جدول ششم محدودیت‌ها و چالش‌هایی را که اغلب در استفاده از مدل‌های EL، ML و DL برای مدیریت CVD ها مواجه می‌شوند، ارائه می‌دهد.

جدول ۶. محدودیت‌ها و چالش‌های مکرر در به‌کارگیری مدل‌های EL، ML و DL برای مدیریت CVD ها.

دسته بندی	زیرشاخه	نمونه	توضیحات
محدودیت‌های مجموعه داده	مجموعه داده های کوچک و داده های محدود	MIT-Statlog HD, BIH, بیمارستان جیانگشی	مجموعه داده های کوچک استحکام آماری و تعمیم پذیری را کاهش می‌دهند.
	عدم تعادل داده ها	نسبت بیماری به غیر بیماری،	توزیع نابرابر کلاس آموزش مدل را سوگیری می‌کند و دقت کلاس اقلیت را کاهش می‌دهد. برای رفع آن تکنیک‌های نمونه برداری مجدد استفاده می‌شود.
	داده های گمشده	شکاف های سیگنال ECG و جزئیات از دست رفته بیماران	اطلاعات دموگرافیک/بالینی ناقص یا گم شده بر دقت تأثیر می‌گذارد.
	کیفیت داده ها و داده های نویز دار	تصاویر ECG نویز دار، مانند مجموعه داده ACD، تصاویر شبکه با وضوح پایین	داده های با کیفیت پایین یا مصنوعی استحکام مدل را کاهش می‌دهد.
	وابستگی خاص داده	مجموعه داده های منطقه ای (دانشگاه سچنوف و بیمارستان سانتا ماریا)	مدل های آموزش دیده بر روی جمعیت های خاص ممکن است به خوبی تعمیم نگیرند.
	نوع داده ها	مجموعه داده های کلیولند، مجارستان و سوئیس	چالش های یکپارچه سازی داده های ساختاریافته و بدون ساختار.
	مجموعه داده های عمومی در مقابل خصوصی	مانند مجموعه داده MIT-BIH و کلیولند قلب	استفاده بیش از حد از مجموعه داده های عمومی خطر برآزش بیش از حد را به همراه دارد. مجموعه داده های اختصاصی تکرارپذیری را محدود می‌کند.
	مهندسی و انتخاب ویژگی	استخراج ویژگی محدود در برخی مطالعات	داده های با ابعاد بالا ممکن است نویز ایجاد کنند. انتخاب ویژگی محدود بر دقت تأثیر می‌گذارد.

محدودیت های مدل	پیچیدگی محاسباتی	XGB ، CNN-LSTM ، CapsuleNet ، RESNet۵۰	مدل های DL و EL به منابع بالایی نیاز دارند و کاربردهای بلادرنگ را محدود می کنند.
	حساسیت مدل	عملکرد تأثیر سیگنال های ECG نویزدار	برخی از مدل ها به نویز بسیار حساس هستند و قابلیت اطمینان را کاهش می دهند.
	مسائل مقیاس پذیری	قابلیت های پردازش ECG بلادرنگ محدود	مشکل در مدیریت داده های در مقیاس بزرگ یا بلادرنگ.
	قابلیت تفسیر مدل	XGB و گروه های عمیق	مدل های DL و گروهی فاقد قابلیت تفسیر هستند و بر پذیرش بالینی تأثیر می گذارند.
محدودیت های عملکرد	محدودیت های عملکرد	به عنوان مثال، مشکلات در معیار حساسیت در مدل AdaBoost	حساسیت کمتر برای نمونه های نادر؛ طبقه بندی باینری پیش بینی های ظریف را محدود می کند.
	نظارت بلادرنگ	مسائل پردازشی سیگنال های قلبی همچون ECG به صورت بلادرنگ	مطالعات کمی چالش های استقرار بلادرنگ را بررسی می کنند.
	مراحل اولیه توسعه	مانند تکنیک های برچسب دار	برخی از روش ها هنوز در مراحل آزمایشی با مجموعه داده های محدود هستند.
	عدم اعتبارسنجی خارجی	عدم وجود ارزیابی اعتبارسنجی متقابل در مجموعه داده های خارجی	بسیاری از مطالعات فاقد اعتبارسنجی خارجی هستند، که قابلیت اطمینان را کاهش می دهند.
کاربرد در دنیای واقعی	چالش های یکپارچه سازی در مراقبت بهداشتی	ادغام محدود با سیستم های EHR	مشکل در همسویی خروجی های مدل با گردش کار بالینی.
	عملکرد در محیط های نویزدار	مانند مسائل مربوط به ثبات در تشخیص ضربان قلب مبتنی بر ویدئو	مدل های آزمایش شده در تنظیمات ایده آل ممکن است در سناریوهای دنیای واقعی شکست بخورند.
	هزینه محاسباتی	مدل های DL که به خوشه های GPU نیاز دارند	نیازهای پردازش بالا برخی از مدل ها را در محیط های دارای منابع محدود غیرعملی می کند.
	محدودیت های مجموعه داده	مجموعه داده های MIT-BIH و Statlog	مطالعات تک مرکزی فاقد قابلیت تعمیم به جمعیت های مختلف هستند.
تعمیم پذیری و اعتبارسنجی	نمایندگی کم گروه های اقلیت	برخی از مجموعه داده ها بر گروه های قومی خاص تمرکز می کنند	تنوع جمعیتی محدود بر کاربرد گسترده تر تأثیر می گذارد.

شکاف های اعتبارسنجی	عدم اعتبار سنجی متقابل مجموعه داده	بسیاری از روش ها فاقد آزمایش بر روی مجموعه داده های دنیای واقعی هستند که کاربرد آن را محدود می کند.
چالش های تجزیه و تحلیل زیر گروه	ناتوانی در تجزیه و تحلیل گروه های خطر خاص	حجم نمونه کوچک تجزیه و تحلیل زیرگروه را محدود می کند.

در میان این محدودیت ها، شایع ترین محدودیت، اندازه مجموعه داده کوچک بود که اغلب با مسائل مربوط به تعمیم پذیری محدود همراه بوده است. این دو اغلب با هم اتفاق می افتند، زیرا مجموعه داده های کوچک ذاتا توانایی یک مدل را برای تعمیم محدود می کنند. یکی دیگر از ترکیب های رایج از محدودیت ها، پیچیدگی محاسباتی همراه با محدودیت های داده بود که در مجموع مانع مقیاس پذیری و کارایی مدل ها می شد. داده های پر سر و صدا یا همان نویزی یکی از تکرارشونده ترین مسائل است و پس از آن حجم نمونه کوچک قرار دارد. یک چالش کلیدی این است که ترکیب منابع داده ناهمگن به دلیل مسائل استانداردسازی و یکپارچه سازی دشوار است. پیشنهادات عملی می تواند شامل افزایش تنوع مجموعه داده، بهبود تکنیک های پیش پردازش داده ها و استفاده از یادگیری فدرال برای مجموعه داده های اختصاصی باشد که می تواند محدودیت های زیادی را برطرف کند.

مهندسی داده و ویژگی و پیش پردازش در مدیریت CVD

مطالعات بررسی شده بر اهمیت یک رویکرد سیستماتیک برای پیش پردازش داده ها در مدیریت بیماری های قلبی عروقی تأکید دارند، با تمرکز ویژه بر مهندسی ویژگی، آماده سازی داده ها، تعادل داده ها و تنظیم فرایارامتر. این تکنیک ها با هدف بهبود عملکرد، قابلیت اطمینان، قابلیت تفسیر و تعمیم پذیری مدل های EL و ML/DL به کار رفته اند. این شامل مقابله با مسائلی نظیر داده های پر سر و صدا، عدم تعادل کلاس، ابعاد بالا و برآزش بیش از حد^۱ () است. ترکیب استراتژی های پیشرفته ای مانند کاهش ابعاد و بهینه سازی ابرپارامترها با روش های اساسی نظیر اعتبارسنجی متقابل و تجزیه و تحلیل داده های اکتشافی^۲ (EDA)،

^۱. Overfitting

^۲. Exploratory data analysis (EDA)

به نتایج قوی و عملکرد پیش‌بینی‌کننده مطلوب منجر می‌شود. علاوه بر این، ادغام منابع داده‌های مختلف و روش‌های پیش‌پردازش وابسته به دامنه، مانند کروماتوگرافی مایع برای پروفایل متابولیت، تنوع استراتژی‌های ابداع شده بر اساس پیچیدگی داده‌های قلبی عروقی را نشان می‌دهد.

روش‌های کلیدی در پیش‌پردازش داده‌ها به چهار دسته اصلی تقسیم می‌شوند: آماده‌سازی داده‌ها، انتخاب ویژگی، استخراج ویژگی، و روش‌های ترکیبی. انتخاب ویژگی، مانند روش‌های Filter، Wrapper، Embedded، Boruta، با کاهش ابعاد مجموعه داده از طریق انتخاب متغیرهای مرتبط، کارایی مدل را افزایش می‌دهد. تکنیک‌های نمونه‌گیری مانند SMOTE، MICE و انواع آن، عدم تعادل کلاس را برای انصاف بهتر مدل مدیریت می‌کنند، و SMOTE اغلب با روش‌های SHAP یا VIM برای تجزیه و تحلیل اهمیت ویژگی ترکیب می‌شود. تنظیم ابرپارامترها با استفاده از روش‌هایی نظیر RandomizedSearchCV و Optuna، پارامترهای مدل را برای عملکرد بهینه تنظیم می‌کند، و اعتبارسنجی متقابل k-fold همراه با آن برای اطمینان از استحکام و قابلیت اطمینان مدل استفاده می‌شود. روش‌های کاهش ابعاد مانند تجزیه و تحلیل مؤلفه‌های اصلی^۱ (PCA) و تجزیه و تحلیل مولفه‌های مستقل^۲ (ICA)، ابعاد را ضمن حفظ اطلاعات مهم کاهش می‌دهند. تکنیک‌های نرمال‌سازی و مقیاس‌بندی، داده‌ها را برای بهترین عملکرد مدل استاندارد می‌کنند، و فیلترها و تبدیل‌های فوری کیفیت داده‌ها را از طریق حذف نویز بهبود می‌بخشند. در نهایت، تکنیک‌های تفسیر مدل مانند SHAP و هوش مصنوعی قابل توضیح^۳ (XAI)، شفافیت در تصمیم‌گیری مدل را افزایش می‌دهند و پیش‌بینی‌های مبتنی بر هوش مصنوعی را قابل توضیح‌تر و قابل اعتمادتر می‌کنند. جدول هفتم چندین روش مورد استفاده در این زمینه یعنی تکنیک‌های مرتبط با مهندسی و پیش‌پردازش ویژگی‌ها و داده‌ها را طبقه‌بندی می‌کند که در اکثر مقالات مورد استفاده قرار گرفته‌اند.

جدول ۷. طبقه‌بندی وظایف و تکنیک‌های مهندسی داده و ویژگی به همراه فراوانی آنها در مطالعات بررسی شده.

^۱. Principal component analysis (PCA)

^۲. Independent component analysis (ICA)

^۳. Explainable artificial intelligence (XAI)

تکنیک ها	دسته بندی های مهندسی داده و ویژگی
Shearing (۱), Shifting (۱), Rotating (۱), Zooming (۱)	تکنیک های تقویت ویژگی و داده
Flipping and Cropping (۱)، الگوریتم پان تامپکینز (۲)، روش غربالگری بوروتا (۲)، روش تعبیه شده (۳)، روش Wrapper (۴)، روش فیلتر (۵)، Relief (۱)، InfoGain (۲)، حذف ویژگی بازگشتی (RFE) (۱)، آنتروپی (۲)، از دست دادن اطلاعات (۱)، تکنیک ارزیابی ویژگی ReliefF (۱)، تکنیک انتخاب ویژگی مبتنی بر همبستگی (۱)،	تکنیک های انتخاب و استخراج ویژگی
PCA (۷)، حداقل مربعات جزئی (PLS) (۱)، تجزیه و تحلیل مؤلفه های مستقل (ICA) (۱)،	کاهش داده ها و ویژگی ها
فیلتر پایین گذر (۱)، فیلتر میانه (۲)، فیلتر Butterworth برای نویز زدایی (۳)، تبدیل سریع فوری (FFT) (۱)، تبدیل فوری گسسته (DFT) (۱)، تبدیل موجک گسسته (DWT) (۲)، آمار مرتبه بالاتر (HOS) (۱)، ۱ D-local binary pattern (LBP) (۱)، فواصل R-R (۱)، گشتاور مرکزی (CM) (۱)، تابع پایه هرمیت (HBF) (۱)،	حذف داده ها؛ پردازش تصویر و سیگنال؛ و تکنیک های تشخیص الگو
SMOTE (۱۰)، روش بوت استرپ برای نامتعادل (۱)، نمونه برداری بیش از حد پیاده روی تصادفی (RWOS) (۱)، تکنیک نمونه گیری بیش از حد اقلیت وزن دار اکثریت (MWMOTE) (۱)، SMOTE مرزی (B-SMOTE) (۱)، مصنوعی تطبیقی (ADASYN) (۱)، نمونه گیری بیش از حد تصادفی (ROS) (۱)، نمونه برداری سایه افین تصادفی موضعی (LoRAS) (۱)، نمونه گیری سایه افین تصادفی وزنی مجاورت (ProWRAS) (۱)،	تکنیک های تعادل داده و کلاس
نرمال سازی / استاندارد سازی Z-score (۲)، نرمال سازی (۱)، Scaler Abs-Max (۱)، مقیاس بندی استاندارد (SS) (۳)، مقیاس بندی مقاوم (۱)، مقیاس بندی و نرمال سازی Min-Max (۵)،	پیش پردازش داده ها
هوش مصنوعی قابل توضیح (XAI) (۲)، SHAP (۱۴)،	قابلیت تفسیر مدل

- استخراج ویژگی

استخراج ویژگی یک گام حیاتی در بهبود عملکرد مدل های یادگیری است که داده های خام را به اطلاعات معنادار تبدیل می کند تا مدل بتواند الگوهای مهم را بیاموزد. در این فرآیند، تکنیک های مختلفی به کار گرفته می شوند. تجزیه و تحلیل حوزه زمان و زمان-فرکانس، به ویژه برای سیگنال های صوتی قلب، ویژگی های مهمی را استخراج می کند. ویژگی های مورفولوژیکی، مانند ماتریس های هم رخدادی سطح خاکستری (GLCM) (Lee et al. ۲۰۲۱)، ساختارهای داده ای را از ECG و صدای قلب آشکار می سازند.

تبدیلات موجک، از جمله حذف نویز موجک و تبدیل موجک گسسته^۱ (DWT)، برای پاکسازی داده‌ها و برجسته کردن الگوهای مهم استفاده می‌شوند. تبدیل فوریه سریع^۲ (FFT) نیز با ترکیب با CNN ها و LSTM ها برای استخراج ویژگی و مدل‌سازی توالی به کار رفته اند (Eleyan and Alboghbaish ۲۰۲۴). مدل‌های یادگیری عمیق مانند CNN در استخراج ویژگی و طبقه‌بندی چند کلاسه برتری دارند و از روش‌هایی مانند مهندسی ویژگی فضایی کانولوشن^۳ (CSFE) برای استخراج الگوهای معنی‌دار از مجموعه داده‌های پیچیده استفاده شده اند (Zafar et al. ۲۰۲۳). تکنیک‌هایی مانند تجزیه حالت تجربی^۴ (EMD) (Lilda et al. ۲۰۲۲) برای استخراج ویژگی‌های غیر خطی از سیگنال‌های ECG استفاده شده و به تشخیص دقیق تر CVD ها کمک می‌کند. همچنین، در (Yang et al. ۲۰۱۸) PCANet اعمال شده است که به افزایش استخراج ویژگی از سیگنال‌های ECG نویزدار کمک کرده است. در مطالعه (Rustam et al. ۲۰۲۲)، مدل ConvSGLV با ترکیب CNN با مجموعه‌ای از مدل‌های مختلف مانند SGDC، LR و SVM برای استخراج و طبقه‌بندی ویژگی‌ها توسعه داده شده است. در مطالعه دیگری، (Baseer et al. ۲۰۲۴) از TabNet برای استخراج ویژگی‌های پیچیده استفاده کرد که منجر به بهبود دقت پیش‌بینی شد. علاوه بر این، ویژگی‌های مهندسی دستی اغلب با فرآیندهای خودکار ترکیب می‌شوند تا دقت مدل‌ها افزایش یابد. در نهایت، ویژگی‌های استخراج شده به عنوان ورودی برای مدل‌های یادگیری برای مدیریت بیماری‌های قلبی-عروقی به کار می‌روند که این امر نقش کلیدی در افزایش دقت و کارایی مدل‌ها در تشخیص و مدیریت این بیماری‌ها دارد.

- انتخاب ویژگی و کاهش ابعاد

انتخاب ویژگی و کاهش ابعاد، دو رویکرد کلیدی در حوزه پیش‌پردازش داده‌ها هستند که نقشی حیاتی در بهینه‌سازی عملکرد و کارایی مدل‌های یادگیری ایفا می‌کنند. هدف اصلی این تکنیک‌ها، شناسایی و گزینش مهم‌ترین و مرتبط‌ترین ویژگی‌ها از میان انبوه داده‌ها، و همزمان حذف ویژگی‌های غیرضروری یا نویز دار است. این فرآیند، علاوه بر کاهش ابعاد مجموعه داده و پیچیدگی محاسباتی، به بهبود چشمگیر دقت و سرعت آموزش مدل‌ها

^۱. Discrete wavelet transforms (DWT)

^۲. Fast Fourier Transform (FFT)

^۳. Convolutional spatial feature engineering (CSFE)

^۴. Empirical Mode Decomposition (EMD)

کمک می‌کند و در نهایت، منجر به افزایش قابلیت تفسیر نتایج و درک بهتر روابط بین ویژگی‌ها و خروجی مدل می‌شود. برای دستیابی به این اهداف، روش‌های متنوعی توسعه یافته‌اند که هر یک نقاط قوت خاص خود را دارند. روش‌های مبتنی بر فیلتر^۱، مانند استفاده از اطلاعات متقابل^۲ (Dritsas and Trigka ۲۰۲۳) یا الگوریتم mRMR، ویژگی‌ها را بر اساس معیارهای آماری و مستقل از مدل ارزیابی می‌کنند (Hsu ۲۰۱۸). در مقابل، روش‌های مبتنی بر Wrapper، نظیر حذف ویژگی بازگشتی^۳ (RFE)، از عملکرد یک مدل یادگیری برای ارزیابی زیرمجموعه‌های مختلف ویژگی‌ها بهره می‌برند (Noor et al. ۲۰۲۳). روش‌های تعبیه‌شده^۴ نیز، مانند تکنیک عملگر گزینش و انقباض کمترین قدرمطلق^۵ (LASSO)، فرآیند انتخاب ویژگی را مستقیماً در روند آموزش مدل ادغام می‌کنند. علاوه بر این، تکنیک‌هایی مانند PCA برای کاهش ابعاد با تبدیل داده‌ها به فضایی با ابعاد کمتر کاربرد داشته، در حالی که ابزارهایی مانند SHAP به ارزیابی اهمیت هر ویژگی در تصمیم‌گیری مدل کمک کرده‌اند (Alfaidi et al. ۲۰۲۲; Islam et al. ۲۰۲۳). بهینه‌سازی ژنتیکی نیز به حذف ویژگی‌های اضافی برای افزایش کارایی مدل کمک می‌کند. فرایند انتخاب ویژگی در مدیریت بیماری‌های قلبی عروقی حائز اهمیت است، جایی که مدل‌های EL (مانند RF، SGD، GB، XGBoost) با استفاده از انتخاب ویژگی‌های بهینه شده، عملکرد بالاتری از خود نشان می‌دهند.

- پیش پردازش داده‌ها

پیش‌پردازش داده‌ها یک مرحله حیاتی و تکرارشونده در مدل‌های یادگیری به ویژه مدل‌های EL است که برای افزایش کیفیت داده‌ها و تضمین سازگاری آن‌ها با الگوریتم‌ها ضروری است، به ویژه در مدیریت بیماری‌های قلبی عروقی. این فرآیند شامل چندین وظیفه کلیدی است: ابتدا، مدیریت داده‌های از دست رفته با استفاده از تکنیک‌هایی مانند MICE، میانگین‌گیری و غیره برای حفظ یکپارچگی داده‌ها انجام می‌شود. در مرحله بعد، تشخیص داده‌های پرت و حذف نویز از طریق فیلترهای مختلف (مانند FIR، IIQR، Butterworth)

^۱. Filter Methods

^۲. Mutual Information

^۳. Recursive Feature Elimination (RFE)

^۴. Embedded Methods

^۵. Least Absolute Shrinkage and Selection Operator (LASSO)

برای بهبود کیفیت سیگنال (به ویژه در ECG و سیگنال‌های صوتی قلب) صورت می‌گیرد (Kokol et al. ۲۰۲۱) (Ardehi et al. ۲۰۲۲) (Aarthy and Iqbal ۲۰۲۴). سپس، نرمال‌سازی و مقیاس‌بندی داده‌ها با روش‌هایی مانند نرمال‌سازی حداقل-حداکثر و امتیاز Z برای جلوگیری از تسلط ویژگی‌های با محدوده عددی بزرگتر و بهبود عملکرد مدل اعمال می‌شود (Baghdadi et al. ۲۰۲۳). علاوه بر این، تکنیک‌های تعادل کلاس/داده مانند SMOTE (Jiang et al. ۲۰۲۳; Ogunpola et al. ۲۰۲۴) و SMOTE-ENN (Moreno-Sanchez ۲۰۲۳; Sinha et al. ۲۰۲۳; Shchetinin and Glushkova ۲۰۲۲) برای رسیدگی به مسائل عدم تعادل کلاس، که در مجموعه داده‌های پزشکی رایج است، به کار گرفته شده‌اند. همچنین، تقسیم‌بندی داده‌ها (مانند استفاده از روش U-Net) برای جداسازی بخش‌های مرتبط سیگنال استفاده می‌شوند. این فرآیندهای جامع پیش‌پردازش، داده‌ها را برای ورودی به مدل‌های یادگیری آماده می‌کنند و در نهایت به توسعه مدل‌های مدیریت CVD دقیق‌تر و قابل اطمینان‌تر کمک می‌کنند.

- بهینه‌سازی هایپرپارامتر

بهینه‌سازی هایپرپارامتر یک گام اساسی در بهبود دقت و پایداری مدل‌های یادگیری است که با هدف شناسایی بهینه‌ترین پارامترها انجام می‌شود. مطالعات مختلف از تکنیک‌های متنوعی برای این منظور بهره برده‌اند، از جمله روش‌های تنظیم فرایارامتر عمومی مانند RandomizedSearchCV و Optuna که به طور گسترده برای تنظیم دقیق مدل‌ها و افزایش عملکرد آن‌ها به کار رفته‌اند (Dhanka et al. ۲۰۲۳) و (Simegn et al. ۲۰۲۲). علاوه بر این، رویکردهای پیشرفته‌تری نظیر بهینه‌سازی‌های اصلاح‌شده (مانند بهینه‌سازی اصلاح شده ی پروانه-شعله^۱ (MMFO) برای افزایش همگرایی شبکه‌های عصبی MLP (Ahmadian et al. ۲۰۲۲) و بهینه‌سازی خودکار طبقه‌بندی شده ازدحام ذرات بهبود یافته^۲ (IACPSO) (Dubey et al. ۲۰۲۲) برای بهینه‌سازی انتخاب ویژگی)، و الگوریتم‌های خاص مانند TPE (Cocianu et al. ۲۰۲۳) (برای بهینه‌سازی شبکه‌های

^۱. Modified Moth-Flame Optimization (MMFO)

^۲. Improved Auto Categorical Particle Swarm Optimization (IACPSO)

(LSTM)، بهینه‌سازی مبتنی بر جغرافیای زیستی^۱ (BBO)، الگوریتم بهینه‌سازی سنجاقک^۲ (DFO) و خوشه‌بندی فضایی مبتنی بر چگالی^۳ (DBSCAN) (Sheeba et al. ۲۰۲۲) نیز با مدل‌های EL/ML ادغام شده‌اند. این تکنیک‌ها، شامل تنظیم معیارهایی مانند اندازه دسته، نرخ یادگیری و توابع فعال‌سازی هستند. همچنین تعدادی از الگوریتم‌های بهینه‌سازی پیشرفته‌ای همچون بهینه‌سازی شکارچیان دریایی^۴ (MPA)، بهینه‌سازی فراابتنکاری^۵ (MHO) و بهینه‌سازی کرم شب‌تاب^۶ (FFO) در مطالعات دیگر استفاده شده‌اند. در نهایت، بهینه‌سازهایی مانند RMSprop، Adam Optimizer، Grid Search CV و Adagrad، همراه با روش‌های سیستماتیک‌تر مانند برآورد پارزن با ساختار درختی^۷ (TPE) و RandomizedSearchCV، نقش مهمی در بهبود دقت و استحکام مدل‌های مورد استفاده در مدیریت بیماری‌های قلبی عروقی (CVD) ایفا می‌کنند.

جدول هشتم یک نمای کلی از دسته‌بندی‌های ابرپارامترها و تکنیک‌های مربوط به آنها که در مدل‌های EL و ML/DL برای مدیریت CVD استفاده شده‌اند، ارائه می‌دهد. دسته اول، معیارهای تنظیم فراپارامتر، اهمیت تنظیم دقیق پارامترها را برای دستیابی به عملکرد بهینه مدل برجسته می‌کند. در مجموع ۳۴ مقاله از تکنیک‌های تنظیم فراپارامتر عمومی در کار خود استفاده کرده‌اند که برخی از قابل توجه‌ترین آنها شامل اصلاح اندازه دسته، نرخ یادگیری و کاربرد تابع فعال‌سازی SoftMax است. دسته مدل‌های بهینه‌سازی، استفاده از الگوریتم‌های پیشرفته برای یافتن پیکربندی پارامترهای بهینه با استفاده از روش‌های بهینه‌سازی را به نمایش می‌گذارد. در نهایت، دسته بهینه‌سازها طیف متنوعی از استراتژی‌ها را نشان می‌دهد که Grid Search CV و Adam Optimizer متداول‌ترین روش‌هایی هستند که برای انتخاب هایپرپارامتر استفاده می‌شوند.

جدول ۸ طبقه‌بندی ابرپارامترها، بهینه‌سازها و تکنیک‌های مرتبط بکارگرفته شده در مدل‌های EL و ML/DL برای مدیریت CVD ها در مطالعات، همراه با فراوانی آنها.

تکنیک	دسته بندی هایپرپارامترها
-------	-----------------------------

^۱. Biogeographic-based optimization (BBO)

^۲. Dragonfly Optimization (DFO)

^۳. Density-Based Spatial Clustering of Applications with Noise (DBSCAN)

^۴. Marine Predators Algorithm (MPA)

^۵. Metaheuristic Optimization (MHO)

^۶. Firefly Optimization (FFO)

^۷. Tree-structured Parzen Estimation (TPE)

معیارهای تنظیم هایپرپارامتر	تنظیم هایپرپارامتر (۳۴)، Batch Size (۳)، نرخ یادگیری (۳)، Max Soft (۳)، Maxpooling (۲)، اندازه تست (۱)، نرمال سازی دسته ای (۱)، تکرار (۱)
مدل های بهینه سازی	الگوریتم های MPA (۱)، MHO (۱)، FFO (۱)، DFO (۱)، BBO (۱)، IACPSO (۱)، PSO (۱)
بهینه سازها	Grid search CV (۵)، بهینه ساز آدام (۴)، بهینه ساز RMSprop (ریشه میانگین مربع) (۳)، بهینه ساز آداگراد (۲)، برآورد پارزن ساختار درختی (TPE) (۱)، بهینه ساز AdaDelta (۱)، بهینه ساز AMSGrad Adam (۱)، بهینه ساز RMSProp Centered (۱)، بهینه ساز Nesterov SGD (۱)، بهینه ساز Ftrl (۱)، بهینه ساز AdaMax (۱)، جستجوی شبکه تصادفی (۱)، RandomizedSearchCV (۱)، بهینه سازی جستجوی تصادفی (۱)

- تقسیم و اعتبارسنجی داده ها

تقسیم و اعتبارسنجی داده ها برای ارزیابی معتبر مدل ها ضروری است. تکنیک های اعتبارسنجی، تخمین قوی از عملکرد مدل را بدون برازش بیش از حد و برآوردهای دقیق ارائه می دهد. روش هایی مانند اعتبارسنجی متقابل k-fold (به عنوان مثال، ۳-۵-۱۰ برابر و ۲۰-تکرار)، تقسیمات آزمون-تست (به عنوان مثال، ۷۰-۳۰٪ یا ۸۰-۲۰٪) و رویکرد بوت استرپ برای جلوگیری از برازش بیش از حد و ارائه برآوردهای عملکرد سازگار در اکثر مطالعات استفاده شده اند. این تکنیک های اعتبارسنجی معمولاً با تکنیک های بهینه سازی مانند جستجوی شبکه و تنظیم فرایامترها برای بهینه سازی پارامترهای مدل و به حداکثر رساندن دقت ترکیب شده اند. جدول نهم توزیع چین های اعتبارسنجی متقابل مورد استفاده در مطالعات مختلف را برجسته می کند. اعتبارسنجی متقابل ۱۰ برابری متداول ترین تکنیک است که در ۳۸ مقاله استفاده شده است که تعادل را بین کارایی محاسباتی و ارزیابی مدل قابل اعتماد منعکس می کند. اعتبارسنجی متقابل ۳-۵ برابری، که ۷ بار بکار قرار گرفته شده، زمانی استفاده می شود که اندازه مجموعه داده محدود باشد، یا مدل ها باید فوراً آزمایش شوند. علاوه بر این، روش های پیشرفته مانند اعتبارسنجی متقابل ۱۱ برابر، ۱۹ برابر و حتی k-fold Leave-Pair-Out (LPOCV) (Al-Yarimi ۲۰۲۲) یک بار اعمال شده اند که در حوزه های خاصی که ارزیابی عملکرد بالا ضروری است، صورت گرفته اند.

جدول ۹. توزیع فولد های اعتبارسنجی متقابل مورد استفاده در مطالعات مختلف.

فولد های اعتبارسنجی متقابل	فراوانی
۱۰ فولد	۳۸
۳-۵ فولد	۷

اعتبارسنجی متقابل ۱۱ فولد، ۱۹ فولد، Leave-Pair-Out (LPOCV) k-fold	هر مورد یک بار
---	----------------

- ابزارهای توضیح پذیری

تکنیک مقادیر SHAP در مطالعات مختلف به طور گسترده برای قابلیت تفسیر مدل ها استفاده شده اند که چگونگی نقش و اهمیت ویژگی های مختلف در مدل های تشخیصی و پیش بینی را توضیح می دهد. به عنوان مثال، در مطالعات (Islam et al. ۲۰۲۳; Torres et al. ۲۰۲۳) SHAP برای تجزیه و تحلیل اهمیت ویژگی استفاده شده، در حالی که در مطالعه (Moreno-Sanchez ۲۰۲۳)، از تکنیک XAI هوش برای همین منظور استفاده شده است. در مطالعه (Baghdadi et al. ۲۰۲۳) از مقادیر SHAP برای شناسایی و تبیین ویژگی های تأثیرگذار در پیش بینی و تحلیل اهمیت ویژگی بهره برده است. بعضی از مطالعات نیز تکنیک های دیگری را استفاده کرده اند. به عنوان مثال، در مطالعه (Al-Absi et al. ۲۰۲۲)، از تکنیک GradCAM به عنوان ابزاری برای تفسیر مدل های DL و شناسایی ویژگی های مهم استفاده شده است.

- تکنیک های همجوشی و تحلیل چند سطحی

تجزیه و تحلیل چند سطحی با ادغام اطلاعات از منابع متعدد، دقت طبقه بندی را افزایش می دهد. استراتژی های ادغام در سطوح مختلف - مانند ادغام در داده ها، ویژگی ها و تصمیم گیری - برای استفاده از اطلاعات مکمل استفاده شده اند. ترکیب داده ها و ویژگی ها از منابع متعدد (به عنوان مثال، سیگنال های ECG) به طور قابل توجهی دقت پیش بینی را بهبود می بخشد. به عنوان مثال، برای یکپارچه سازی ویژگی، مطالعه (Guo et al. ۲۰۲۰) از مدل های خطی و تکنیک RF جهت افزایش عملکرد پیش بینی استفاده کرده است. به طور مشابه، مطالعه (Shi et al. ۲۰۲۱) تحلیل همبستگی متعارف تشخیصی (DCCA) را همراه با استخراج ویژگی ها از طریق شبکه های ResNet و GRU بکار گرفته است. علاوه بر این، مدل های مبتنی بر CNN که با تکنیک های مهندسی ویژگی ها ادغام شده اند، به دقت بالایی دست یافته اند، مانند مطالعه (Zhanquan et al. ۲۰۲۲) که دقت طبقه بندی ۹۸٫۴۱ درصدی را با استفاده از ترکیبی از ویژگی های استخراج شده به صورت دستی و مبتنی بر DL را گزارش داده است (Zhanquan et al. ۲۰۲۲). یادگیری چندوجهی یکی دیگر از راهبردهای مؤثر برای بهبود عملکرد پیش بینی کننده با ترکیب روش های مختلف داده

است. به عنوان مثال، مطالعه (Al-Absi et al. ۲۰۲۲) از یک رویکرد یادگیری چند وجهی جهت ادغام تصاویر شبکه‌ای و DXA استفاده کرده که منجر به افزایش دقت و قابلیت تفسیر مدل شده است. با تکنیک‌های یادگیری چند وجهی، می‌توان منابع داده ناهمگن را برای پیش‌بینی بهتر ترکیب کرد. به عنوان مثال، تکنیک تجزیه حالت تجربی که عمدتاً یک تکنیک پردازش سیگنال برای تجزیه و تحلیل سیگنال‌های غیر ثابت، مانند داده‌های ECG است، می‌تواند برای استخراج و طبقه‌بندی ویژگی‌های پیچیده استفاده شود. اگرچه EMD به طور مستقیم به عنوان یک پارادایم یادگیری چند وجهی استفاده نمی‌شود، اما می‌تواند مدل‌های چند وجهی را با افزایش کیفیت ویژگی استخراج شده توسط ECG قبل از ادغام با روش‌های اضافی، مانند تاریخچه بالینی یا داده‌های تصویری، تسهیل کند. با بهبود سیگنال ECG و عملکرد طبقه‌بندی بالایی، EMD می‌تواند به عنوان یک ابزار پیش‌پردازش در پارادایم‌های یادگیری چند وجهی استفاده شود که در آن اطلاعات از منابع مختلف ترکیب می‌شوند تا تشخیص قوی CVD را امکان‌پذیر کند (Lilda et al. ۲۰۲۲). علاوه بر این تکنیک‌ها، تکنیک‌های افزایش داده‌ها به رفع مشکلات داده‌های از دست رفته کمک می‌کند. مطالعه (Kokol et al. ۲۰۲۱) استفاده از تولید داده‌های مصنوعی را به عنوان روشی برای جبران سوابق از دست رفته، اطمینان از کامل بودن داده‌ها و استحکام مدل، اهمیت این تکنیک را برجسته کرده است.

بحث و نتیجه گیری

نتایج این مطالعه نشان می دهد که مدل های یادگیری گروهی (EL) نقش کلیدی در مدیریت بیماری های قلبی عروقی دارند و در چهار طبقه اصلی شامل پیش بینی، تشخیص، شناسایی و طبقه بندی بیماری ها مورد استفاده قرار می گیرند. یافته ها نشان می دهد که ترکیب مدل های EL با یادگیری ماشین (ML) و یادگیری عمیق (DL) می تواند دقت و کارایی تشخیصی را بهبود بخشد. به طور خاص، مدل های جنگل تصادفی (RF) و روش های تقویت کننده مانند XGBoost به دلیل قابلیت بالای آن ها در مدیریت داده های نامتوازن، مقاومت در برابر نویز، جلوگیری از بیش برآزش و توانایی ترکیب تصمیم گیری چند درخت در قالب یادگیری گروهی، دقت بالاتری نسبت به مدل های منفرد به دست آورده اند و از موفق ترین روش های مورد استفاده در مطالعات مرتبط بوده اند. با استفاده از مجموعه داده های بزرگ و متنوع، این رویکردها امکان شناسایی الگوهای پیچیده و ارائه پیش بینی های دقیق تر را فراهم می کنند که می تواند منجر به بهبود فرآیندهای تصمیم گیری بالینی و کاهش خطاهای تشخیصی شود. همچنین نقش DL در دو دسته شناسایی و طبقه بندی بیماری های قلبی عروقی نسبت به دو دسته دیگر مشهودتر بود. دلیل این امر پیچیدگی داده ها و سیگنال ها از جمله ECG ها، تصاویر و صداهای قلب است که در شناسایی و طبقه بندی CVD ها استفاده می شود. معماری های مختلف DL در این دو دسته بسیار قدرتمندتر از الگوریتم های ML و مدل های گروهی ظاهر شده اند و به دقت و عملکرد بسیار بهتری دست یافته اند. یافته ها نشان می دهد که در این دو مقوله بیشترین الگوریتم هایی که با مدل های DL ترکیب شده اند شامل CNN ها و LSTM ها نتایج خوبی داشته اند، مدل های RF و XGB هستند که نشان دهنده درهم تنیدگی مدل های DL با EL است که روند نوظهوری را نشان می دهد.

یکی دیگر از ترکیب های مهم در مطالعات بررسی شده، نقش پیش پردازش داده ها و مهندسی ویژگی است که تکنیک های آن نقش اساسی در بهبود عملکرد، دقت، قابلیت اطمینان و تعمیم پذیری مدل های مورد استفاده در مدیریت CVD ها دارد. تکنیک های حذف نویز و داده های پرت، متعادل کردن مجموعه داده ها و کلاس ها و قابلیت تفسیر مدل

به این پیشرفت ها کمک کرده است. در این میان ، نقش تکنیک هایی مانند SMOTE و SHAP بسیار مشهود است.

علاوه بر این، این مطالعه به محدودیت های مدل های EL نیز پرداخته است. بسیاری از مطالعات بررسی شده از مجموعه داده های کوچک و غیرمتوازن استفاده کرده اند که می تواند بر استحکام آماری و تعمیم پذیری مدل ها تأثیر منفی بگذارد. کیفیت داده ها، از جمله وجود نویز در سیگنال های ECG یا تصاویر پزشکی با وضوح پایین، نیز یکی از چالش های رایج بوده است. علاوه بر این، فقدان داده های دموگرافیک و بالینی کامل و نبود اعتبارسنجی خارجی در برخی مطالعات، ارزیابی عملکرد مدل ها را محدود می کند. از منظر فنی، مدل های یادگیری عمیق و گروهی اگرچه دقت بالایی دارند، اما به منابع محاسباتی زیاد نیاز دارند و اغلب فاقد قابلیت تفسیر کافی هستند که این امر بر پذیرش بالینی آن ها تأثیر می گذارد. همچنین، در بسیاری از موارد، مدل ها بر روی جمعیت های خاص آموزش داده شده اند و در نتیجه ممکن است در محیط های دنیای واقعی یا داده های چندمرکزی عملکرد مشابهی نداشته باشند. در نهایت، هزینه محاسباتی بالا، محدودیت در پردازش بلادرنگ، و شکاف در ارزیابی کاربردهای واقعی از دیگر چالش هایی است که باید در پژوهش های آینده مورد توجه قرار گیرد.

به طور کلی، یافته های این مطالعه می تواند بینش های ارزشمندی را برای پژوهشگران و متخصصان حوزه هوش مصنوعی و سلامت فراهم کند و مسیر پژوهش های آینده را در استفاده از مدل های ترکیبی یادگیری روشن نماید. با توجه به روندهای نوظهور در مطالعات اخیر، پیشنهاد می شود پژوهش های آتی بر توسعه چارچوب های ترکیبی هوشمند مبتنی بر تلفیق مدل های ML، DL و EL متمرکز شوند که بتوانند به صورت تطبیقی و خودکار، نوع داده های ورودی (مانند سیگنال های ECG، تصاویر پزشکی، و داده های بالینی چندوجهی) را شناسایی و مناسب ترین الگوریتم را برای تحلیل آن انتخاب کنند. به طور خاص، پژوهش های آینده می توانند به طراحی مدل های هیبریدی مبتنی بر CNN-RF، LSTM-XGBoost یا AutoML-EL بپردازند که علاوه بر بهبود دقت، قابلیت تفسیر بالینی، کارایی محاسباتی، و پردازش بلادرنگ را افزایش دهند. همچنین، توسعه ی این چارچوب ها بر روی مجموعه داده های چندمرکزی و ناهمگون می تواند گامی مهم در جهت افزایش تعمیم پذیری و قابلیت کاربرد در محیط های واقعی سلامت دیجیتال باشد.

در مسیر تحقیقاتی آینده، بهره‌گیری از هوش مصنوعی مولد (Generative AI) و هوش مصنوعی عاملی (Agentic AI) می‌تواند مسیر نوینی در ترکیب با مدل‌های یادگیری گروهی برای مدیریت CVDها فراهم کند. GenAI با توانایی تولید داده‌های واقع‌گرایانه و رفع عدم توازن مجموعه داده‌ها، می‌تواند عملکرد مدل‌های ترکیبی را ارتقا دهد. در مقابل، Agentic AI با قابلیت تصمیم‌گیری تطبیقی و تعامل پویا با داده‌های بالینی، می‌تواند نقش مؤثری در توسعه‌ی سامانه‌های تشخیص و نظارت بلادرنگ ایفا کند. ادغام این رویکردها با چارچوب‌های EL، زمینه‌ساز نسل جدیدی از سیستم‌های هوشمند سلامت مبتنی بر یادگیری خودکار و داده‌محور خواهد بود.

تعارض منافع

نویسندگان اعلام می‌دارند هیچ تعارض منفعی در رابطه با نویسندگی و یا انتشار این مقاله وجود ندارد.

ORCID

Mahmoud Zahedian		https://orcid.org/0000-0001-6466-9828
Nezhad		
Mohammad Mehraeen		http://orcid.org/0000-0002-4154-8975
Rouhollah Bagheri		https://orcid.org/0000-0003-3154-7466
Seyyed Mohammad Tabatabaei		https://orcid.org/0000-0002-3153-8968

١. Aarthy, S. T., and J. L. Mazher Iqbal. ٢٠٢٤. "A Novel Deep Learning Approach for Early Detection of Cardiovascular Diseases from ECG Signals." *MEDICAL ENGINEERING & PHYSICS* ١٢٥. doi: ١٠.١٠١٦/j.medengphy.٢٠٢٤.١٠٤١١١.
٢. Abdullah, Azween, S. Nithya, M. Mary Shanthi Rani, S. Vijayalakshmi, and Balamurugan Balusamy. ٢٠٢٣. "Stacked LSTM and Kernel-PCA-Based Ensemble Learning for Cardiac Arrhythmia Classification." *INTERNATIONAL JOURNAL OF ADVANCED COMPUTER SCIENCE AND APPLICATIONS* ١٤(٩):٣٩-٤٨.
٣. Ahamed, Jameel, Abdul Manan Koli, Khaleel Ahmad, Mohd Alam Jamal, and B. B. Gupta. ٢٠٢٢. "CDPS-IoT: Cardiovascular Disease Prediction System Based on IoT Using Machine Learning." *INTERNATIONAL JOURNAL OF INTERACTIVE MULTIMEDIA AND ARTIFICIAL INTELLIGENCE* ٧(٤):٧٨-٨٦. doi: ١٠.٩٧٨١/ijimai.٢٠٢١.٠٩.٠٠٢.
٤. Ahamed, Jameel, Roohie Naaz Mir, and Mohammad Ahsan Chishti. ٢٠٢٢. "Industry ٤.٠ Oriented Predictive Analytics of Cardiovascular Diseases Using Machine Learning, Hyperparameter Tuning and Ensemble Techniques." *INDUSTRIAL ROBOT-THE INTERNATIONAL JOURNAL OF ROBOTICS RESEARCH AND APPLICATION* ٤٩(٣, SI):٥٤٤-٥٤. doi: ١٠.١١٠٨/IR-١٠.٢٠٢١.٠٢٤٠.
٥. Ahmadian, Sajad, Seyed Mohammad Jafar Jalali, Saeid Raziani, and Abdollah Chalechale. ٢٠٢٢. "An Efficient Cardiovascular Disease Detection Model Based on Multilayer Perceptron and Moth-Flame Optimization." *EXPERT SYSTEMS* ٣٩(٤, SI). doi: ١٠.١١١١/exsy.١٢٩١٤.
٦. Ahn, Imjin, Hansle Gwon, Heejun Kang, Yunha Kim, Hyeram Seo, Heejung Choi, Ha Na Cho, Minkyung Kim, Tae Joon Jun, and Young-Hak Kim. ٢٠٢١. "Machine Learning-Based Hospital Discharge Prediction for Patients With Cardiovascular Diseases: Development and Usability Study." *JMIR MEDICAL INFORMATICS* ٩(١١). doi: ١٠.٢١٩٦/٣٢٦٦٢.
٧. Al-Absi, Hamada R. H., Mohammad Tariqul Islam, Mahmoud Ahmed Refaee, Muhammad E. H. Chowdhury, and Tanvir Alam. ٢٠٢٢. "Cardiovascular Disease Diagnosis from DXA Scan and Retinal Images Using Deep Learning." *SENSORS* ٢٢(١٢). doi: ١٠.٣٣٩٠/s٢٢١٢٤٣١٠.
٨. Al-Yarimi, Fuad Ali Mohammed. ٢٠٢٢. "Handling High Dimensionality in Ensemble Learning for Arrhythmia Prediction." *INTELLIGENT AUTOMATION AND SOFT COMPUTING* ٣٢(٣):١٧٢٩-٤٢. doi: ١٠.٣٢٦٠٤/iasc.٢٠٢٢.٠٢٢٤١٨.
٩. Alfaidi, Aseel, Reem Aljuhani, Bushra Alshehri, Hajer Alwadei, and Sahar Sabbeh. ٢٠٢٢. "Machine Learning: Assisted Cardiovascular Diseases Diagnosis." *INTERNATIONAL JOURNAL OF ADVANCED COMPUTER SCIENCE AND APPLICATIONS*

۱۳(۲):۱۳۵-۴۱.

۱۰. Ali, Md Mamun, Bikash Kumar Paul, Kawsar Ahmed, Francis M. Bui, Julian M. W. Quinn, and Mohammad Ali Moni. ۲۰۲۱. "Heart Disease Prediction Using Supervised Machine Learning Algorithms: Performance Analysis and Comparison." *COMPUTERS IN BIOLOGY AND MEDICINE* ۱۳۶. doi: ۱۰.۱۰۱۶/j.compbimed.۲۰۲۱.۱۰۴۶۷۲.
۱۱. Alotaibi, Aziz. ۲۰۲۰. "Ensemble Deep Learning Approaches in Health Care: A Review." *Computers, Materials & Continua* ۸۲(۳).
۱۲. Alreshidi, Fayez Saud, Mohammad Alsaffar, Rajeswari Chengoden, and Naif Khalaf Alshammari. ۲۰۲۴. "Fed-CL-an Atrial Fibrillation Prediction System Using ECG Signals Employing Federated Learning Mechanism." *Scientific Reports* ۱۴(۱):۲۱۰۳۸.
۱۳. An, Ying, Nengjun Huang, Xianlai Chen, FangXiang Wu, and Jianxin Wang. ۲۰۲۱. "High-Risk Prediction of Cardiovascular Diseases via Attention-Based Deep Neural Networks." *IEEE-ACM TRANSACTIONS ON COMPUTATIONAL BIOLOGY AND BIOINFORMATICS* ۱۸(۳):۱۰۹۳-۱۱۰۵. doi: ۱۰.۱۱۰۹/TCBB.۲۰۱۹.۲۹۳۵.۰۵۹.
۱۴. Anbumani, Menaga, and Kannan Kaniyaiah. ۲۰۲۴. "Soft Set-Based Parameter Reduction Algorithm Through a Discernibility Matrix and the Hybrid Approach for the Risk-Factor Prediction of Cardiovascular Diseases by Various Machine Learning Techniques." *PERTANIKA JOURNAL OF SCIENCE AND TECHNOLOGY* ۳۲(۱):۲۶۵-۸۴. doi: ۱۰.۴۷۸۳۶/pjst.۳۲,۱,۱۶.
۱۵. Ansari, Yaqoob, Omar Mourad, Khalid Qaraqe, and Erchin Serpedin. ۲۰۲۳. "Deep Learning for ECG Arrhythmia Detection and Classification: An Overview of Progress for Period ۲۰۱۷-۲۰۲۳." *FRONTIERS IN PHYSIOLOGY* ۱۴. doi: ۱۰.۳۳۸۹/fphys.۲۰۲۳.۱۲۴۶۷۴۶.
۱۶. Ardeti, Venkata Anuhya, Venkata Ratnam Kolluru, George Tom Varghese, and Rajesh Kumar Patjoshi. ۲۰۲۲. "An Outlier Detection and Feature Ranking Based Ensemble Learning for ECG Analysis." *INTERNATIONAL JOURNAL OF ADVANCED COMPUTER SCIENCE AND APPLICATIONS* ۱۳(۶):۷۲۷-۳۷.
۱۷. Azmi, Javed, Muhammad Arif, Md Tabrez Nafis, M. Afshar Alam, Safdar Tanweer, and Guojun Wang. ۲۰۲۲. "A Systematic Review on Machine Learning Approaches for Cardiovascular Disease Prediction Using Medical Big Data." *MEDICAL ENGINEERING & PHYSICS* ۱۰۵. doi: ۱۰.۱۰۱۶/j.medengphy.۲۰۲۲.۱۰۳۸۲۵.
۱۸. Baghdadi, Nadiah A., Sally Mohammed Farghaly Abdelaliem, Amer Malki, Ibrahim Gad, Ashraf Ewis, and Elsayed Atlam. ۲۰۲۳. "Advanced Machine Learning Techniques for Cardiovascular Disease Early Detection and Diagnosis." *JOURNAL OF BIG DATA* ۱۰(۱). doi: ۱۰.۱۱۸۶/s۴۰۵۳۷-۰۲۳-۰۰۸۱۷-۱.
۱۹. Balaha, Hossam Magdy, Ahmed Osama Shaban, Eman M. El-Gendy, and Mahmoud M. Saafan. ۲۰۲۲. "A Multi-Variate Heart Disease Optimization and Recognition

- Framework.” *NEURAL COMPUTING & APPLICATIONS* 33(18, SI):10907–22. doi: 10.1007/s00521-022-07241-1.
20. Barriada, Ruben G., Olga Simo-Servat, Alejandra Planas, Cristina Hernandez, Rafael Simo, and David Masip. 2022. “Deep Learning of Retinal Imaging: A Useful Tool for Coronary Artery Calcium Score Prediction in Diabetic Patients.” *APPLIED SCIENCES-BASEL* 12(3). doi: 10.3390/app12031401.
21. Baseer, K. K., K. Sivakumar, Duggineni Veeraiah, Gunjan Chhabra, Prasanna Kumar Lakineni, M. Jahir Pasha, Ramu Gandikota, and Gopakumar Harikrishnan. 2022. “Healthcare Diagnostics with an Adaptive Deep Learning Model Integrated with the Internet of Medical Things (IoMT) for Predicting Heart Disease.” *BIOMEDICAL SIGNAL PROCESSING AND CONTROL* 92. doi: 10.1016/j.bspc.2022.109888.
22. Bertsimas, Dimitris, Alison Rose Ann Borenstein, Antonin Dauvin, and Agni Orfanoudaki. 2022. “Ensemble Machine Learning for Personalized Antihypertensive Treatment.” *NAVAL RESEARCH LOGISTICS* 69(5):669–88. doi: 10.1002/nav.22040.
23. Bhatt, Chintan M., Parth Patel, Tarang Ghetia, and Pier Luigi Mazzeo. 2023. “Effective Heart Disease Prediction Using Machine Learning Techniques.” *ALGORITHMS* 16(2). doi: 10.3390/a16020088.
24. Boulkaboul, Sahar, Samira Bouchama, Syphax Kasser, and Belkacem Ait Si Ali. 2022. “D-Resnet: Deep Resnet Based Approach for ECG Classification.” *CONTROL ENGINEERING AND APPLIED INFORMATICS* 26(1):64–71.
25. Cao, Kerang, Chang Liu, Siqi Yang, Yuxin Zhang, Lili Li, Hoekyung Jung, and Shuo Zhang. 2020. “Prediction of Cardiovascular Disease Based on Multiple Feature Selection and Improved PSO-XGBoost Model.” *Scientific Reports* 10(1):12406.
26. Celermajer, David S., Clara K. Chow, Eloi Marijon, Nicholas M. Anstey, and Kam S. Woo. 2012. “Cardiovascular Disease in the Developing World: Prevalences, Patterns, and the Potential of Early Disease Detection.” *Journal of the American College of Cardiology* 60(14):1207–16.
27. Chalapathi, Mukkoti Maruthi venkata, Dudekula Khasim Vali, Yellapragada Venkata Pavan Kumar, Challa Pradeep Reddy, and Purna Prakash Kasaraneni. 2022. “AN INSIGHT INTO VIABLE MACHINE LEARNING MODELS FOR EARLY DIAGNOSIS OF CARDIOVASCULAR DISEASE.” *SCALABLE COMPUTING-PRACTICE AND EXPERIENCE* 20(1, SI). doi: 10.12694/scpe.v20i1.2326.
28. Chen, Liang, Zhijun Han, Junhong Wang, and Chengjian Yang. 2022. “The Emerging Roles of Machine Learning in Cardiovascular Diseases: A Narrative Review.” *ANNALS OF TRANSLATIONAL MEDICINE* 10(10). doi: 10.21037/atm-22-1803.
29. Chen, Shuwen, Jiaming Qi, Shicheng Fan, Zheng Qiao, Joo Chuan Yeo, and Chwee Teck Lim. 2021. “Flexible Wearable Sensors for Cardiovascular Health Monitoring.” *Advanced Healthcare Materials* 10(17):2100116.
30. Chicco, Davide, and Giuseppe Jurman. 2020. “Machine Learning Can Predict Survival

- of Patients with Heart Failure from Serum Creatinine and Ejection Fraction Alone.” *BMC MEDICAL INFORMATICS AND DECISION MAKING* ۲۰(۱). doi: ۱۰,۱۱۸۶/s۱۲۹۱۱-۰۲۰-۱۰۲۳-۵.
۳۱. Chorney, Wesley, and Sing Hui Ling. ۲۰۲۵. “Federated Learning Strategies for Atrial Fibrillation Detection.” *Journal of Experimental and Theoretical Analyses* ۳(۳):۲۳.
۳۲. Chun, Matthew, Robert Clarke, Benjamin J. Cairns, David Clifton, Derrick Bennett, Yiping Chen, Yu Guo, Pei Pei, Jun Lv, Canqing Yu, Ling Yang, Liming Li, Zhengming Chen, Tingting Zhu, and China Kadoorie Biobank Collaborati. ۲۰۲۱. “Stroke Risk Prediction Using Machine Learning: A Prospective Cohort Study of ۰,۵ Million Chinese Adults.” *JOURNAL OF THE AMERICAN MEDICAL INFORMATICS ASSOCIATION* ۲۸(۸):۱۷۱۹-۲۷. doi: ۱۰,۱۰۹۳/jamia/ocab۰۶۸.
۳۳. Cocianu, Catalina-Lucia, Cristian Razvan Uscatu, Konstantinos Kofidis, Sorin Muraru, and Alin Gabriel Vaduva. ۲۰۲۳. “Classical, Evolutionary, and Deep Learning Approaches of Automated Heart Disease Prediction: A Case Study.” *ELECTRONICS* ۱۲(۷). doi: ۱۰,۳۳۹۰/electronics۱۲۰۷۱۶۶۳.
۳۴. Daghistani, Tahani A., Radwa Elshawy, Sherif Sakr, Amjad M. Ahmed, Abdullah Al-Thwayee, and Mouaz H. Al-Mallah. ۲۰۱۹. “Predictors of In-Hospital Length of Stay among Cardiac Patients: A Machine Learning Approach.” *INTERNATIONAL JOURNAL OF CARDIOLOGY* ۲۸۸:۱۴۰-۴۷. doi: ۱۰,۱۰۱۶/j.ijcard.۲۰۱۹,۰۱,۰۴۶.
۳۵. Datta, Suparno, Ariane Morassi Sasso, Nina Kiwit, Subhronil Bose, Girish Nadkarni, Riccardo Miotto, and Erwin P. Boettinger. ۲۰۲۲. “Predicting Hypertension Onset from Longitudinal Electronic Health Records with Deep Learning.” *JAMIA OPEN* ۵(۴). doi: ۱۰,۱۰۹۳/jamiaopen/ooac۰۹۷.
۳۶. Desai, Forum, Deepraj Chowdhury, Rupinder Kaur, Marloes Peeters, Rajesh Chand Arya, Gurpreet Singh Wander, Sukhpal Singh Gill, and Rajkumar Buyya. ۲۰۲۲. “HealthCloud: A System for Monitoring Health Status of Heart Patients Using Machine Learning and Cloud Computing.” *INTERNET OF THINGS* ۱۷. doi: ۱۰,۱۰۱۶/j.iot.۲۰۲۱,۱۰,۰۴۸۵.
۳۷. Dhanka, Sanjay, Vibhor Kumar Bhardwaj, and Surita Maini. ۲۰۲۳. “Comprehensive Analysis of Supervised Algorithms for Coronary Artery Heart Disease Detection.” *EXPERT SYSTEMS* ۴۰(۷). doi: ۱۰,۱۱۱۱/exsy.۱۳۳۰۰.
۳۸. Dinh, An, Stacey Miertschin, Amber Young, and Somya D. Mohanty. ۲۰۱۹. “A Data-Driven Approach to Predicting Diabetes and Cardiovascular Disease with Machine Learning.” *BMC MEDICAL INFORMATICS AND DECISION MAKING* ۱۹(۱). doi: ۱۰,۱۱۸۶/s۱۲۹۱۱-۰۱۹-۰۹۱۸-۵.
۳۹. Divya, R., Finney Daniel Shadrach, and S. Padmaja. ۲۰۲۳. “Cardiovascular Risk Detection Using Harris Hawks Optimization with Ensemble Learning Model on PPG Signals.” *SIGNAL IMAGE AND VIDEO PROCESSING* ۱۷(۸):۴۵۰۳-۱۲. doi: ۱۰,۱۰۰۷/s۱۱۷۶۰-۰۲۳-۰۲۶۸۴-y.
۴۰. Dritsas, Elias, and Maria Trigka. ۲۰۲۳. “Efficient Data-Driven Machine Learning

- Models for Cardiovascular Diseases Risk Prediction.” *SENSORS* 22(3). doi: 10.3390/s22031161.
- ξ1. Du, Jinsong, Xiao Chang, Chunhong Ye, Yijun Zeng, Sijia Yang, Shan Wu, and Li Li. 2023. “Developing a Hypertension Visualization Risk Prediction System Utilizing Machine Learning and Health Check-up Data.” *SCIENTIFIC REPORTS* 13(1). doi: 10.1038/s41598-023-26281-y.
 - ξ2. Dubey, Animesh Kumar, Amit Kumar Sinhal, and Richa Sharma. 2022. “An Improved Auto Categorical PSO with ML for Heart Disease Prediction.” *ENGINEERING TECHNOLOGY & APPLIED SCIENCE RESEARCH* 12(3):8067-73.
 - ξ3. Eleyan, Alaa, and Ebrahim Alboghbaish. 2023. “Electrocardiogram Signals Classification Using Deep-Learning-Based Incorporated Convolutional Neural Network and Long Short-Term Memory Framework.” *COMPUTERS* 13(2). doi: 10.3390/computers1302000.
 - ξ4. Elvas, Luis B., Miguel Nunes, Joao C. Ferreira, Miguel Sales Dias, and Luis Bras Rosario. 2023. “AI-Driven Decision Support for Early Detection of Cardiac Events: Unveiling Patterns and Predicting Myocardial Ischemia.” *JOURNAL OF PERSONALIZED MEDICINE* 13(9). doi: 10.3390/jpm13091421.
 - ξ5. Enad, Huda Ghazi, and Mazin Abed Mohammed. 2023. “Cloud Computing-Based Framework for Heart Disease Classification Using Quantum Machine Learning Approach.” *JOURNAL OF INTELLIGENT SYSTEMS* 33(1). doi: 10.1010/jisys-2023-0261.
 - ξ6. Essa, Ehab, and Xianghua Xie. 2021. “An Ensemble of Deep Learning-Based Multi-Model for ECG Heartbeats Arrhythmia Classification.” *IEEE ACCESS* 9:103402-74. doi: 10.1109/ACCESS.2021.3098986.
 - ξ7. Ghosh, Pronab, Sami Azam, Mirjam Jonkman, Asif Karim, F. M. Javed Mehedi Shamrat, Eva Ignatious, Shahana Shultana, Abhijith Reddy Beeravolu, and Friso De Boer. 2021. “Efficient Prediction of Cardiovascular Disease Using Machine Learning Algorithms With Relief and LASSO Feature Selection Techniques.” *IEEE ACCESS* 9:19304-26. doi: 10.1109/ACCESS.2021.303709.
 - ξ8. Gour, Sanjeev, Rajendra Randa, Karuna Nidhi Pandagre, Arpit Neema, Sangeeta Jat, and Deepak Malakar. 2020. “A Comprehensive Review on Predicting Heart Failure with Ensemble Learning: Integrating Environmental Variables and Health Perspectives.” *International Journal of Environmental Sciences* 11(ξs):909-73.
 - ξ9. Guarneros-Nolasco, Luis Rolando, Nancy Aracely Cruz-Ramos, Giner Alor-Hernandez, Lisbeth Rodriguez-Mazahua, and Jose Luis Sanchez-Cervantes. 2021. “Identifying the Main Risk Factors for Cardiovascular Diseases Prediction Using Machine Learning Algorithms.” *MATHEMATICS* 9(20). doi: 10.3390/math9202037.
 - ο0. Guo, Chunyan, Jiabing Zhang, Yang Liu, Yaying Xie, Zhiqiang Han, and Jianshe Yu. 2020. “Recursion Enhanced Random Forest With an Improved Linear Model (RERF-ILM) for Heart Disease Detection on the Internet of Medical Things Platform.” *IEEE*

- ACCESS ۸:۵۹۲۴۷-۵۶. doi: ۱۰,۱۱۰۹/ACCESS.۲۰۲۰,۲۹۸۱۱۵۹.
۵۱. Gutierrez-Esparza, Guadalupe, Tomas Pulido, Mireya Martinez-Garcia, Tania Ramirez-delReal, Lucero E. Groves-Miralrio, Manlio F. Marquez-Murillo, Luis M. Amezcua-Guerra, Gilberto Vargas-Alarcon, and Enrique Hernandez-Lemus. ۲۰۲۳. "A Machine Learning Approach to Personalized Predictors of Dyslipidemia: A Cohort Study." *FRONTIERS IN PUBLIC HEALTH* ۱۱. doi: ۱۰,۳۳۸۹/fpubh.۲۰۲۳,۱۲۱۳۹۲۶.
 ۵۲. Hassaballah, Mahmoud, Yaser M. M. Wazery, Ibrahim E. E. Ibrahim, and Aly Farag. ۲۰۲۳. "ECG Heartbeat Classification Using Machine Learning and Metaheuristic Optimization for Smart Healthcare Systems." *BIOENGINEERING-BASEL* ۱۰(۴). doi: ۱۰,۳۳۹۰/bioengineering۱۰۰۴۰۴۲۹.
 ۵۳. Hassan, Md. Mehedi, Sadika Zaman, Md. Mushfiqur Rahman, Anupam Kumar Bairagi, Walid El-Shafai, Rajkumar Singh Rathore, and Deepak Gupta. ۲۰۲۴. "Efficient Prediction of Coronary Artery Disease Using Machine Learning Algorithms with Feature Selection Techniques." *COMPUTERS & ELECTRICAL ENGINEERING* ۱۱۵. doi: ۱۰,۱۰۱۶/j.compeleceng.۲۰۲۴,۱۰۹۱۳۰.
 ۵۴. He, Yi, Wuyou Li, Wangqi Zhang, Sheng Zhang, Xitian Pi, and Hongying Liu. ۲۰۲۱. "Research on Segmentation and Classification of Heart Sound Signals Based on Deep Learning." *APPLIED SCIENCES-BASEL* ۱۱(۲). doi: ۱۰,۳۳۹۰/app۱۱۰۲۰۶۵۱.
 ۵۵. Henni, Fouad, Baghdad Atmani, Fatiha Atmani, and Fatima Saadi. ۲۰۲۳. "Improving Coronary Artery Disease Prediction: Use of Random Forest, Feature Importance and Case-Based Reasoning." *INTERNATIONAL JOURNAL OF DECISION SUPPORT SYSTEM TECHNOLOGY* ۱۵(۱). doi: ۱۰,۴۰۱۸/ijdsst.۳۱۹۳۰۷.
 ۵۶. Hosni, Mohamed, Juan M. de Gea, Ali Idri, Manal El Bajta, Jose Luis Fernandez Aleman, Gines Garcia-Mateos, and Ittissam Abnane. ۲۰۲۱. "A Systematic Mapping Study for Ensemble Classification Methods in Cardiovascular Disease." *ARTIFICIAL INTELLIGENCE REVIEW* ۵۴(۴):۲۸۲۷-۶۱. doi: ۱۰,۱۰۰۷/s۱۰۴۶۲-۰۲۰-۰۹۹۱۴۶.
 ۵۷. Hossain, Ekramul, Shahadat Uddin, and Arif Khan. ۲۰۲۱. "Network Analytics and Machine Learning for Predictive Risk Modelling of Cardiovascular Disease in Patients with Type ۲ Diabetes." *EXPERT SYSTEMS WITH APPLICATIONS* ۱۶۴. doi: ۱۰,۱۰۱۶/j.eswa.۲۰۲۰,۱۱۳۹۱۸.
 ۵۸. Hsu, Wei-Yen. ۲۰۱۸. "A Decision-Making Mechanism for Assessing Risk Factor Significance in Cardiovascular Diseases." *DECISION SUPPORT SYSTEMS* ۱۱۵:۶۴-۷۷. doi: ۱۰,۱۰۱۶/j.dss.۲۰۱۸,۰۹,۰۰۴.
 ۵۹. Hu, Zhixu, Hang Qiu, Ziqi Su, Minghui Shen, and Ziyu Chen. ۲۰۲۰. "A Stacking Ensemble Model to Predict Daily Number of Hospital Admissions for Cardiovascular Diseases." *IEEE ACCESS* ۸:۱۳۸۷۱۹-۲۹. doi: ۱۰,۱۱۰۹/ACCESS.۲۰۲۰,۳۰۱۲۱۴۳.
 ۶۰. Husain, Mohammad, Pankaj Kumar, Mohammad Nadeem Ahmed, Arshad Ali, Mohammad Ashiquee Rasool, Mohammad Rashid Hussain, and Muhammad Shahid

- Dildar. ۲۰۲۳. "Harnessing Ensemble in Machine Learning for Accurate Early Prediction and Prevention of Heart Disease." *INTERNATIONAL JOURNAL OF ADVANCED COMPUTER SCIENCE AND APPLICATIONS* ۱۴(۱۰):۱۸۲-۹۵.
۶۱. Ibarra, Rodrigo, Jaime Leon, Ivan Avila, and Hiram Ponce. ۲۰۲۲. "Cardiovascular Disease Detection Using Machine Learning." *COMPUTACION Y SISTEMAS* ۲۶(۴):۱۶۶۱-۶۸. doi: ۱۰.۱۳۰۵۳/CyS-۲۶-۴-۴۴۲۲.
۶۲. Islam, Md. Merajul, Md. Jahangir Alam, Md Maniruzzaman, N. A. M. Faisal Ahmed, Md Sujun Ali, Md. Jahanur Rahman, and Dulal Chandra Roy. ۲۰۲۳. "Predicting the Risk of Hypertension Using Machine Learning Algorithms: A Cross Sectional Study in Ethiopia." *PLOS ONE* ۱۸(۸). doi: ۱۰.۱۳۷۱/journal.pone.۰۲۸۹۶۱۳.
۶۳. Islam, Sheikh Mohammed Shariful, Ashis Talukder, Md. Abdul Awal, Md. Muhammad Umer Siddiqui, Md. Martuza Ahamad, Benojir Ahammed, Lal B. Rawal, Roohallah Alizadehsani, Jemal Abawajy, Liliana Laranjo, Clara K. Chow, and Ralph Maddison. ۲۰۲۲. "Machine Learning Approaches for Predicting Hypertension and Its Associated Factors Using Population-Level Data From Three South Asian Countries." *FRONTIERS IN CARDIOVASCULAR MEDICINE* ۹. doi: ۱۰.۳۳۸۹/fcvm.۲۰۲۲.۸۳۹۳۷۹.
۶۴. Jentzer, Jacob C., Yogesh N. V Reddy, Sabri Soussi, Ruben Crespo-Diaz, Parag C. Patel, Patrick R. Lawler, Alexandre Mebazaa, and Shannon M. Dunlay. ۲۰۲۴. "Unsupervised Machine Learning to Identify Subphenotypes among Cardiac Intensive Care Unit Patients with Heart Failure." *ESC Heart Failure* ۱۱(۶):۴۲۴۲-۵۶.
۶۵. Jiang, Lili, Sirong Chen, Yuanhui Wu, Da Zhou, and Lihua Duan. ۲۰۲۳. "Prediction of Coronary Heart Disease in Gout Patients Using Machine Learning Models." *MATHEMATICAL BIOSCIENCES AND ENGINEERING* ۲۰(۳):۴۵۷۴-۹۱. doi: ۱۰.۳۹۳۴/mbe.۲۰۲۳۲۰۲.
۶۶. Karboub, Kaouter, and Mohamed Tabaa. ۲۰۲۲. "A Machine Learning Based Discharge Prediction of Cardiovascular Diseases Patients in Intensive Care Units." *HEALTHCARE* ۱۰(۶). doi: ۱۰.۳۳۹۰/healthcare۱۰۰۶.۹۶۶.
۶۷. Karthikeyan, N. K., and others. ۲۰۲۴. "A Novel Attention-Based Cross-Modal Transfer Learning Framework for Predicting Cardiovascular Disease." *Computers in Biology and Medicine* ۱۷۰:۱۰۷۹۷۷.
۶۸. Kim, Joung Ouk (Ryan), Yong-Suk Jeong, Jin Ho Kim, Jong-Weon Lee, Dougho Park, and Hyoung-Seop Kim. ۲۰۲۱. "Machine Learning-Based Cardiovascular Disease Prediction Model: A Cohort Study on the Korean National Health Insurance Service Health Screening Database." *DIAGNOSTICS* ۱۱(۶). doi: ۱۰.۳۳۹۰/diagnostics۱۱۰۶.۹۴۳.
۶۹. Kirboga, Kevser Kubra, and Ecir Ugur Kucuksille. ۲۰۲۳. "Identifying Cardiovascular Disease Risk Factors in Adults with Explainable Artificial Intelligence." *ANATOLIAN JOURNAL OF CARDIOLOGY* ۲۷(۱۱):۶۵۷-۶۳. doi: ۱۰.۱۴۷۴۴/AnatolJCardiol.۲۰۲۳.۳۲۱۴.
۷۰. Kokol, Peter, Jan Jurman, Tajda Bogovic, Tadej Zavrsnik, Jernej Zavrsnik, and Helena

- Blazun Vosner. ۲۰۲۱. "Supporting Real World Decision Making in Coronary Diseases Using Machine Learning." *INQUIRY-THE JOURNAL OF HEALTH CARE ORGANIZATION PROVISION AND FINANCING* ۵۸. doi: ۱۰.۱۱۷۷/۰.۰۴۶۹۵۸.۲۱۹۹۷۳۳۸.
۷۱. Krittanawong, Chayakrit, Hafeez Ul Hassan Virk, Sripal Bangalore, Zhen Wang, Kipp W. Johnson, Rachel Pinotti, HongJu Zhang, Scott Kaplin, Bharat Narasimhan, Takeshi Kitai, Usman Baber, Jonathan L. Halperin, and W. H. Wilson Tang. ۲۰۲۰. "Machine Learning Prediction in Cardiovascular Diseases: A Meta-Analysis." *SCIENTIFIC REPORTS* ۱۰(۱). doi: ۱۰.۱۰۳۸/۴۱۵۹۸۰۲۰.۷۲۶۸۵۰۱.
۷۲. Kumar, Deepak, Chaman Verma, Sanjay Dahiya, Pradeep Kumar Singh, Maria Simona Raboaca, Zoltan Illes, and Brijesh Bakariya. ۲۰۲۱. "Cardiac Diagnostic Feature and Demographic Identification (CDF-DI): An IoT Enabled Healthcare Framework Using Machine Learning." *SENSORS* ۲۱(۱۹). doi: ۱۰.۳۳۹۰/s۲۱۱۹۶۵۸۴.
۷۳. Kumar, Pardeep, and Ankit Kumar. ۲۰۲۳. "Heart Disease Classification and Recommendation by Optimized Features and Adaptive Boost Learning." *INTERNATIONAL JOURNAL OF ADVANCED COMPUTER SCIENCE AND APPLICATIONS* ۱۴(۳):۹۰۹-۱۴.
۷۴. Lacson, Ronilda C., Bowen Baker, Harini Suresh, Katherine Andriole, Peter Szolovits, and Eduardo Lacson Jr. ۲۰۱۹. "Use of Machine-Learning Algorithms to Determine Features of Systolic Blood Pressure Variability That Predict Poor Outcomes in Hypertensive Patients." *CLINICAL KIDNEY JOURNAL* ۱۲(۲):۲۰۶-۱۲. doi: ۱۰.۱۰۹۳/ckj/sfy.۰۴۹.
۷۵. Lambay, Muhib Anwar, and S. Pakkır Mohideen. ۲۰۲۴. "Applying Data Science Approach to Predicting Diseases and Recommending Drugs in Healthcare Using Machine Learning Models - A Cardio Disease Case Study." *MULTIMEDIA TOOLS AND APPLICATIONS*. doi: ۱۰.۱۰۰۷/s۱۱۰۴۲۰۲۳۰۱۸۰۳۵۰.
۷۶. Latha, S., P. Muthu, Khin Wee Lai, Azira Khalil, and Samiappan Dhanalakshmi. ۲۰۲۲. "Performance Analysis of Machine Learning and Deep Learning Architectures on Early Stroke Detection Using Carotid Artery Ultrasound Images." *FRONTIERS IN AGING NEUROSCIENCE* ۱۳. doi: ۱۰.۳۳۸۹/fnagi.۲۰۲۱.۸۲۸۲۱۴.
۷۷. Lee, Heechang, Taeyoung Yoon, Chaeyun Yeo, HyeonYoung Oh, Yebin Ji, Seongwoo Sim, and Daesung Kang. ۲۰۲۱. "Cardiac Arrhythmia Classification Based on One-Dimensional Morphological Features." *APPLIED SCIENCES-BASEL* ۱۱(۲۰). doi: ۱۰.۳۳۹۰/app۱۱۲۰۹۴۶۰.
۷۸. Lee, Joonyub, Yera Choi, Taehoon Ko, Kanghyuck Lee, Juyoung Shin, and Hun-Sung Kim. ۲۰۲۴. "Prediction of Cardiovascular Complication in Patients with Newly Diagnosed Type ۲ Diabetes Using an XGBoost/ GRU-ODE-Bayes-Based Machine-Learning Algorithm." *ENDOCRINOLOGY AND METABOLISM* ۳۹(۱):۱۷۶-۸۵. doi: ۱۰.۳۸۰۳/EnM.۲۰۲۳.۱۷۳۹.
۷۹. Lee, Woojoo, Youn-Hee Lim, Eunhee Ha, Yeonjin Kim, and Won Kyung Lee. ۲۰۲۲.

“Forecasting of Non-Accidental, Cardiovascular, and Respiratory Mortality with Environmental Exposures Adopting Machine Learning Approaches (Vol 14, Pg 1, 2022).” *ENVIRONMENTAL SCIENCE AND POLLUTION RESEARCH* 29(08):8833. doi: 10.1007/s11356-022-2222-4.

80. Lertampaiporn, Supatcha, Apiradee Hongsthong, Warin Wattanapornprom, and Chinae Thammarongtham. 2022. “Ensemble-AHTPpred: A Robust Ensemble Machine Learning Model Integrated With a New Composite Feature for Identifying Antihypertensive Peptides.” *FRONTIERS IN GENETICS* 13. doi: 10.3389/fgene.2022.883766.
81. Li, Qingqing, Wenhui Xie, Liping Li, Lijing Wang, Qinyi You, Lu Chen, Jing Li, Yilang Ke, Jun Fang, Libin Liu, and Huashan Hong. 2021. “Development and Validation of a Prediction Model for Elevated Arterial Stiffness in Chinese Patients With Diabetes Using Machine Learning.” *FRONTIERS IN PHYSIOLOGY* 12. doi: 10.3389/fphys.2021.714190.
82. Li, Yung-Hui, Latifa Nabila Harfiya, Kartika Purwandari, and Yue-Der Lin. 2020. “Real-Time Cuffless Continuous Blood Pressure Estimation Using Deep Learning Model.” *SENSORS* 20(19). doi: 10.3390/s20190716.
83. Lilda, S. Daphin, R. Jayaparvathy, and Akshaya Balaji. 2022. “An Efficient Machine Learning Based Ventricular Late Potential Detection and Classification Technique for Cardiac Healthcare.” *CONCURRENCY AND COMPUTATION-PRACTICE & EXPERIENCE* 34(26). doi: 10.1002/cpe.7279.
84. Lo, Ying-Tsang, Hamido Fujita, and Tun-Wen Pai. 2016. “PREDICTION OF CORONARY ARTERY DISEASE BASED ON ENSEMBLE LEARNING APPROACHES AND CO-EXPRESSED OBSERVATIONS.” *JOURNAL OF MECHANICS IN MEDICINE AND BIOLOGY* 16(1, SI). doi: 10.1142/S0219019016501601.
85. Louridi, Nabaouia, Samira Douzi, and Bouabid El Ouahidi. 2021. “Machine Learning-Based Identification of Patients with a Cardiovascular Defect.” *JOURNAL OF BIG DATA* 8(1). doi: 10.1186/s13037-021-0024-9.
86. Lv, Haichen, Xiaolei Yang, Bingyi Wang, Shaobo Wang, Xiaoyan Du, Qian Tan, Zhujiang Hao, Ying Liu, Jun Yan, and Yunlong Xia. 2021. “Machine Learning-Driven Models to Predict Prognostic Outcomes in Patients Hospitalized With Heart Failure Using Electronic Health Records: Retrospective Study.” *JOURNAL OF MEDICAL INTERNET RESEARCH* 23(4). doi: 10.2196/24996.
87. Mahajan, Palak, Shahadat Uddin, Farshid Hajati, and Mohammad Ali Moni. 2023. “Ensemble Learning for Disease Prediction: A Review.” P. 1808 in *Healthcare*. Vol. 11.
88. Maheshwari, Danyal, Ubaid Ullah, Pablo A. Osorio Marulanda, Alain Garcia-Olea Jurado, Ignacio Diez Gonzalez, Jose M. Ormaetxe Merodio, and Begonya Garcia-Zapirain. 2023. “Quantum Machine Learning Applied to Electronic Healthcare Records for

- Ischemic Heart Disease Classification.” *HUMAN-CENTRIC COMPUTING AND INFORMATION SCIENCES* ۱۳. doi: ۱۰,۲۲۹۶۷/HCIS.۲۰۲۳,۱۳,۰۰۶.
۸۹. Maqsood, Sumbal, Shuxiang Xu, Matthew Springer, and Rami Mohawesh. ۲۰۲۱. “A Benchmark Study of Machine Learning for Analysis of Signal Feature Extraction Techniques for Blood Pressure Estimation Using Photoplethysmography (PPG).” *IEEE ACCESS* ۹:۱۳۸۸۱۷-۳۳. doi: ۱۰,۱۱۰۹/ACCESS.۲۰۲۱,۳۱۱۷۹۶۹.
۹۰. Miao, Kathleen H., Julia H. Miao, and George J. Miao. ۲۰۱۶. “Diagnosing Coronary Heart Disease Using Ensemble Machine Learning.” *International Journal of Advanced Computer Science and Applications* ۷(۱۰).
۹۱. Mishra, Aditya Dev, Ajay Kumar Shrivastava, and Megha Bhushan. ۲۰۲۰. “Disease Diagnosis and Control by Reinforcement Learning Techniques: A Systematic Literature Review.” *Discover Computing* ۲۸(۱):۱۲۹.
۹۲. Mondal, Subhash, Ranjan Maity, Yachang Omo, Soumadip Ghosh, and Amitava Nag. ۲۰۲۴. “An Efficient Computational Risk Prediction Model of Heart Diseases Based on Dual-Stage Stacked Machine Learning Approaches.” *IEEE ACCESS* ۱۲:۷۲۵۵-۷۰. doi: ۱۰,۱۱۰۹/ACCESS.۲۰۲۴,۳۳۵۰۹۹۶.
۹۳. Moradi, Hamed, Akram Al-Hourani, Gianmarco Concilia, Ebnaz Khoshmanesh, Farhad R. Nezami, Scott Needham, Sara Baratchi, and Khoshayar Khoshmanesh. ۲۰۲۳. “Recent Developments in Modeling, Imaging, and Monitoring of Cardiovascular Diseases Using Machine Learning.” *BIOPHYSICAL REVIEWS* ۱۵(۱, SI):۱۹-۳۳. doi: ۱۰,۱۰۰۷/s۱۲۵۵۱-۰۲۲-۰۱۰۴۰۰۷.
۹۴. Moreno-Sanchez, Pedro A. ۲۰۲۳. “Improvement of a Prediction Model for Heart Failure Survival through Explainable Artificial Intelligence.” *FRONTIERS IN CARDIOVASCULAR MEDICINE* ۱۰. doi: ۱۰,۳۳۸۹/fcvm.۲۰۲۳,۱۲۱۹۵۸۶.
۹۵. Moskaleva, Natalia E., Ksenia M. Shestakova, Alexey V Kukharensko, Pavel A. Markin, Maria V Kozhevnikova, Ekaterina O. Korobkova, Alex Brito, Sabina N. Baskhanova, Natalia V Mesonzhnik, Yuri N. Belenkov, Natalia V Pyatigorskaya, Elena Tobolkina, Serge Rudaz, and Svetlana A. Appolonova. ۲۰۲۲. “Target Metabolome Profiling-Based Machine Learning as a Diagnostic Approach for Cardiovascular Diseases in Adults.” *METABOLITES* ۱۲(۱۲). doi: ۱۰,۳۳۹۰/metabo۱۲۱۲۱۱۸۵.
۹۶. Muhammad, Ghulam, Saad Naveed, Lubna Nadeem, Tariq Mahmood, Amjad R. Khan, Yasar Amin, and Saeed Ali Omer Bahaj. ۲۰۲۳. “Enhancing Prognosis Accuracy for Ischemic Cardiovascular Disease Using K Nearest Neighbor Algorithm: A Robust Approach.” *IEEE ACCESS* ۱۱:۹۷۸۷۹-۹۵. doi: ۱۰,۱۱۰۹/ACCESS.۲۰۲۳,۳۳۱۲۰۴۶.
۹۷. Nabel, Elizabeth G. ۲۰۰۳. “Cardiovascular Disease.” *New England Journal of Medicine* ۳۴۹(۱):۶۰-۷۲.
۹۸. Najafi, Amirhossein, Alireza Nemati, Mahdi Ashrafzadeh, and Sarfaraz Hashemkhani Zolfani. ۲۰۲۳. “Multiple-Criteria Decision Making, Feature Selection, and Deep Learning: A Golden Triangle for Heart Disease Identification.” *ENGINEERING*

APPLICATIONS OF ARTIFICIAL INTELLIGENCE 120. doi:
10.1016/j.engappai.2023.106662.

99. Najim, Ahmed Haitham, and Nejah Nasri. 2023. "Artificial Intelligence for Heart Disease Prediction and Imputation of Missing Data in Cardiovascular Datasets." *COGENT ENGINEERING* 11(1). doi: 10.1080/23311916.2023.2320630.
100. Navaei, Reza Lotfi, Mohammad Safarzadeh, and Seyed Mohammad Jafar Sobhani. 2020. "Optimizing Flamelet Generated Manifold Models: A Machine Learning Performance Study." *ArXiv Preprint ArXiv*: 20.07.01030.
101. Nissa, Najmu, Sanjay Jamwal, Mehdi Neshat, Yudong Zhang, and Francesco Cauteruccio. 2023. "A Technical Comparative Heart Disease Prediction Framework Using Boosting Ensemble Techniques." *COMPUTATION* 12(1). doi: 10.3390/computation1201010.
102. Noor, Ayesha, Nadeem Javaid, Nabil Alrajeh, Babar Mansoor, Ali Khaqan, and Safdar Hussain Bouk. 2023. "Heart Disease Prediction Using Stacking Model With Balancing Techniques and Dimensionality Reduction." *IEEE ACCESS* 11:116026-40. doi: 10.1109/ACCESS.2023.2320681.
103. Nucifora, Gaetano, Daniele Muser, Joshua Bradley, Zoi Tsoumani, Giulia De Angelis, Thomas Caiffa, Matthias Schmitt, Gianfranco Sinagra, and Chris Miller. 2020. "Unsupervised Phenotypic Clustering of Cardiac MRI Data Reveals Distinct Subgroups Associated with Outcomes in Ischemic Cardiomyopathy." *The International Journal of Cardiovascular Imaging* 1-11.
104. Oanh, Tran Thi, and Nguyen Thanh Tung. 2022. "Predicting Hypertension Based on Machine Learning Methods: A Case Study in Northwest Vietnam." *MOBILE NETWORKS & APPLICATIONS* 27(2, SI):2013-23. doi: 10.1007/s11036-022-01984-w.
105. Ogunpola, Adedayo, Faisal Saeed, Shadi Basurra, Abdullah M. Albarrak, and Sultan Noman Qasem. 2023. "Machine Learning-Based Predictive Models for Detection of Cardiovascular Diseases." *DIAGNOSTICS* 13(2). doi: 10.3390/diagnostics13020244.
106. Ohashi, Yukitaka, Tomohiko Ihara, Kazutaka Oka, Yuya Takane, and Yukihiro Kikegawa. 2023. "Machine Learning Analysis and Risk Prediction of Weather-Sensitive Mortality Related to Cardiovascular Disease during Summer in Tokyo, Japan." *SCIENTIFIC REPORTS* 13(1). doi: 10.1038/s41598-023-44181-9.
107. Oliveira, Bruno Alberto Soares, Giulia Zanon Castro, Giovanna Luiza Medina Ferreira, and Frederico Gadelha Guimaraes. 2023. "CML-Cardio: A Cascade Machine Learning Model to Predict Cardiovascular Disease Risk as a Primary Prevention Strategy." *MEDICAL & BIOLOGICAL ENGINEERING & COMPUTING* 61(6):1409-20. doi: 10.1007/s11517-022-02707-z.
108. Oliveira, Mariana, Joana Seringa, Fausto Jose Pinto, Roberto Henriques, and Teresa

- Magalhaes. ۲۰۲۳. "Machine Learning Prediction of Mortality in Acute Myocardial Infarction." *BMC MEDICAL INFORMATICS AND DECISION MAKING* ۲۳(۱). doi: ۱۰.۱۱۸۶/s۱۲۹۱۱-۰۲۳-۰۲۱۶۸-۶.
۱۰۹. Ospina, Raydonal, Adenice G. O. Ferreira, Helio M. de Oliveira, Victor Leiva, and Cecilia Castro. ۲۰۲۳. "On the Use of Machine Learning Techniques and Non-Invasive Indicators for Classifying and Predicting Cardiac Disorders." *BIOMEDICINES* ۱۱(۱۰). doi: ۱۰.۳۳۹۰/biomedicines۱۱۱۰۲۶۰۴.
۱۱۰. Ott, Gabriel, Yannik Schaubelt, Juan Miguel Lopez Alcaraz, Wilhelm Haverkamp, and Nils Strodthoff. ۲۰۲۴. "Using Explainable AI to Investigate Electrocardiogram Changes during Healthy Aging-From Expert Features to Raw Signals." *PLOS ONE* ۱۹(۴). doi: ۱۰.۱۳۷۱/journal.pone.۰۳۰۲۰۲۴.
۱۱۱. Oyewola, David Opeoluwa, Emmanuel Gbenga Dada, and Sanjay Misra. ۲۰۲۴. "Diagnosis of Cardiovascular Diseases by Ensemble Optimization Deep Learning Techniques." *INTERNATIONAL JOURNAL OF HEALTHCARE INFORMATION SYSTEMS AND INFORMATICS* ۱۹(۱). doi: ۱۰.۴۰۱۸/IJHISI.۳۳۴۰۲۱.
۱۱۲. Pan, Chandan, Arnab Poddar, Rohan Mukherjee, and Ajoy Kumar Ray. ۲۰۲۲. "Impact of Categorical and Numerical Features in Ensemble Machine Learning Frameworks for Heart Disease Prediction." *BIOMEDICAL SIGNAL PROCESSING AND CONTROL* ۷۶. doi: ۱۰.۱۰۱۶/j.bspc.۲۰۲۲.۱۰۳۶۶۶.
۱۱۳. Pandey, Saroj Kumar, Anupam Shukla, Surbhi Bhatia, Thippa Reddy Gadekallu, Ankit Kumar, Arwa Mashat, Mohd Asif Shah, and Rekh Ram Janghel. ۲۰۲۳. "Detection of Arrhythmia Heartbeats from ECG Signal Using Wavelet Transform-Based CNN Model." *INTERNATIONAL JOURNAL OF COMPUTATIONAL INTELLIGENCE SYSTEMS* ۱۶(۱). doi: ۱۰.۱۰۰۷/s۴۴۱۹۶-۰۲۳-۰۰۲۵۶-z.
۱۱۴. Paul Navarrete, Jean, Jose Pinto, Rosa Liliana Figueroa, Maria Elena Lagos, Qing Zeng, and Carla Taramasco. ۲۰۲۲. "Supervised Learning Algorithm for Predicting Mortality Risk in Older Adults Using Cardiovascular Health Study Dataset." *APPLIED SCIENCES-BASEL* ۱۲(۲۲). doi: ۱۰.۳۳۹۰/app۱۲۲۲۱۱۵۳۶.
۱۱۵. Pavic, Ognjen, Lazar Dasic, Tijana Geroski, Mina Vaskovic Jovanovic, and Nenad Filipovic. ۲۰۲۳. "RISK CLASSIFICATION FOR SUDDEN CARDIAC DEATH IN PATIENTS WITH HYPERTROPHIC CARDIOMYOPATHY BASED ON MACHINE LEARNING ALGORITHMS." *JOURNAL OF THE SERBIAN SOCIETY FOR COMPUTATIONAL MECHANICS* ۱۷(۲):۱۵۳-۶۵. doi: ۱۰.۲۴۸۷۴/jsscm.۲۰۲۳.۱۷.۰۲.۱۱.
۱۱۶. Priyadarshi, Rahul, Rakesh Ranjan, Anish Kumar Vishwakarma, Tiansheng Yang, and Rajkumar Singh Rathore. ۲۰۲۴. "Exploring the Frontiers of Unsupervised Learning Techniques for Diagnosis of Cardiovascular Disorder: A Systematic Review." *IEEE Access*.
۱۱۷. Rahim, Aqsa, Yawar Rasheed, Farooque Azam, Muhammad Waseem Anwar, Muhammad Abdul Rahim, and Abdul Wahab Muzaffar. ۲۰۲۱. "An Integrated

- Machine Learning Framework for Effective Prediction of Cardiovascular Diseases.” *IEEE ACCESS* 9:1.6070-88. doi: 10.1109/ACCESS.2021.3098688.
118. Rahman, Md Saifur, Chandan Karmarkar, and Sheikh Mohammed Shariful Islam. 2022. “Application of Federated Learning in Cardiology: Key Challenges and Potential Solutions.” *Mayo Clinic Proceedings: Digital Health* 2(2):090-90.
 119. Rane, Nitin, Saurabh P. Choudhary, and Jayesh Rane. 2022. “Ensemble Deep Learning and Machine Learning: Applications, Opportunities, Challenges, and Future Directions.” *Studies in Medical and Health Sciences* 1(2):18-21.
 120. Reddy, Karna Vishnu Vardhana, Irraivan Elamvazuthi, Azrina Abd Aziz, Sivajothi Paramasivam, Hui Na Chua, and S. Pranavanand. 2021. “Heart Disease Risk Prediction Using Machine Learning Classifiers with Attribute Evaluators.” *APPLIED SCIENCES-BASEL* 11(18). doi: 10.3390/app11188302.
 121. Ribeiro, Pedro, Joana Sa, Daniela Paiva, and Pedro Miguel Rodrigues. 2022. “Cardiovascular Diseases Diagnosis Using an ECG Multi-Band Non-Linear Machine Learning Framework Analysis.” *BIOENGINEERING-BASEL* 11(1). doi: 10.3390/bioengineering11010058.
 122. Roth, Gregory A., George A. Mensah, Catherine O. Johnson, Giovanni Addolorato, Enrico Ammirati, Larry M. Baddour, Noël C. Barengo, Andrea Z. Beaton, Emelia J. Benjamin, Catherine P. Benziger, and others. 2020. “Global Burden of Cardiovascular Diseases and Risk Factors, 1990-2019: Update from the GBD 2019 Study.” *Journal of the American College of Cardiology* 76(25):2982-3021.
 123. Rustam, Furqan, Abid Ishaq, Kashif Munir, Mubarak Almutairi, Naila Aslam, and Imran Ashraf. 2022. “Incorporating CNN Features for Optimizing Performance of Ensemble Classifier for Cardiovascular Disease Prediction.” *DIAGNOSTICS* 12(6). doi: 10.3390/diagnostics12060942.
 124. Sajid, Mirza Rizwan, Arshad Ali Khan, Haitham M. Albar, Noryanti Muhammad, Waqas Sami, Syed Ahmad Chan Bukhari, and Iram Wajahat. 2022. “Exploration of Black Boxes of Supervised Machine Learning Models: A Demonstration on Development of Predictive Heart Risk Score.” *COMPUTATIONAL INTELLIGENCE AND NEUROSCIENCE* 2022. doi: 10.1100/2022/0470313.
 125. Sajid, Mirza Rizwan, Noryanti Muhammad, Roslinazairimah Zakaria, Ahmad Shahbaz, Syed Ahmad Chan Bukhari, Seifedine Kadry, and A. Suresh. 2021. “Nonclinical Features in Predictive Modeling of Cardiovascular Diseases: A Machine Learning Approach.” *INTERDISCIPLINARY SCIENCES-COMPUTATIONAL LIFE SCIENCES* 13(2, SI):201-11. doi: 10.1007/s12039-021-00423-w.
 126. Samaras, Agorastos-Dimitrios, Serafeim Moustakidis, Ioannis D. Apostolopoulos, Elpiniki Papageorgiou, and Nikolaos Papandrianos. 2023. “Uncovering the Black Box of Coronary Artery Disease Diagnosis: The Significance of Explainability in Predictive Models.” *APPLIED SCIENCES-BASEL* 13(12). doi: 10.3390/app13128120.
 127. Sandelowski, Margarete, Julie Barroso, and Corrine I. Voils. 2007. “Using Qualitative

- Metasummary to Synthesize Qualitative and Quantitative Descriptive Findings.” *Research in Nursing & Health* ۳۰(۱):۹۹-۱۱۱.
۱۲۸. Sapra, Varun, Luxmi Sapra, Akashdeep Bhardwaj, Salil Bharany, Akash Saxena, Faten Khalid Karim, Sara Ghorashi, and Ali Wagdy Mohamed. ۲۰۲۳. “Integrated Approach Using Deep Neural Network and CBR for Detecting Severity of Coronary Artery Disease.” *ALEXANDRIA ENGINEERING JOURNAL* ۶۸:۷۰۹-۲۰. doi: ۱۰.۱۰۱۶/j.aej.۲۰۲۳.۰۱.۰۲۹.
۱۲۹. Saputra, Jayson, Cindy Lawrencya, Jecky Mitra Saini, and Suharjito Suharjito. ۲۰۲۳. “Hyperparameter Optimization for Cardiovascular Disease Data-Driven Prognostic System.” *VISUAL COMPUTING FOR INDUSTRY BIOMEDICINE AND ART* ۶(۱). doi: ۱۰.۱۱۸۶/s۴۲۴۹۲-۰۲۳-۰۰۱۴۳-۶.
۱۳۰. Sharma, Prerna, Krishna Choudhary, Kshitij Gupta, Rahul Chawla, Deepak Gupta, and Arun Sharma. ۲۰۲۰. “Artificial Plant Optimization Algorithm to Detect Heart Rate & Presence of Heart Disease Using Machine Learning.” *ARTIFICIAL INTELLIGENCE IN MEDICINE* ۱۰۲. doi: ۱۰.۱۰۱۶/j.artmed.۲۰۱۹.۱۰۱۷۵۲.
۱۳۱. Shchetinin, E. Y., and A. G. Glushkova. ۲۰۲۲. “Arrhythmia Detection Using Resampling and Deep Learning Methods on Unbalanced Data.” *COMPUTER OPTICS* ۴۶(۶):۹۸۰-۸۷. doi: ۱۰.۱۸۲۸۷/۲۴۱۲-۶۱۷۹-CO-۱۱۱۲.
۱۳۲. Sheeba, Adlin, S. Padmakala, C. A. Subasini, and S. P. Karuppiah. ۲۰۲۲. “MKELM: Mixed Kernel Extreme Learning Machine Using BMDA Optimization for Web Services Based Heart Disease Prediction in Smart Healthcare.” *COMPUTER METHODS IN BIOMECHANICS AND BIOMEDICAL ENGINEERING* ۲۵(۱۰):۱۱۸۰-۹۴. doi: ۱۰.۱۰۸۰/۱۰۲۵۵۸۴۲.۲۰۲۲.۲۰۳۴۷۹۵.
۱۳۳. Sherazi, Syed Waseem Abbas, Huilin Zheng, and Jong Yun Lee. ۲۰۲۳. “A Machine Learning-Based Applied Prediction Model for Identification of Acute Coronary Syndrome (ACS) Outcomes and Mortality in Patients during the Hospital Stay.” *SENSORS* ۲۳(۳). doi: ۱۰.۳۳۹۰/s۲۳۰۳۱۳۵۱.
۱۳۴. Shi, Jingjing, Chao Chen, Hui Liu, Yinglong Wang, Minglei Shu, and Qing Zhu. ۲۰۲۱. “Automated Atrial Fibrillation Detection Based on Feature Fusion Using Discriminant Canonical Correlation Analysis.” *COMPUTATIONAL AND MATHEMATICAL METHODS IN MEDICINE* ۲۰۲۱. doi: ۱۰.۱۱۵۵/۲۰۲۱/۶۶۹۱۱۷۷.
۱۳۵. Sidey-Gibbons, Jenni A. M., and Chris J. Sidey-Gibbons. ۲۰۱۹. “Machine Learning in Medicine: A Practical Introduction.” *BMC Medical Research Methodology* ۱۹:۱-۱۸.
۱۳۶. Simegn, Gizeaddis Lamesgin, Worku Birhanie Gebeyehu, and Mizanu Zelalem Degu. ۲۰۲۲. “Computer-Aided Decision Support System for Diagnosis of Heart Diseases.” *RESEARCH REPORTS IN CLINICAL CARDIOLOGY* ۱۳:۳۹-۵۴. doi: ۱۰.۲۱۴۷/RRCC.S۳۶۶۳۸۰.
۱۳۷. Sinha, Nidhi, M. A. Ganesh Kumar, Amit M. Joshi, and Linga Reddy Cenkeramaddi. ۲۰۲۳. “DASMcC: Data Augmented SMOTE Multi-Class Classifier for Prediction of Cardiovascular Diseases Using Time Series Features.” *IEEE ACCESS* ۱۱:۱۱۷۶۴۳-

٥٥. doi: ١٠,١١٠٩/ACCESS.٢٠٢٣,٣٣٢٥٧.٥.

١٣٨. Sozen, Mert Erkan, Gorkem Sariyer, Mustafa Yigit Sozen, Gaurav Kumar Badhotiya, and Lokesh Vijavargy. ٢٠٢٣. "Machine Learning Implementations for Multi-Class Cardiovascular Risk Prediction in Family Health Units." *INTERNATIONAL JOURNAL OF MATHEMATICAL ENGINEERING AND MANAGEMENT SCIENCES* ٨(٦):١١٧١-٨٧. doi: ١٠,٣٣٨٨٩/IJMEMS.٢٠٢٣,٨,٦,٠٦٦.
١٣٩. Stepanyan, I. V, Ch. A. Alimbayev, M. O. Savkin, D. Lyu, and M. Zidun. ٢٠٢٢. "Comparative Analysis of Machine Learning Methods for Prediction of Heart Diseases." *JOURNAL OF MACHINERY MANUFACTURE AND RELIABILITY* ٥١(٨):٧٨٩-٩٩. doi: ١٠,٣١٠٣/S١٠٥٢٦١٨٨٢٢٠٨٠٢١٠.
١٤٠. Subramani, Sivakannan, Neeraj Varshney, M. Vijay Anand, Manzoore Elahi M. Soudagar, Lamya Ahmed Al-keridis, Tarun Kumar Upadhyay, Nawaf Alshammari, Mohd Saeed, Kumaran Subramanian, Krishnan Anbarasu, and Karunakaran Rohini. ٢٠٢٣. "Cardiovascular Diseases Prediction by Machine Learning Incorporation with Deep Learning." *FRONTIERS IN MEDICINE* ١٠. doi: ١٠,٣٣٨٩/fmed.٢٠٢٣,١١٥٠٩٣٣.
١٤١. Sun, Jin-Yu, Hui Shen, Qiang Qu, Wei Sun, and Xiang-Qing Kong. ٢٠٢١. "The Application of Deep Learning in Electrocardiogram: Where We Came from and Where We Should Go?" *International Journal of Cardiology* ٣٣٧:٧١-٧٨.
١٤٢. Tachie, Christabel Y. E., Daniel Obiri-Ananay, Nii Adjetey Tawiah, Nii Attoh-Okine, and Alberta N. A. Aryee. ٢٠٢٣. "Machine Learning Approaches for Predicting Fatty Acid Classes in Popular US Snacks Using NHANES Data." *NUTRIENTS* ١٥(١٥). doi: ١٠,٣٣٩٠/nu١٥١٥٣٣١.
١٤٣. Tadege, Melaku, Awoke Seyoum Tegegne, and Zelalem G. Dessie. ٢٠٢٤. "Post-Surgery Survival and Associated Factors for Cardiac Patients in Ethiopia: Applications of Machine Learning, Semi-Parametric and Parametric Modelling." *BMC MEDICAL INFORMATICS AND DECISION MAKING* ٢٤(١). doi: ١٠,١١٨٦/s١٢٩١١-٠٢٤-٠٢٤٨٠-٩.
١٤٤. Tadesse, Girmaw Abebe, Hamza Javed, Komminist Weldemariam, Yong Liu, Jin Liu, Jiyan Chen, and Tingting Zhu. ٢٠٢١. "DeepMI: Deep Multi-Lead ECG Fusion for Identifying Myocardial Infarction and Its Occurrence-Time." *ARTIFICIAL INTELLIGENCE IN MEDICINE* ١٢١. doi: ١٠,١٠١٦/j.artmed.٢٠٢١,١٠٢١٩٢.
١٤٥. Teresa Garcia-Ordas, Maria, Martin Bayon-Gutierrez, Carmen Benavides, Jose Aveleira-Mata, and Jose Alberto Benitez-Andrades. ٢٠٢٣. "Heart Disease Risk Prediction Using Deep Learning Techniques with Feature Augmentation." *MULTIMEDIA TOOLS AND APPLICATIONS* ٨٢(٢٠):٣١٧٥٩-٧٣. doi: ١٠,١٠٠٧/s١١٠٤٢-٠٢٣-١٤٨١٧-z.
١٤٦. Thakur, Gajendra Singh, Sunil Kumar Sahu, N. Kumar Swamy, Manish Gupta, Tony Jan, Mukesh Prasad, and Andrea Prati. ٢٠٢٣. "Review of Soft Computing Techniques in Monitoring Cardiovascular Disease in the Context of South Asian Countries." *APPLIED SCIENCES-BASEL* ١٣(١٧). doi: ١٠,٣٣٩٠/app١٣١٧٩٥٥٥.

۱۴۷. Tiwari, Achyut, Aryan Chugh, and Aman Sharma. ۲۰۲۲. "Ensemble Framework for Cardiovascular Disease Prediction." *COMPUTERS IN BIOLOGY AND MEDICINE* ۱۴۶. doi: ۱۰.۱۰۱۶/j.compbiomed.۲۰۲۲.۱۰۵۶۲۴.
۱۴۸. Tolentino, Lean Karlo S., John Erick L. Isoy, Kayne Adriane A. Bulawan, Mary Claire T. Co, Caryl Faye C. Monreal, Ian Joshua W. Vitto, Maria Victoria C. Padilla, Jay Fel C. Quijano, Romeo Jr L. Jorda, and Jessica S. Velasco. ۲۰۲۰. "CAREdio: Health Screening and Heart Disease Prediction System for Rural Communities in the Philippines." *INTERNATIONAL JOURNAL OF ADVANCED COMPUTER SCIENCE AND APPLICATIONS* ۱۱(۸): ۴۶۴-۷۲.
۱۴۹. Torres, Romina, Christopher Zurita, Diego Mellado, Orietta Nicolis, Carolina Saavedra, Marcelo Tuesta, Matias Salinas, Ayleen Bertini, Oneglio Pedemonte, Marvin Querales, and Rodrigo Salas. ۲۰۲۳. "Predicting Cardiovascular Rehabilitation of Patients with Coronary Artery Disease Using Transfer Feature Learning." *DIAGNOSTICS* ۱۳(۳). doi: ۱۰.۳۳۹۰/diagnostics۱۳۰۳۰۵۰۸.
۱۵۰. Tsai, Dung-Jang, Shih-Hung Tsai, Hui-Hsun Chiang, Chia-Cheng Lee, and Sy-Jou Chen. ۲۰۲۲. "Development and Validation of an Artificial Intelligence Electrocardiogram Recommendation System in the Emergency Department." *JOURNAL OF PERSONALIZED MEDICINE* ۱۲(۵). doi: ۱۰.۳۳۹۰/jpm۱۲۰۵۰۷۰۰.
۱۵۱. Varale, Viraj S., and Kalpana S. Thakre. ۲۰۲۰. "Prediction of Heart Disease Using Machine Learning Algorithm." *BIOSCIENCE BIOTECHNOLOGY RESEARCH COMMUNICATIONS* ۱۳(۱۴, SI): ۲۸۷-۹۰. doi: ۱۰.۲۱۷۸۶/bbrc/۱۳,۱۴/۶۷.
۱۵۲. Virani, Salim S., Alvaro Alonso, Emelia J. Benjamin, Marcio S. Bittencourt, Clifton W. Callaway, April P. Carson, Alanna M. Chamberlain, Alexander R. Chang, Susan Cheng, Francesca N. Delling, and others. ۲۰۲۰. "Heart Disease and Stroke Statistics—۲۰۲۰ Update: A Report from the American Heart Association." *Circulation* ۱۴۱(۹): e۱۳۹-۵۹۶.
۱۵۳. Wu, Mengze, Yongdi Lu, Wenli Yang, and Shen Yuong Wong. ۲۰۲۱. "A Study on Arrhythmia via ECG Signal Classification Using the Convolutional Neural Network." *FRONTIERS IN COMPUTATIONAL NEUROSCIENCE* ۱۴. doi: ۱۰.۳۳۸۹/fncom.۲۰۲۰.۵۶۴۰۱۵.
۱۵۴. Wu, Xingwei, Chang Huan, Mengting Li, Qin Lv, Jiaying Zhang, Enwu Long, Jiuqun Zhu, and Rongsheng Tong. ۲۰۲۲. "A Machine Learning-Based Risk Warning Platform for Potentially Inappropriate Prescriptions for Elderly Patients with Cardiovascular Disease." *FRONTIERS IN PHARMACOLOGY* ۱۳. doi: ۱۰.۳۳۸۹/fphar.۲۰۲۲.۸۰۴۵۶۶.
۱۵۵. Xiang, Menghui, Junbin Zang, Juliang Wang, Haoxin Wang, Chenzheng Zhou, Ruiyu Bi, Zhidong Zhang, and Chenyang Xue. ۲۰۲۳. "Research of Heart Sound Classification Using Two-Dimensional Features." *BIOMEDICAL SIGNAL PROCESSING AND CONTROL* ۷۹(۲). doi: ۱۰.۱۰۱۶/j.bspc.۲۰۲۲.۱۰۴۱۹۰.
۱۵۶. Yadav, Anjali, Anushikha Singh, Malay Kishore Dutta, and Carlos M. Travieso. ۲۰۲۰.

- “Machine Learning-Based Classification of Cardiac Diseases from PCG Recorded Heart Sounds.” *NEURAL COMPUTING & APPLICATIONS* 32(24, SI):17843–56. doi: 10.1007/s00521-019-04047-0.
156. Yang, Weiyi, Yujuan Si, Di Wang, and Buhao Guo. 2018. “Automatic Recognition of Arrhythmia Based on Principal Component Analysis Network and Linear Support Vector Machine.” *COMPUTERS IN BIOLOGY AND MEDICINE* 101:22–32. doi: 10.1016/j.combiomed.2018.08.003.
157. Yazdi, Fatemeh, and Shahrokh Asadi. 2020. “Enhancing Cardiovascular Disease Diagnosis: The Power of Optimized Ensemble Learning.” *IEEE Access* 13:46747–62.
158. Yongcharoenchaiyasit, Khomkrit, Sujitra Arwatchananukul, Punnarumol Temdee, and Ramjee Prasad. 2023. “Gradient Boosting Based Model for Elderly Heart Failure, Aortic Stenosis, and Dementia Classification.” *IEEE ACCESS* 11:48677–96. doi: 10.1109/ACCESS.2023.3276468.
159. You, Jia, Yu Guo, Ju-Jiao Kang, Hui-Fu Wang, Ming Yang, Jian-Feng Feng, Jin-Tai Yu, and Wei Cheng. 2023. “Development of Machine Learning-Based Models to Predict 10-Year Risk of Cardiovascular Disease: A Prospective Cohort Study.” *STROKE AND VASCULAR NEUROLOGY* 8(7):470–80. doi: 10.1136/svn-2023-002332.
160. Zafar, Kainat, Hafeez Ur Rehman Siddiqui, Abdul Majid, Adil Ali Saleem, Ali Raza, Furqan Rustam, and Sandra Dudley. 2023. “Deep Learning-Based Feature Engineering to Detect Anterior and Inferior Myocardial Infarction Using UWB Radar Data.” *IEEE ACCESS* 11:97740–57. doi: 10.1109/ACCESS.2023.3312948.
161. Zhanquan, Sun, Wang Chaoli, Tian Engang, and Yin Zhong. 2022. “ECG Signal Classification via Combining Hand-Engineered Features with Deep Neural Network Features.” *MULTIMEDIA TOOLS AND APPLICATIONS* 81(10):13467–88. doi: 10.1007/s11042-021-11023-6.
162. Zou, Congyu, Alexander Mueller, Utschick Wolfgang, Daniel Rueckert, Phillip Mueller, Matthias Becker, Alexander Steger, and Eimo Martens. 2022. “Heartbeat Classification by Random Forest With a Novel Context Feature: A Segment Label.” *IEEE JOURNAL OF TRANSLATIONAL ENGINEERING IN HEALTH AND MEDICINE* 10. doi: 10.1109/JTEHM.2022.3202749.