

بهبود دقت تشخیص در انتخاب کاندید درمان همزمان‌سازی مجدد قلبی با استفاده از بیش‌نمونه‌برداری اقلیت تعمیم‌یافته و یادگیری گروهی

آرش ساغری^۱، فارغ‌التحصیل کارشناسی ارشد، محمد نژاده^{۲*}، استادیار

^۱گروه مهندسی کامپیوتر، واحد فومن و شفت، دانشگاه آزاد اسلامی، فومن، ایران Saghari@guilan.ac.ir

^۲گروه مهندسی کامپیوتر، واحد فومن و شفت، دانشگاه آزاد اسلامی، فومن، ایران mohammad.nejadedh@iau.ac.ir

چکیده: با گسترش کاربردهای هوش مصنوعی به‌ویژه یادگیری ماشین در حوزه پزشکی، امکان ارتقای دقت و اثربخشی در تشخیص و درمان بیماری‌ها فراهم شده است. یکی از کاربردهای مهم آن، بهبود تشخیص پاسخ بیماران به درمان همزمان‌سازی مجدد قلبی است؛ روشی درمانی برای بیماران مبتلا به نارسایی قلبی که در آن با تحریک هماهنگ بطن‌ها عملکرد قلب بهبود می‌یابد. یکی از چالش‌های مهم در این زمینه، عدم تعادل کلاس‌ها میان بیماران پاسخ‌دهنده و غیرپاسخ‌دهنده به درمان است که دقت مدل‌های طبقه‌بندی را کاهش می‌دهد. در این پژوهش، روشی ترکیبی مبتنی بر بیش‌نمونه‌برداری مصنوعی تعمیم‌یافته از کلاس اقلیت و یادگیری گروهی پیشنهاد شده است. برای ارتقای کیفیت نمونه‌های مصنوعی، از الگوریتم جستجوی کلاخ به‌عنوان یک روش فراابتکاری استفاده شده و برای کاهش بُعد و انتخاب ویژگی‌های مؤثر از الگوریتم ژنتیک بهره گرفته شده است. در مرحله طبقه‌بندی، دو مدل یادگیری گروهی شامل گرادیان تقویت‌شده و جنگل تصادفی به کار گرفته شدند. مجموعه داده مورد استفاده شامل ۶۰ ویژگی اولیه بود که با استفاده از الگوریتم ژنتیک به ۴۱ ویژگی منتخب بهینه‌شده کاهش یافت. معیار بهینگی، حداکثرسازی دقت مدل در شناسایی بیماران غیرپاسخ‌دهنده بود. نتایج ارزیابی مدل‌ها نشان داد روش پیشنهادی با این ۴۱ ویژگی منتخب، میانگین هارمونیک برابر با ۸۹٫۰۷٪ و دقت طبقه غیرپاسخ‌دهنده به درمان برابر با ۹۳٫۵۹٪ را به دست آورده است. یافته‌های این پژوهش نشان می‌دهد که ترکیب روش‌های بهینه‌سازی، بیش‌نمونه‌برداری و یادگیری گروهی می‌تواند به‌طور مؤثری دقت تشخیص بیماران غیرپاسخ‌دهنده به درمان همزمان‌سازی مجدد قلبی را افزایش داده و به تصمیم‌گیری‌های درمانی مبتنی بر داده کمک کند.

واژه‌های کلیدی: تشخیص همزمان‌سازی مجدد قلبی، الگوریتم جستجوی کلاخ، یادگیری گروهی، بیش‌نمونه‌برداری اقلیت، نارسایی قلبی

Enhancing Accuracy in Cardiac Resynchronization Therapy Candidate Selection using Generalized Minority Oversampling and Ensemble Learning

Arash Saghari ¹, Master's degree graduate, A Mohammad Nejadeh ^{2*}, Assistant professor

¹ Department of Computer Engineering, FSh.C., Islamic Azad University, Fouman, Iran, Saghari@guilan.ac.ir

² Department of Computer Engineering, FSh.C., Islamic Azad University, Fouman, Iran, mohammad.nejadeh@iau.ac.ir

Abstract: The spreading of artificial intelligence, especially machine learning applications in the medical field, has opened that possibility of raising the precision and effectiveness of diagnosis and treatment of diseases. One of the important applications is the facilitation of the detection of the patient's response to the treatment by cardiac resynchronization therapy, a treatment method for patients with heart failure in which the function of the heart is improved by the coordinated stimulation of the ventricles. One of the main challenges in this area is class imbalance among the numbers of responders and non-responders to the treatment, which lowers the accuracy of the classification models. This research proposes a hybrid method based on Synthetic Minority Over-sampling with generalized class and ensemble learning. To improve the quality of synthetic samples, the Crow Search Algorithm has been used as a metaheuristic method and the Genetic Algorithm has been used for dimensionality reduction and efficient feature selection. For the classification stage, two ensemble learning models of Gradient Boosting and Random Forest were implemented. The utilized dataset includes 60 initial features, which have been reduced to 41 optimized selected features using the genetic algorithm. The criterion of optimality was the maximization of model accuracy in the identification of non-responders. The performance of the proposed method with these 41 selected features showed the average harmonic with the value of 89.07% and the accuracy of the non-response class of 93.59%. The results of this study indicated the combination of optimization, oversampling, and ensemble learning methods could effectively increase the identification accuracy of the non-responder patients of cardiac resynchronization therapy and thus provide assistive data-driven medical decision-making.

Keywords: Cardiac resynchronization therapy (CRT) diagnosis, Crow search algorithm, Ensemble learning, Minority oversampling, Heart failure.

۱. مقدمه

همسایه‌ها و تعداد نمونه‌های مصنوعی بهره می‌برد. برخلاف روش‌های کلاسیک که مقادیر ثابتی برای این پارامترها در نظر می‌گیرند، در اینجا با جستجوی هوشمند، بهترین ترکیب برای داده‌های پزشکی خاص درمان CRT شناسایی می‌شود. افزون بر این، استفاده از یادگیری گروهی^۵ برای طبقه‌بندی نهایی داده‌ها، با هدف دستیابی به تعادلی بین دقت^۶ تشخیص در هر دو کلاس و کاهش بایاس مدل طراحی شده است. همچنین استفاده از میانگین هارمونیک (F1)^۷ حساسیت و ویژگی به‌عنوان معیار ارزیابی، تضمین می‌کند که مدل در برابر هر دو نوع خطا (مثبت کاذب و منفی کاذب) مقاوم باشد [۵]. در حوزه پزشکی جاری که حجم داده‌های آموزش بالا نیست و طبقه‌ها بسیار نامتعادل است رویکرد طبقه‌بندی گروهی می‌تواند یک رویکرد موثر باشد. از دیدگاه به روزترین فناوری، این تحقیق بر مجموعه‌ای از دستاوردهای روش‌شناسی پیشین بنا شده، اما در عین حال تفاوت‌های کلیدی با آن‌ها دارد:

- اغلب پژوهش‌ها تنها از نسخه پایه SMOTE یا تغییرات ساده‌ی آن بهره برده‌اند، بدون تنظیم تطبیقی پارامترها برای داده‌های پزشکی خاص.

- در بسیاری از تحقیقات، الگوریتم‌های یادگیری گروهی مورد استفاده قرار گرفته‌اند، اما به ندرت با الگوریتم‌های بهینه‌سازی برای بهبود تعادل نمونه‌ها ترکیب شده‌اند.

- کارهای پیشین کمتر از معیار میانگین هارمونیک برای سنجش تعادل عملکرد دو طبقه استفاده کرده‌اند، درحالی‌که این معیار برای تحلیل دقیق سیستم‌های حساس مانند درمان‌های پزشکی بسیار مؤثر است.

- ترکیب الگوریتم CSA با SMOTE برای استخراج نمونه‌های اقلیت، یک نوآوری محسوب می‌شود که پیش‌تر در مطالعات مرتبط مورد بررسی قرار نگرفته است.

در نهایت، ساختار این مقاله به شکل زیر سازمان‌دهی شده است: در بخش دوم، پیشینه پژوهش و مرور ادبیات شامل رویکردهای رایج در طبقه‌بندی داده‌های نامتعادل و کاربردهای پزشکی درمان CRT ارائه می‌شود. بخش سوم، روش‌شناسی تحقیق و

طبقه‌بندی دقیق در حوزه سلامت، به‌ویژه در تحلیل پاسخ به درمان، همواره یکی از چالش‌های مهم یادگیری ماشین (ML)^۱ بوده است. این چالش به‌خصوص در مجموعه داده‌هایی که از نظر توزیع طبقات نامتوازن هستند، برجسته‌تر می‌شود. در این شرایط، کلاس اقلیت (مثلاً بیماران عدم پاسخ‌دهنده به درمان خاص) سهم بسیار کمی از داده‌ها را به خود اختصاص می‌دهد و همین امر موجب می‌شود که مدل‌های سنتی ML به دلیل تمرکز بیش از حد بر کلاس اکثریت، عملکرد ضعیفی در تشخیص نمونه‌های اقلیت داشته باشند [۱]. در کاربردهای حیاتی مانند تشخیص کلاهدرداری، پیش‌بینی نقص‌های نرم‌افزاری، تشخیص بیماری و به‌ویژه پیش‌بینی اثربخشی درمان‌های پزشکی مانند درمان همزمان‌سازی مجدد قلبی (CRT)^۲، این عدم تعادل می‌تواند منجر به بروز خطاهایی با پیامدهای جدی شود [۲].

به‌طور خاص، در درمان CRT که یکی از شیوه‌های پیشرفته در بهبود عملکرد قلبی بیماران نارسای قلبی است، بررسی‌ها نشان داده‌اند که حدود ۳۰ تا ۳۵ درصد بیماران پاسخ مناسبی به این درمان نمی‌دهند [۳]. لذا طراحی سیستم‌هایی هوشمند برای پیش‌بینی دقیق این پاسخ، یک ضرورت حیاتی برای تصمیم‌گیری پزشکی است. از سوی دیگر، چالش‌هایی مانند کم‌بودن داده‌های مربوط به کلاس اقلیت، هم‌پوشانی ویژگی‌ها میان طبقات، و عدم وجود تعادل ساختاری در داده‌ها، ضرورت بهره‌گیری از راهکارهای نوین ترکیبی را دوچندان می‌سازد [۴].

در مواجهه با این چالش‌ها، پژوهش حاضر مدلی ترکیبی ارائه می‌دهد که بر بهبود کیفیت داده‌های طبقه‌ی اقلیت از طریق بیش‌نمونه‌برداری هدفمند تمرکز دارد. نوآوری اصلی پژوهش در طراحی و پیاده‌سازی یک نسخه تعمیم‌یافته از الگوریتم بیش‌نمونه‌برداری مصنوعی از طبقه‌ی اقلیت (SMOTE)^۳ نهفته است که از یک الگوریتم فراابتکاری به نام الگوریتم جستجوی کلاغ (CSA)^۴ برای تنظیم پارامترهای حیاتی SMOTE مانند تعداد

⁵ Ensemble Learning

⁶ Accuracy

⁷ F1-Measure

¹ Machine Learning (ML)

² Cardiac Resynchronization Therapy (CRT)

³ Synthetic Minority Over-sampling Technique (SMOTE)

⁴ Crow Search Algorithm (CSA)

چارچوب پیشنهادی شامل تنظیمات الگوریتمی، ساختار مدل، و فرآیند اجرای آزمایشات را تشریح می‌کند. بخش چهارم به تحلیل نتایج و مقایسه عملکرد مدل پیشنهادی با سایر روش‌ها اختصاص دارد. سرانجام، بخش پنجم به نتیجه‌گیری و پیشنهادهایی برای تحقیقات آتی می‌پردازد.

۲. ادبیات تحقیق

مطالعات متعددی به چالش طبقه‌بندی داده‌های نامتوازن پرداخته‌اند. بخش قابل توجهی از این پژوهش‌ها، به‌ویژه در حوزه داده‌های پزشکی و قلبی‌عروقی، تمرکز ویژه‌ای بر این موضوع داشته‌اند. در این راستا، رویکردهای گوناگونی برای تولید داده‌های مصنوعی معرفی شده‌اند که شامل الگوریتم‌های کلاسیک، روش‌های ترکیبی و تکنیک‌های یادگیری عمیق (DL)^۸ می‌شوند. این ابزارها با هدف تسهیل و ارزیابی فرآیند متعادل‌سازی داده‌ها توسعه یافته‌اند. در ادامه، برخی از جدیدترین و مرتبط‌ترین تحقیقات در این زمینه مرور می‌شوند. روش‌های نمونه‌برداری، از جمله تکنیک SMOTE (نمونه‌برداری بیش‌ازحد اقلیت مصنوعی)، معمولاً برای متعادل‌سازی مجموعه داده‌های مصنوعی بر روی داده‌های بدون برچسب به‌کار گرفته می‌شوند. متیو و همکارانش [۶]، نسخه‌ای بهبودیافته از SMOTE با عنوان KWLSMOTE (نمونه‌برداری بیش‌ازحد اقلیت مصنوعی وزن‌دار هسته با دقت محلی) را پیشنهاد کردند که با تمرکز بر فضای ویژگی ماشین بردار پشتیبان (SVM)، محدودیت‌های نسخه اصلی SMOTE را در مواجهه با مرزهای تصمیم‌گیری پیچیده برطرف می‌سازد. مزیت اصلی این روش افزایش دقت طبقه‌بندی در محیط‌های غیرخطی است، هرچند پیچیدگی محاسباتی بالا و نیاز به تنظیم دقیق پارامترهای هسته از محدودیت‌های آن به‌شمار می‌رود. بنین و همکارانش [۷]، الگوریتم ماهاکیل^۹ را به‌عنوان روشی نوین برای نمونه‌برداری بیش‌ازحد مثبتی بر کروموزوم‌های وراثتی در داده‌های نرم‌افزاری نامتوازن معرفی کردند. این الگوریتم با تولید نمونه‌های مصنوعی بر اساس شباهت ژنتیکی، می‌تواند تنوع کلاس اقلیت و در نتیجه عملکرد طبقه‌بندی‌کننده را بهبود بخشد. با وجود افزایش دقت

طبقه‌بندی، این روش به منابع محاسباتی قابل توجهی نیاز دارد و برای کاربردهای غیرنرم‌افزاری بهینه‌سازی نشده است. چنگ و همکاران [۸]، الگوریتمی اصلاح‌شده از ماشین بردار پشتیبان (SVM)^{۱۰} را برای طبقه‌بندی داده‌های سرطان سینه ارائه کردند که ترکیبی از انتخاب هوشمند نمونه و SVM است. این روش با حذف نمونه‌های نویزی و مرزی، دقت طبقه‌بندی داده‌های نامتوازن را ارتقا می‌دهد. نقطه قوت این رویکرد در ادغام یادگیری نظارت‌شده با نمونه‌برداری پنهان نهفته است، هرچند قابلیت تعمیم آن به سایر حوزه‌ها محدود است. حسین و همکاران [۹]، روش SMOTE تطبیقی را معرفی کردند که با استفاده از معیار فاصله میان نمونه‌های کلاس اقلیت، داده‌های مصنوعی واقع‌بینانه‌تری تولید می‌کند. این رویکرد با هدف کاهش نویز و افزایش تعمیم‌پذیری^{۱۱} طراحی شده و نسبت به نسخه‌های کلاسیک SMOTE عملکرد بهتری نشان داده است. با این حال، پیچیدگی محاسباتی و دشواری در تنظیم معیار فاصله از چالش‌های اصلی آن محسوب می‌شود. لاوتی و همکاران [۱۰]، از رویکرد «نمونه‌برداری کمتر مثبتی بر آنتروپی همراه با گروه پویا» برای تشخیص تقلب بهره گرفتند. این روش با انتخاب نمونه‌های نماینده‌تر از کلاس اکثریت، تعادل بهتری در داده‌ها ایجاد کرده و دقت سیستم‌های تشخیص را افزایش داده است. با وجود نتایج امیدوارکننده، مقیاس‌پذیری این مدل در مجموعه داده‌های بزرگ همچنان محدودیت دارد. شارما و همکاران [۱۱]، به بررسی محدودیت‌های شبکه‌های مولد متخاصم (GANs)^{۱۲} در شرایط کمبود داده در کلاس اقلیت پرداختند. یافته‌های آنان نشان داد که در صورت ناکافی بودن داده‌های آموزشی، GANها به‌جای یادگیری توزیع واقعی داده‌ها، صرفاً نمونه‌های موجود را بازتولید می‌کنند که این امر موجب افت عملکرد مدل می‌شود. این مطالعه بر ضرورت فراهم‌سازی حجم کافی داده برای آموزش GANها تأکید دارد، هرچند راهکار جایگزینی ارائه نشده است. سیسودیا و همکاران [۱۲]، الگوریتم نمونه‌گیری احتمالی شبه‌قطعی K-نزدیک‌ترین

¹⁰ Support Vector Machine (SVM)

¹¹ Generalizability

¹² Generative Adversarial Networks (GANs)

⁸ Deep Learning (DL)

⁹ MAHAKIL

همسایه‌ها (QDPSKNN)^{۱۳} را برای مدیریت توزیع نامتوازن داده‌های کلیک کاربر بهبود دادند. این روش با در نظر گرفتن توزیع واقعی و احتمالی داده‌ها، نتایج پایدارتر و قابل‌اعتمادتری نسبت به روش‌های مرسوم ارائه می‌دهد. مزیت اصلی آن حفظ پایداری عملکرد در داده‌های پرتکرار است، هرچند وابستگی به انتخاب صحیح پارامترها از محدودیت‌های کلیدی آن محسوب می‌شود. مطالعه‌ی جانباشا و همکاران [۱۳]، به بررسی استفاده از شبکه‌های GAN و DL در متعادل‌سازی داده‌های فیزیولوژیکی نظیر ECG پرداخت. آن‌ها با معرفی Cardio GAN توانستند دقتی بالاتر از ۹۹٫۰٪ در طبقه‌بندی سیگنال‌های قلبی کسب کنند. مزیت بارز این روش در ادغام DL با تولید داده‌های مصنوعی است، با این حال، پیچیدگی مدل‌ها و نیاز به آموزش طولانی مدت از نقاط ضعف آن است. ویدمن و همکاران [۱۴]، تأثیر کیفیت داده را بر عملکرد طبقه‌بندی در داده‌های نامتعادل بررسی کردند. این پژوهش نشان داد که حتی خطاهای اندک در داده می‌توانند باعث کاهش چشمگیر دقت شوند، در حالی که پاک‌سازی هدفمند و انتخاب ویژگی مناسب می‌تواند نتایج را بهبود بخشد. سهم نوآورانه‌ی این تحقیق در تمرکز بر جنبه کیفیت داده بود، گرچه پیشنهادات عملی برای بهبود داده محدود بود. دوان و همکاران [۱۵]، در پاسخ به مشکل عدم تعادل در یادگیری چندبرجسبی، روشی مبتنی بر خوشه‌بندی سلسله‌مراتبی^{۱۴} ارائه کردند که با تقسیم فضای برجسب به زیر فضاهای همبسته، تبدیل به مسائل چندطبقه می‌شود. نتایج تجربی نشان داد که این روش نسبت به مدل‌های پیشین عملکرد بهتری در طبقه‌بندی دارد. مزیت مهم آن در بهره‌گیری از همبستگی بین برجسب‌هاست، اگرچه پیچیدگی مدل و مراحل رمزگشایی ممکن است کارایی آن را در کاربردهای لحظه‌ای کاهش دهد. وزنیک و همکاران [۱۶]، مدل مبتنی بر ترکیب حافظه‌ی طولانی کوتاه مدت دوطرفه (Bi-LSTM)^{۱۵} و درخت تصمیم (DT)^{۱۶} را در سیستم‌های اینترنت اشیا معرفی کردند که به کمک روش‌های متعادل‌سازی چون SMOTE-Tomek

عملکرد تشخیصی بالایی را ثبت کرد. دقت تشخیص بیشتر از ۹۶٪ و جامعیت بالای ۹۶٪ نشان از توانایی بالای این سیستم دارد. این پژوهش، نتایج مطلوبی برای شناسایی سرطان دهانه‌ی رحم را در بستر اینترنت اشیا گزارش کرد، اگر چه وابستگی زیاد به پیش‌پردازش داده از جمله محدودیت‌های آن بود. تابناه و همکاران [۱۷]، تأثیر نسبت‌های مختلف عدم تعادل طبقه بر دقت مدل‌های طبقه‌بندی را بررسی کردند. آن‌ها نشان دادند که رابطه بین نسبت عدم تعادل و دقت مدل، به صورت منحنی محدب است، یعنی در برخی سطوح عدم تعادل، کاهش عملکرد شدیدتر می‌شود. مزیت این مطالعه در ارائه چارچوب تحلیلی جامع است، هرچند بر ارائه راهکار عملی تمرکز نداشت. ژو و همکاران [۱۸]، الگوریتمی ترکیبی متشکل از خوشه‌بندی و SMOTE را معرفی کردند که با درج خطی "نمونه‌های امن"، توانست در ۸ مجموعه داده UCI عملکردی فراتر از سایر روش‌ها ثبت کند. حساسیت ۹۹٫۸۴٪ و اختصاصیت^{۱۷} ۹۹٫۵۶٪ در داده‌های پزشکی از مزایای برجسته آن بود. با این حال، اتکای بالا به دقت خوشه‌بندی اولیه، ریسک تولید داده‌های گمراه‌کننده را افزایش می‌دهد. اووه و همکاران [۱]، کارایی روش‌های مختلف بیش‌نمونه‌برداری و زیرنمونه‌برداری را در چهار مجموعه داده بهداشتی بررسی کردند. آن‌ها دریافتند که روش نزدیک به حادثه^{۱۸} در ترکیب با رگرسیون لجستیک، در مواجهه با انحرافات داده بسیار مؤثر است. این مطالعه راهنمای عملی برای انتخاب روش نمونه‌برداری فراهم کرد، اگرچه اثربخشی در سایر حوزه‌ها نیازمند بررسی بیشتر است. نظر و همکاران [۱۹]، با مرور نظام‌مند ۲۲ مطالعه، عملکرد مدل‌های AI را در پیش‌بینی پاسخ به درمان CRT ارزیابی کردند. مدل‌های نظارت‌شده دقت بالای ۸۵٪ و AUC برابر ۰٫۸۶ داشتند که از عملکرد دستورالعمل‌های سنتی بهتر بود. این پژوهش پتانسیل AI را در تصمیم‌گیری کلینیکی نشان داد، ولی تأکید داشت که یافته‌ها باید در مطالعات تصادفی کنترل‌شده تأیید شوند. نوگویان و همکاران [۲۰]، نیز به بررسی روش‌های ارزیابی ناهم‌زمانی الکتریکی در درمان CRT پرداختند. آن‌ها مزایا و معایب

13 Quad Division Prototype Selection-based k-Nearest Neighbor classifier (QDPSKNN)

14 Hierarchical clustering

15 Bidirectional Long Short-Term Memory (Bi-LSTM)

16 Decision Tree (DT)

17 Specificity

18 NearMiss

روش‌های جایگزین از جمله وکتورهای گراف و نگاشت الکترو-آناتومیک را برشمردند. با وجود نتایج امیدبخش، این روش‌ها هنوز به‌طور گسترده در عمل بالینی پذیرفته نشده‌اند. پیشنهاد این مطالعه بر استفاده از هوش مصنوعی (AI)^{۱۹} برای بهبود پیش‌بینی درمان بود، گرچه جنبه‌های عملی آن کمتر مورد بحث قرار گرفت. در مطالعه دویفبری و همکاران [۲۱]، روش‌های SMOTE سنتی و ترکیبی تطبیقی برای داده‌های با عدم تعادل بالا استفاده شد که نتایج دقت متعادل تا ۹۹٫۰۳٪ را در طبقه‌بندی جنگل تصادفی (RF)^{۲۰} ثبت کرد. این پژوهش اثربخشی بیش‌نمونه‌برداری بر اساس دشواری یادگیری را تأیید کرد، ولی در داده‌های با نسبت عدم تعادل بسیار بالا عملکرد روش‌ها کاهش یافت. در نهایت، سلیمانی و همکاران [۲۲]، با پیشنهاد یک روش نمونه‌برداری پویا برای DL، به بهبود دقت و کاهش خطای شبکه‌های عصبی (NN)^{۲۱} در داده‌های چندطبقه پرداختند. آن‌ها با آزمایش بر ۱۳ مجموعه داده، نشان دادند که روش پیشنهادی از سایر رویکردهای متداول مؤثرتر عمل می‌کند. هرچند این رویکرد نیاز به تنظیم دقیق پارامترها داشته و برای داده‌های لحظه‌ای کاربرد فوری ندارد.

در حوزه‌ی تشخیص بیماری توسط الگوریتم‌های یادگیری ماشین، پژوهش شریفی مهرجرد و همکاران [۲۳]، از سیگنال EEG و الگوریتم‌های یادگیری ماشین برای تشخیص اوتیسم بهره برد. روش کار این تحقیق، مرور سیستماتیک مقالات و مقایسه آن‌ها از نظر مراحل پردازش سیگنال (پیش‌پردازش، استخراج و انتخاب ویژگی) و نوع طبقه‌بندی به‌کاررفته است. در مقالات بررسی‌شده از الگوریتم‌هایی مثل SVM، KNN، DT، RF و مخصوصاً شبکه‌های عصبی کانولوشنی^{۲۲} و روش‌های یادگیری عمیق استفاده شده و نتیجه کلی این است که CNN‌ها معمولاً بهترین دقت را برای تشخیص اوتیسم از EEG می‌دهند. با این حال، معایبی مثل وابستگی زیاد مدل‌ها به دیتاست‌های کوچک و غیراستاندارد، دشواری تعمیم به جمعیت‌های مختلف، نیاز به داده با کیفیت و حجم زیاد، و نبود یک پروتکل واحد

برای پیش‌پردازش و استخراج ویژگی به‌عنوان محدودیت‌های مهم این حوزه مطرح شده است. نژاده و همکارانش [۲۴]، یک رویکرد بردار ویژگی مبتنی بر تقویت بردار ویژگی برای پیش‌بینی پاسخ به درمان CRT پیاده‌سازی کردند. این مطالعه با استفاده از خوشه‌بندی وزنی خودکار، انتخاب و ترکیب ویژگی‌های بالینی را بهبود بخشیده و توانسته است دقت پیش‌بینی را در مقایسه با روش‌های سنتی افزایش دهد. یکی از نقاط قوت این مطالعه، تمرکز بر بهینه‌سازی بردار ویژگی با خوشه‌بندی هوشمند است که به یافتن الگوهای پیچیده در داده‌های بالینی کمک می‌کند. با این حال، مطالعه حاضر همچنین به دلیل فقدان یک رویکرد مستقل و جامع برای مدیریت و بهبود کلاس اقلیت (بیماران بدون پاسخ) دارای محدودیت است که می‌تواند بر عملکرد مدل در مورد داده‌های بسیار نامتوازن تأثیر بگذارد. علاوه بر این، اتکای زیاد به اندازه‌ی ابعاد دارد ممکن است کارایی روش را در مجموعه داده‌هایی با توزیع پیچیده‌تر محدود کند. در پژوهشی دیگر نژاده و همکاران [۲۵]، در پژوهش خود به توسعه مدلی برای پیش‌بینی پاسخ به CRT با بهره‌گیری از روش‌های DL و تحلیل ویژگی‌ها پرداختند. آن‌ها داده‌های مربوط به ۲۰۹ بیمار مبتلا به نارسایی قلبی را که شامل ۶۰ ویژگی دموگرافیک، بالینی، آزمایشگاهی، نوار قلب و اکوکاردیوگرافی بود، گردآوری کردند. مراحل پردازش شامل استخراج ویژگی‌های پایه با استفاده از شاخص‌های کیفی، بیش‌نمونه‌برداری برای مقابله با عدم تعادل داده‌ها، خوشه‌بندی سلسله‌مراتبی و انتخاب ویژگی‌های کلیدی از خوشه‌های با خطای پایین بود. مدل DL پیشنهادی، متشکل از لایه‌های متراکم و کانولوشن، توانست با دقت ۹۱٫۸۵٪، ناحیه زیر منحنی ROC برابر ۰٫۹۵۷ و حساسیت ۹۴٫۲۲٪ عملکرد بسیار خوبی در پیش‌بینی پاسخ به درمان ارائه دهد. با این حال، این پژوهش دارای برخی محدودیت‌ها نیز می‌باشد. از جمله مهم‌ترین آن‌ها می‌توان به اندازه محدود نمونه (۲۰۹ بیمار) اشاره کرد که ممکن است بر قابلیت تعمیم‌پذیری نتایج به جمعیت‌های بزرگ‌تر تأثیرگذار باشد. همچنین، نبود اعتبارسنجی خارجی با استفاده از داده‌های مستقل از مراکز درمانی دیگر، احتمال بیش‌برازش مدل را افزایش می‌دهد. علاوه بر این، اگرچه روش‌های

19 Artificial Intelligence (AI)

20 Random Forest (RF)

21 Neural Networks (NNs)

22 Convolutional Neural Network (CNN)

است. در مسأله‌های طبقه‌بندی نامتوازن، معیارهایی مانند صحت، جامعیت^{۲۴} و به ویژه میانگین هارمونیک کاربرد بیشتری دارند، زیرا تمرکز اصلی آن‌ها بر روی عملکرد مدل در شناسایی نمونه‌های کلاس اقلیت است. سوم این‌که به منظور افزایش دقت و قابلیت تعمیم مدل طبقه‌بندی، از روش‌های یادگیری گروهی استفاده شده است. تجمیع چند مدل ضعیف یا قوی در قالب یک مدل گروهی (مثلاً روش‌های مبتنی بر تجمیع بوت‌استرپ^{۲۵} یا تقویتی^{۲۶}) معمولاً عملکرد بهتری نسبت به یک مدل منفرد ارائه می‌دهد. به کمک این نوآوری‌ها، هدف نهایی این پژوهش بهبود معیارهای تشخیص پاسخ به درمان CRT به شکلی دقیق و پایدار دنبال شده است. شکل ۱ روند کلی مراحل روش پیشنهادی را نشان می‌دهد.



بیش نمونه‌برداری برای مقابله با عدم تعادل داده‌ها به کار گرفته شده‌اند، اما این رویکردها ممکن است در برخی موارد به تکرار الگوهای مصنوعی و کاهش تنوع داده‌ها منجر شوند. تلاش‌های پژوهش‌های اخیر نشان‌دهنده تلاش‌های مختلف در حوزه پزشکی برای مقابله با چالش داده‌های نامتعادل است و چگونگی بهبود عملکرد مدل‌های یادگیری ماشینی را با استفاده از روش‌های مختلف نمونه‌برداری و الگوریتم‌های پیشرفته نشان می‌دهد. در بخش بعد روشی نوآورانه برای بهبود این نواقص مطرح می‌شود.

۳. روش پیشنهادی

با توجه به نتایج به دست آمده از بررسی مطالعات پیشین و شناسایی چالش‌های موجود در تشخیص دقیق پاسخ بیماران به درمان CRT، می‌توان دریافت که نبود توازن در داده‌ها، انتخاب غیربینه ویژگی‌ها، و کارایی پایین برخی مدل‌های طبقه‌بندی، از موانع اصلی در بهبود عملکرد سیستم‌های پشتیبان تصمیم‌گیری در این حوزه هستند. در پاسخ به این چالش‌ها، این پژوهش یک چارچوب جدید و ترکیبی را پیشنهاد می‌دهد که با بهره‌گیری از روش‌های هوشمند بهینه‌سازی، نمونه‌سازی، و یادگیری گروهی، تلاش دارد دقت و پایداری پیش‌بینی پاسخ به درمان CRT را بهبود بخشد. در ادامه، مراحل این روش پیشنهادی به تفصیل تشریح شده‌اند. در روش پیشنهادی ما، سه نوآوری اصلی لحاظ شده است: اول این‌که الگوریتم SMOTE سنتی برای نمونه‌سازی اقلیت با الگوریتم CSA ترکیب شده است تا فرآیند نمونه‌سازی نامتوازن بهینه گردد. روش SMOTE به‌عنوان یک تکنیک شناخته شده برای حل مشکل عدم تعادل کلاس‌ها به کار می‌رود، اما پارامترهای آن (مانند تعداد همسایه‌ها و تعداد نمونه‌های افزوده) معمولاً به صورت تجربی تعیین می‌شوند. در این پژوهش با استفاده از الگوریتم CSA، پارامترهای مربوط به SMOTE به شکلی بهینه انتخاب شده‌اند تا تعادل کلاس‌ها با کمترین خطای افزوده ممکن به دست آید. الگوریتم CSA یک الگوریتم فراابتکاری جدید مبتنی بر رفتار اجتماعی کلاغ‌ها در ذخیره و بازیابی غذا است. دوم این‌که از معیار میانگین هارمونیک، صحت^{۲۳} و حساسیت به‌عنوان معیار ارزیابی نهایی استفاده شده

24 Recall

25 Bagging

26 Boosting

23 Precision

شکل ۱- روش پیشنهادی جهت بهبود معیار میانگین هارمونیک تشخیص CRT

در اولین گام، مجموعه داده بیماران CRT بارگذاری و معرفی می شود. داده های مورد استفاده در این پژوهش شامل اطلاعات بالینی بیمارانی است که تحت درمان CRT قرار گرفته اند. این داده ها می توانند شامل مشخصات فردی (سن، جنسیت)، نتایج آزمایش های بالینی (مانند کسر جهشی بطن چپ، طول QRS در الکتروکاردیوگرافی) و سایر شاخص های مرتبط (نظیر جرم اسکار قلب) باشند. پاسخ دهی مثبت به CRT برای هر بیمار مطابق با معیار تعریف شده در مطالعات مشابه (به عنوان مثال، افزایش حداقل ۱۰٪ در کسر جهشی بطن چپ پس از یک سال) تعیین می شود. به این ترتیب، هر نمونه در مجموعه داده دارای ویژگی های ورودی و برچسب دوگانه «پاسخ دهنده» یا «غیر پاسخ دهنده» است. پس از بارگذاری داده ها، پیش پردازش داده ها انجام می شود. ابتدا ویژگی های پیوسته موجود (نظیر مقادیر مداوم آزمایشگاهی) با استفاده از روش های گسسته سازی مناسب به بازه های گسسته تقسیم شدند. انجام گسسته سازی روی متغیرهای پیوسته باعث بهبود عملکرد بسیاری از طبقه بندها می شود؛ به طوری که نشان داده شده این مرحله می تواند کیفیت طبقه بندی را به طرز چشمگیری افزایش دهد. پس از گسسته سازی، استخراج ویژگی های مفید صورت می گیرد. در این مرحله، شاخص های مرتبط با پاسخ به CRT (مانند اندازه اسکار میوکاردا، نمایه های دینامیکی الکتریکی و پارامترهای هم تراز قلب) شناسایی و استخراج می شوند. انتخاب یا استخراج ویژگی های مرتبط پیش از آموزش مدل ضروری است، چرا که تمرکز بر مهم ترین ویژگی ها می تواند دقت پیش بینی را بهبود دهد. در نهایت یک مجموعه ویژگی های نهایی بدست آمده که ابعاد آن نسبت به فضای اولیه کاهش یافته و کارایی الگوریتم های بعدی را افزایش می دهد. گام بعدی مربوط به متعادل سازی کلاس ها با رویکرد ترکیبی CSA-SMOTE است. از آنجایی که در داده های CRT معمولاً تعداد پاسخ دهندگان بسیار کمتر از غیر پاسخ دهندگان است، برای کاهش عدم تعادل داده ها از تکنیک نمونه سازی مصنوعی استفاده می شود. در روش SMOTE سنتی نمونه های جدیدی برای کلاس اقلیت ایجاد

می شود تا تعادل در کلاس ها حاصل گردد. در این پژوهش اما، به جای تعیین ثابت پارامترهای SMOTE، از الگوریتم CSA برای بهینه سازی این پارامترها کمک گرفته ایم. در رویکرد CSA-SMOTE هر «کلاغ» نماینده یک راه حل بالقوه است که می تواند شامل پارامترهای SMOTE (مانند تعداد همسایه ها) یا نمونه های مصنوعی تولیدی باشد. الگوریتم CSA مطابق با رفتار هوشمندانه کلاغ ها عمل می کند؛ به طوری که هر کلاغ با احتمال (AP-1) به موقعیت ذخیره شده غذای کلاغ دیگری حرکت می کند (که معادل جستجوی محلی و بهینه سازی راه حل های موجود است) و با احتمال AP به نقطه ای تصادفی جهش می کند (که موجب کاوش سراسری فضای پاسخ می شود). پارامترهای اصلی CSA شامل اندازه کلاغ^{۲۷}، حداکثر تعداد تکرار^{۲۸}، طول پرواز^{۲۹} و احتمال آگاهی^{۳۰} می باشد. این مکانیسم ترکیبی بین کاوش و بهره برداری امکان پذیر است، به گونه ای که پس از چندین تکرار، بهترین مجموعه پارامترهای SMOTE مطابق با معیار کارایی (مثلاً میانگین هارمونیک) انتخاب شده و نمونه های افزوده نهایی تولید می گردد. به این ترتیب، روش CSA-SMOTE با حفظ تنوع در نمونه سازی، تعادل مناسبی در توزیع کلاس ها فراهم کرده و از بیش برآزش داده های اقلیت جلوگیری می کند. پس از متعادل سازی داده ها، مرحله انتخاب ویژگی انجام می شود. در این مرحله از الگوریتم تقویت گرادیان (GB)^{۳۱} استفاده شده است تا ویژگی های با اهمیت بالا انتخاب گردند. الگوریتم GB به صورت تدریجی چندین DT را آموزش می دهد و بر اساس کاهش خطای مدل اهمیت هر ویژگی را ارزیابی می کند. پس از اجرای الگوریتم تقویتی، ویژگی های با امتیاز اهمیت بالا به عنوان زیرمجموعه نهایی انتخاب می شوند. انتخاب ویژگی بر اساس اهمیت یادگیری شده، علاوه بر حفظ قدرت پیش بینی مدل، باعث کاهش ابعاد مسئله و ساده تر شدن مدل نهایی می شود. در نهایت، مجموعه داده متعادل و کاهش یافته به مدل طبقه بند وارد شده و با روش های یادگیری گروهی آموزش داده می شود. به

27 Number (N)

28 Max Iterations

29 Flight Length

30 Awareness Probability

31 Gradient Boosting (GB)

کارگیری روش‌های یادگیری گروهی مانند میانگین‌گیری گروهی یا **Bagging** یا روش‌های مشابه موجب می‌شود چندین طبقه‌بند پایه با هم ترکیب شده و مدل نهایی مقاوم‌تر و دقیق‌تری شکل بگیرد؛ همان‌طور که مطالعات نشان داده‌اند یادگیری گروهی معمولاً برتری چشمگیری نسبت به مدل‌های منفرد در مسائل نامتوازن دارد. عملکرد مدل نهایی با اعتبارسنجی **KCV** ارزیابی شده و معیارهای دقت، صحت، حساسیت، اختصاصیت و میانگین هارمونیک محاسبه می‌شوند. این معیارها با استفاده از ماتریس پیش‌بینی تعیین شده و توانایی مدل را در شناسایی دقیق پاسخ‌دهندگان **CRT** نشان می‌دهند. در ادامه مراحل پر اهمیت پیش‌پردازش، متعادل‌سازی طبقات به روش پیشنهادی و رویکرد طبقه‌بندی پژوهش شرح داده می‌شود.

۱.۳. پیش‌پردازش در روش پیشنهادی

پیش‌پردازش داده‌ها یکی از مراحل مهم در توسعه یک مدل **ML** است و تأثیر مستقیمی بر کیفیت و دقت نتایج دارد. در این مطالعه، مجموعه‌ای از تکنیک‌های پیش‌پردازش برای تهیه داده‌های بالینی بیماران **CRT** اعمال شد.

- حذف ویژگی‌های غیرموثر: در مرحله اول، ویژگی‌های غیرموثر از بردار اصلی حذف می‌شوند. برای این منظور، بهره اطلاعاتی (**IG**)^{۳۲} هر ویژگی محاسبه می‌شود. بهره اطلاعاتی معیاری است که میزان کاهش آنتروپی یا عدم قطعیت در طبقه‌بندی پس از استفاده از یک ویژگی خاص را نشان می‌دهد. ویژگی‌هایی که بهره اطلاعاتی بسیار کمی دارند و توانایی قابل توجهی در تمایز بین کلاس‌های پاسخ و غیرپاسخ نشان نمی‌دهند، از مجموعه داده‌ها حذف می‌شوند. این روش باعث کاهش ابعاد فضای ویژگی و افزایش کارایی محاسباتی می‌شود.

- گسسته‌سازی ویژگی‌های پیوسته: پس از حذف ویژگی‌های غیرگسسته، ابتدا ویژگی‌های پیوسته مانند مقادیر آزمایشگاهی حذف می‌شوند و سپس با استفاده از روش‌های گسسته‌سازی مناسب، ویژگی‌ها به فواصل گسسته تقسیم می‌شوند. برای این منظور از روش سطل (سطوح مساوی) استفاده شده است که در آن دامنه هر ویژگی به چندین بازه با عرض یکسان تقسیم

می‌شود. انجام گسسته‌سازی روی متغیرهای پیوسته منجر به عملکرد بهتر بسیاری از طبقه‌بندی‌کننده‌ها می‌شود، به طوری که مطالعات نشان داده‌اند که این مرحله می‌تواند کیفیت طبقه‌بندی را به طور قابل توجهی بهبود بخشد. تعداد فواصل برای هر ویژگی بر اساس توزیع و ماهیت بالینی داده‌ها تعیین خواهد شد. - نرمال‌سازی و مقیاس‌بندی ویژگی: به دلیل تغییر در محدوده ویژگی‌های مختلف، فرآیند نرمال‌سازی ویژگی اعمال می‌شود. برای ویژگی‌های پیوسته که گسسته‌سازی نشده‌اند، از روش استانداردسازی امتیاز-زد^{۳۳} استفاده شده است که در آن هر ویژگی به میانگین صفر و انحراف معیار یک تبدیل می‌شود. این تبدیل تضمین می‌کند که همه ویژگی‌ها در مقیاس یکسانی هستند و هیچ‌کدام، به دلیل داشتن محدوده عددی بزرگتر، به طور نامتناسبی بر فرآیند یادگیری مدل تأثیر نخواهند گذاشت.

- کلاس‌های متعادل با رویکرد ترکیبی **CSA-SMOTE**: رویکردهای ترکیب الگوریتم‌های فراابتکاری در حوزه‌های مختلف در بسیاری از پژوهش‌ها [۲۶، ۲۷]، منجر به بهبود عملکرد شده‌اند. از اینرو پژوهش جاری تکنیک ترکیبی را برای متعادل‌سازی طبقات توصیه می‌کند. معمولاً تعداد نمونه‌های غیر پاسخگو در داده‌های **CRT** بسیار بیشتر از تعداد پاسخ‌دهندگان است. برای رفع این مشکل، عدم تعادل داده‌ها با استفاده از یک تکنیک نمونه‌گیری مصنوعی کاهش می‌یابد. در روش سنتی **SMOTE**، نمونه‌های جدیدی برای کلاس اقلیت تولید می‌شوند تا تعادل کلاس حاصل شود. در این مطالعه، به جای ثابت کردن پارامترهای **SMOTE**، از الگوریتم **CSA** برای بهینه‌سازی این پارامترها استفاده شد. این الگوریتم پارامترهای **SMOTE** مانند تعداد همسایه‌ها را به طور بهینه انتخاب می‌کند تا تعادل کلاس با کمترین خطای اضافی ممکن حاصل شود.

- انتخاب ویژگی بر اساس تکنیک الگوریتم ژنتیک: پس از متعادل‌سازی، مرحله شناسایی ویژگی‌های مفید توسط الگوریتم ژنتیک (**GA**)^{۳۴} تنظیم می‌شود. تابع هدف در این مرحله بر اساس الگوریتم **GBT** تعریف می‌شود. **GA** به عنوان یک روش جستجوی تکاملی، زیرمجموعه‌های مختلفی از ویژگی‌ها را

³³ Z-Score

³⁴ Genetic Algorithm (GA)

³² Information Gain (IG)

ارزیابی می‌کند و با کمک عملگرهای انتخاب، تقاطع و جهش به تدریج به بهترین مجموعه ویژگی‌ها همگرا می‌شود. استفاده از GBT به عنوان یک تابع ارزیابی و تعیین معیار میانگین هارمونیک جهت برازش، یافتن ویژگی‌هایی را که بیشترین قدرت پیش‌بینی را برای تشخیص پاسخ‌دهندگان CRT دارند، امکان‌پذیر می‌سازد. این فرآیند، جدا از حفظ قدرت پیش‌بینی مدل و کاهش ابعاد مسئله، بر ساده‌سازی مدل نهایی نیز تأثیر می‌گذارد.

- کدگذاری و تبدیل ویژگی‌های دسته‌بندی‌شده: متغیرهای دسته‌بندی‌شده غیررتبه‌ای با استفاده از کدگذاری وان‌هات^{۳۵} به ویژگی‌های دودویی تبدیل می‌شوند. برای متغیرهای ترتیبی، کدگذاری عددی به طور مناسب اعمال شد که ترتیب طبیعی دسته‌بندی‌ها را در معماری مدل حفظ می‌کند.

پس از تکمیل این مراحل، یک مجموعه ویژگی بهینه‌شده با ابعادی که در مقایسه با فضای اولیه به طور قابل توجهی کاهش یافته به دست آمد که کارایی الگوریتم طبقه‌بندی پژوهش را بهبود می‌دهد. فرآیند پیش‌پردازش مشخص شده می‌تواند زمینه را برای آموزش مدل‌های گروهی با دقت و اطمینان بالا فراهم کند.

۲.۳. الگوریتم CSA در روش پیشنهادی

الگوریتم CSA یک روش فراابتکاری است که با الهام از رفتار طبیعی کلاغ‌ها در پنهان و پیدا کردن غذا عمل می‌کند. این الگوریتم از این رفتار برای جستجوی بهترین راه‌حل‌ها در یک فضای جستجو استفاده می‌کند. در CSA، هر کلاغ نماینده راه‌حل است و در طول فرآیند بهینه‌سازی، کلاغ‌ها به دنبال بهترین موقعیت‌ها برای پنهان کردن غذا (راه حل) هستند. انتخاب الگوریتم CSA به دلیل سادگی و کارایی‌اش برای حل مسائل بهینه‌سازی مختلف بوده است و تنظیم مناسب پارامترهای آن می‌تواند به بهبود عملکرد الگوریتم منجر گردد. در ادامه مراحل اصلی و پارامترهای مهم الگوریتم CSA آورده شده و سپس نحوه‌ی بکارگیری الگوریتم CSA با الگوریتم SMOTE برای دستیابی به هدف روش پیشنهادی شرح داده می‌شود.

- مقدمه و مقداردهی اولیه:

ابتدا تعداد کلاغ‌ها (که نشان‌دهنده تعداد راه‌حل‌های اولیه است) و موقعیت اولیه آن‌ها در فضای جستجو به صورت تصادفی تعیین می‌شود.

- به‌روزرسانی موقعیت کلاغ‌ها:

هر کلاغ به سمت بهترین مکان شناخته‌شده خود یا مکان‌های دیگر کلاغ‌ها حرکت می‌کند. اگر کلاغی به موقعیت بهتری دست یابد، آن را به عنوان مکان مخفی جدید ذخیره می‌کند.

- حافظه و تصادف:

کلاغ‌ها ممکن است به صورت تصادفی مکان‌های جدیدی پیدا کنند و از حافظه خود برای بهبود موقعیت‌ها استفاده کنند، که این امر موجب جستجوی موثرتر در فضای راه‌حل می‌شود.

- محدودیت‌ها و اعتبارسنجی:

هر کلاغ موقعیت جدید خود را بررسی می‌کند و در صورت معتبر بودن، آن را به‌روزرسانی می‌کند.

- تکرار فرایند:

این مراحل تا رسیدن به شرایط توقف مانند تعداد تکرارهای معین یا یافتن راه‌حل بهینه ادامه دارد.

همچنین پارامترهای الگوریتم عبارتند از:

- تعداد کلاغ‌ها:

تعیین‌کننده تعداد راه‌حل‌های اولیه در فضای جستجو است. افزایش این تعداد ممکن است دقت جستجو را افزایش دهد ولی زمان محاسباتی را نیز بیشتر می‌کند.

- حداکثر تکرارها:

تعداد تکرارهایی که الگوریتم برای یافتن راه‌حل بهینه انجام می‌دهد.

- احتمال آگاهی:

تعیین‌کننده احتمال حرکت یک کلاغ به مکان جدید به جای دنبال کردن مکان بهترین کلاغ است، که بین ۰ و ۱ تنظیم می‌شود.

- طول پرواز:

حداکثر مسافتی که هر کلاغ می‌تواند در یک تکرار طی کند، و بر تغییر موقعیت کلاغ‌ها تأثیر دارد.

- حافظه کلاغ‌ها^{۳۶}:

ویژگی‌ها است که به‌عنوان نمونه‌های مصنوعی دقیق‌تر و مؤثرتر برای مدل‌سازی طبقه اقلیت عمل کنند.

- به‌روزرسانی و جایگزینی نمونه‌ها:

نمونه‌های مصنوعی بهبود یافته توسط CSA جایگزین نمونه‌های اولیه می‌شوند. این نمونه‌ها که حالا بهتر نماینده کلاس اقلیت هستند، به داده‌های اصلی اضافه می‌شوند.

با بهینه‌سازی نمونه‌های مصنوعی، دقت مدل در شناسایی کلاس اقلیت به‌طور قابل‌توجهی افزایش می‌یابد، که این امر به کاهش نرخ خطا و بهبود توانایی مدل در تشخیص موارد نادر کمک می‌کند. الگوریتم CSA توانایی تولید نمونه‌های مصنوعی بهتری را دارد که هم‌پوشانی کمتری با دیگر کلاس‌ها دارند، و این موضوع در داده‌های پیچیده و دارای ویژگی‌های متنوع بسیار اهمیت دارد. علاوه بر این، CSA موجب افزایش جامعیت و صحت طبقه‌بندی می‌شود و توانایی مدل را در شناسایی دقیق موارد کلاس اقلیت به ویژه در کاربردهای حساس، مانند پیش‌بینی‌های پزشکی، بهبود می‌بخشد. اگرچه استفاده از این الگوریتم فرآیند کارایی ممکن است پیچیدگی محاسباتی بیشتری به همراه داشته باشد، اما به دلیل ارائه نتایج بهینه، ارزش افزوده قابل‌توجهی را به همراه دارد و با کاهش تعداد داده‌ها و افزایش کارایی، نتایج بهتری را ارائه می‌دهد.

۳.۳. الگوریتم گرادیان درختان تقویت‌شده

روش تقویتی که بر DT متکی است به‌عنوان درخت تقویتی شناخته می‌شود. مدل گرادیان درختان تقویت‌شده^{۳۸} نشانگر اجرای مؤثر و کارآمد مدل درخت تقویتی است. GBT در واقع لیفتینگ درخت انتها به انتهای به‌شدت مقیاس‌پذیر است. اساساً، مدل درخت تقویتی را می‌توان به‌عنوان گروهی از درختان تصمیم در نظر گرفت. شنگ و چنگ در پژوهش خود توضیح گسترده و کاملی از این روش را ارائه کرده‌اند [۲۸]. با توجه به اینکه درخت لیفتینگ اولیه برابر $\hat{y}_i^0 = f_0(x_i) = 0$ است، مدل حل مشکل (مقدار پیش‌بینی شده) نمونه i^{th} در مرحله t برابر است با:

هر کلاغ بهترین موقعیت پیدا شده را به‌عنوان حافظه ذخیره می‌کند و در تصمیم‌گیری برای حرکات بعدی از آن استفاده می‌کند.

- شرایط توقف^{۳۷}:

معیارهایی که موجب توقف الگوریتم می‌شوند، مانند رسیدن به تعداد معینی از تکرارها یا یافتن راه‌حل با دقت مطلوب.

در روش‌های ML، یکی از چالش‌های اصلی هنگام کار با داده‌های نامتعادل، حضور کلاس‌های با نمونه‌های کمتر است که می‌تواند منجر به آموزش نادرست مدل‌ها شود. برای مقابله با این مشکل، از تکنیک‌های مختلف بیش‌نمونه‌برداری استفاده می‌شود. یکی از روش‌های شناخته‌شده در این زمینه، الگوریتم SMOTE است. الگوریتم SMOTE با تولید نمونه‌های مصنوعی جدید از داده‌های اقلیت تلاش می‌کند تا تعادل بهتری بین کلاس‌ها ایجاد کند. با این حال، نمونه‌های ایجاد شده توسط SMOTE ممکن است بهینه نباشند و همچنان مشکلاتی مانند هم‌پوشانی کلاس‌ها یا تولید نمونه‌های غیردقیق وجود داشته باشد. الگوریتم CSA به‌عنوان یک روش مؤثر می‌تواند به عنوان یک مرحله بهینه‌سازی پس از اجرای SMOTE به کار گرفته شود. در این روش، نمونه‌های مصنوعی ایجاد شده توسط SMOTE به عنوان نقاط شروع در فضای جستجو در نظر گرفته می‌شوند. سپس، CSA به جستجوی نمونه‌های بهتر در فضای ویژگی‌ها می‌پردازد. این کار با استفاده از رفتار جستجوی کلاغ‌ها انجام می‌شود که در جستجوی بهترین مکان برای پنهان کردن غذا هستند، یعنی در اینجا جستجوی نمونه‌هایی که نماینده بهتری از داده‌های اقلیت باشند. مراحل کار بکار رفته در روش پیشنهادی بصورت زیر است:

- بیش‌نمونه‌برداری اولیه با SMOTE:

ابتدا با استفاده از SMOTE، تعدادی نمونه مصنوعی از داده‌های اقلیت تولید می‌شود. این نمونه‌ها به منظور کاهش عدم‌تعادل میان کلاس‌ها به داده‌های اصلی اضافه می‌شوند.

- بهبود نمونه‌ها با CSA:

در مرحله بعد، الگوریتم CSA روی نمونه‌های تولید شده توسط SMOTE اعمال می‌شود. هدف CSA یافتن نقاط بهینه در فضای

³⁸ Gradient Boosted Trees (GBT)

³⁷ Stopping Criteria

داده شده است [29]. این روش به ویژه در مسائل پیچیده و داده‌های متنوع می‌تواند کارایی بسیار بالایی داشته باشد.

GBT: این الگوریتم یادگیری گروهی به‌طور مداوم مدل‌های ضعیف، معمولاً درختان تصمیم، را به صورت ترتیبی آموزش می‌دهد. هر درخت جدید به‌منظور کاهش خطای مدل‌های قبلی و بهبود پیش‌بینی‌ها اضافه می‌شود. فرآیند یادگیری در این الگوریتم بر مبنای کاهش تدریجی خطا و تصحیح اشتباهات مدل‌های قبلی استوار است. الگوریتم GBT به دلیل توانایی بالا در مدل‌سازی روابط پیچیده و انعطاف‌پذیری در برخورد با داده‌های غیرخطی شناخته شده است.

RF: این الگوریتم نیز از یادگیری گروهی درختان تصمیم استفاده می‌کند، اما با رویکردی متفاوت. در RF، مجموعه‌ای از درختان تصمیم به‌طور هم‌زمان آموزش داده می‌شوند، و پیش‌بینی‌ها به‌طور میانگین یا با رأی‌گیری از درختان مختلف ترکیب می‌شود. هر درخت به‌طور تصادفی بر روی زیرمجموعه‌ای از ویژگی‌ها و داده‌ها آموزش می‌بیند، که باعث کاهش همبستگی بین درختان و بهبود تعمیم‌پذیری مدل می‌شود.

روش یادگیری گروهی: در این روش، قدرت هر دو الگوریتم از طریق ترکیب نتایج آنها به حداکثر می‌رسد. این کار در روش پیشنهادی با استفاده از تکنیک رأی‌گیری انجام می‌شود.

ترکیب الگوریتم‌های GBT و RF می‌تواند به بهبود قابل توجهی در دقت مدل‌ها منجر شود، زیرا هر الگوریتم ضعف‌های دیگری را جبران می‌کند. این روش به کاهش واریانس و خطا کمک می‌کند؛ به‌طوری‌که RF با کاهش واریانس و GBT با تمرکز بر کاهش خطا، عملکرد مدل را تقویت می‌کند. علاوه بر این، روش گروهی بهبود قابلیت تعمیم مدل‌ها به داده‌های پیچیده و غیرخطی را نیز تسهیل می‌کند. به‌طور کلی، استفاده از این روش ترکیبی باعث بهره‌برداری بهینه از نقاط قوت هر الگوریتم و کاهش نقاط ضعف آنها می‌شود و به این ترتیب، کارایی مدل‌های ML به‌طور چشمگیری افزایش می‌یابد.

۵.۳. مجموعه داده مورد ارزیابی تحقیق

در فرآیند پیش‌بینی مبتنی بر ML، استفاده از داده‌های پزشکی جمع‌آوری شده ضروری است. این بخش، مجموعه‌ی

$$\hat{y}_i^t = \sum_{k=1}^t f_k(x_i) = \hat{y}_i^{t-1} + f_t(x_i) \quad (1)$$

در اینجا x_i بردار ویژگی نمونه i ام و \hat{y}_i^t مقدار پیش‌بینی مرحله t ام است. متغیر $f_t(x_i)$ ، DT افزوده در مرحله t ام می‌باشد. تابع هدف این الگوریتم به‌صورت زیر نشان داده می‌شود:

$$Obj^t = \sum_{i=1}^n l(y_i, \hat{y}_i^t) + \sum_{i=1}^t \Omega(f_i), \quad (2)$$

که در آن l نشانگر تابع ضرر محدب، $\Omega(f_i)$ تابعی برای محاسبه پیچیدگی درخت f_i برای جلوگیری از بیش‌برازش مدل است. همچنین داریم:

$$\sum_{i=1}^t \Omega(f_i) = \sum_{i=1}^{t-1} \Omega(f_i) + \Omega(f_t) \quad (3)$$

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 \quad (4)$$

که در آن T نشانگر تعداد برگ‌ها، و w امتیاز گره برگ را نشان می‌دهد. λ و γ نیز ضرایب تنظیم برای کنترل پیچیدگی و جریمه مدل‌سازی هستند. رابطه ۲ فرم کلی تابع هدف است. اما رابطه ۵، تقریب درجه دوم از رابطه‌ی ۲ است که آن را بهینه می‌کند و در الگوریتم‌های تقویت درختی استفاده می‌شود.

$$Obj^t = \sum_{i=1}^n \left[g_i f_i(x_i) + \frac{1}{2} h_i f_i^2(x_i) \right] + \Omega(f_i) \quad (5)$$

فرض بر این خواهد بود که $h_i = g_i = l'(y_i, \hat{y}_i^{t-1})$ مقدار حدی را می‌توان با استفاده از تابع دو برابری یگانه نشان داد.

$$w_j^* = -\frac{g_i}{H_i + \lambda} \quad (6)$$

تابع هدف به مقدار حداقل خود می‌رسد.

$$Obj_{min} = -\frac{1}{2} \sum_{j=1}^T \frac{g_i}{H_i + \lambda} + \gamma T \quad (7)$$

۴.۳. تشخیص به روش گروهی طبقه‌بندها

روش یادگیری گروهی^{۳۹} پیشنهادی از تکنیک رأی‌گیری^{۴۰} دو الگوریتم یادگیری گروهی دیگر، مانند GBT و RF بهره می‌برد که به‌طور مؤثر از ترکیب مزایای هر دو الگوریتم برای بهبود عملکرد کلی مدل استفاده می‌کند. یادگیری گروهی، در ترکیب الگوریتم‌های سنتی و حتی یادگیری عمیق مؤثر نشان

³⁹ Ensemble Learning

⁴⁰ Voting

صحت عدم پاسخ به درمان که به عنوان ارزش پیش‌بینی کننده مثبت نیز شناخته می‌شود نسبت شناسایی‌های مثبت واقعی را از تمامی شناسایی‌های مثبت انجام شده به دست می‌آورد که در آن مقدار بالاتر نشانگر عملکرد بهتر است. در رابطه‌ی ۸، صحت برای طبقه‌ی عدم پاسخ به درمان نشان داده شده است.

$$P_p = \frac{TP}{TP + FP} \quad (8)$$

می‌توان مشابه صحت طبقه‌ی مثبت (P_p)، صحت برای طبقه‌ی منفی (P_N) را گزارش کرد که صحت طبقه‌ی پاسخ به درمان را نشان می‌دهد. در رابطه‌ی ۹، تاکید بر طبقه‌ی اکثریت دارد.

$$P_N = \frac{TN}{TN + FN} \quad (9)$$

جامعیت عدم پاسخ به درمان که به عنوان نرخ مثبت واقعی نیز شناخته می‌شود نسبت شناسایی‌های مثبتی که به درستی انجام شده‌اند را از تمامی نمونه‌های مثبتی که باید شناسایی شوند، اندازه‌گیری می‌کند، که در این بخش مقدار بالاتر نشانگر عملکرد بهتر است.

$$R_p = \frac{TP}{TP + FN} \quad (10)$$

می‌توان مشابه جامعیت طبقه‌ی مثبت (R_p)، جامعیت برای طبقه‌ی منفی (R_N) را گزارش کرد که جامعیت طبقه‌ی پاسخ به درمان را نشان می‌دهد. در رابطه‌ی ۱۱، تاکید بر طبقه‌ی اکثریت دارد.

$$R_N = \frac{TN}{TN + FP} \quad (11)$$

میانگین هارمونیک یک معیار ارزیابی پرکاربرد است که صحت و جامعیت را در یک مقدار واحد ترکیب می‌کند، این معیار عموماً هرم صحت و هم جامعیت را با وزن یکسانی ترکیب می‌کند. این معیار زمانی اهمیت ویژه می‌یابد که با داده‌های نامتعادل سروکار داریم، زیرا به جای تمرکز صرف بر یکی از معیارها، هر دو را در نظر می‌گیرد.

$$F - Measure(x) = \frac{2PR}{P + R} \times 100 \quad (12)$$

که در آن TP مثبت صحیح، FP مثبت کاذب و FN منفی کاذب است.

۴. بحث و نتایج

در پژوهش حاضر، یک چالش مهم مرتبط با طبقه‌بندی داده‌های نامتعادل به منظور تشخیص وضعیت پاسخ به CRT در

ویژگی‌های مورد تحلیل پژوهش را معرفی می‌کند که توسط تیم پژوهشی نژاده و همکاران [۳]، جمع‌آوری گردیده است. دسترسی به داده‌ها، بخش نوآورانه تحقیق، تغییرات دیتاست در مراحل مختلف از طریق گیت‌هاب^{۴۱} پروژه در دسترس می‌باشد. این داده‌ها از منابع مختلف از جمله پرونده‌های پزشکی، آزمایشگاهی، بالینی، نوار قلب و اکو، توسط متخصصان پزشکی جمع‌آوری شده‌اند. مجموعه‌ای به تعداد ۲۰۹ نمونه از بیماران مبتلا به نارسایی قلبی که در بیمارستان حشمت شهر رشت، به تداوم فرآیند CRT تحت پوشش قرار گرفته‌اند، در این تحقیق مورد بررسی و ارزیابی قرار گرفته‌اند. تمامی عمل‌های کاشت دستگاه انجام شده، بر اساس استفاده از دستگاه CRT-D بوده است. از اطلاعات مشخصات پزشکی و پرسشنامه‌های تکمیلی بیمار، ۶۰ ویژگی برای تحلیل فرآیند مهیا گردیده است. تقریباً ۸۲ درصد از نمونه‌ها به عنوان پاسخ مثبت به CRT و ۱۸ درصد به عنوان پاسخ منفی به CRT شناسایی شدند.

۶.۳. نحوه‌ی سنجش روش پیشنهادی

همانطور که در پژوهش جاری بصورت پیوسته تاکید شده است چالش اصلی تحقیق عدم تعادل بین طبقات و تعداد نمونه‌های پایین برای آموزش‌پذیری مطلوب است. لذا معیار صحت و جامعیت طبقه‌ی عدم پاسخ به درمان می‌تواند معیارهای پراهمیت برای پژوهش باشند. معیار میانگین هارمونیک که به عنوان امتیاز F1 نیز شناخته می‌شود میانگین هارمونی یا هماهنگی از صحت و جامعیت را نشان می‌دهد که به طور همزمان هر دو معیار را متعادل می‌کند. هرچند معمولاً از دقت برای مجموعه داده‌های متعادل استفاده می‌شود، اما دقت برای مجموعه داده‌های نامتعادل مناسب نیست، چرا که ممکن است با طبقه‌بندی اشتباه اکثر نمونه‌های طبقه اقلیت و طبقه‌بندی تمامی نمونه‌ها به عنوان طبقه اکثریت، دقت بالایی را به دست بیاورد. در مقابل، میانگین هارمونیک در مواردی که پیش‌بینی طبقه مجاور یا کنترل داده‌های نامتعادل ضروری است، مفید است. میانگین هارمونیک زمانی ارزشمند است که نیاز به متعادل‌سازی و در نظر گرفتن همزمان صحت و جامعیت باشد.

41 <https://github.com/nejadeh123/CRT>

مثبت داده‌اند، در حالی که تنها ۳۷ نمونه (معادل ۱۷/۷٪) عدم پاسخ به درمان را نشان می‌دهند. این توزیع نابرابر نمونه‌ها نشان‌دهنده چالش جدی عدم تعادل بین طبقات است که می‌تواند بر عملکرد مدل‌های تشخیصی تأثیر منفی بگذارد. بنابراین، روش پیشنهادی باید به منظور رفع این عدم تعادل، اقداماتی نظیر استفاده از تکنیک‌های SMOTE را به کار گیرد تا از کاهش دقت مدل در پیش‌بینی‌های مرتبط با طبقه "عدم پاسخ به درمان" جلوگیری کند و کارایی کلی سیستم را بهبود بخشد. در ارزیابی اولیه داده‌ها، مقایسه‌ای از نتایج طبقه‌بندی الگوریتم‌های مختلف ML در تشخیص وضعیت "پاسخ به درمان" و "عدم پاسخ به درمان" همزمان‌سازی مجدد قلبی ارائه می‌گردد. شکل ۳، عملکرد پنج الگوریتم ML بر اساس معیارهای مختلف از جمله دقت، صحت، جامعیت، و میانگین هارمونیک مقایسه شده است.



شکل ۲- نمایش تعداد نمونه‌های مرتبط با طبقه‌های پاسخ و عدم پاسخ به درمان CRT

طبقه‌بندی اولیه با استفاده از بردار ویژگی اولیه که شامل ۶۰ ویژگی است، انجام شده است. هدف از این طبقه‌بندی، ارزیابی کیفیت تشخیص وضعیت پاسخ به درمان توسط هر یک از الگوریتم‌ها و تعیین الگوریتمی است که عملکرد مطلوب‌تری ارائه می‌دهد. نتایج نشان می‌دهد که الگوریتم‌های مختلف در شناسایی طبقات "پاسخ به درمان" و "عدم پاسخ به درمان" نتایج متفاوتی ارائه می‌دهند. برای مثال، الگوریتم بیز ساده (NB)^{۴۷} با دقت ۸۱/۸۲٪ عملکرد مناسبی در تشخیص پاسخ به درمان داشته اما در تشخیص عدم پاسخ به درمان ناکام مانده است یعنی صحت عدم پاسخ به درمان آن صفر است. به همین ترتیب، NN نیز با دقت ۸۲/۳٪ توانسته عملکرد مناسبی در تشخیص پاسخ به درمان ارائه دهد اما در تشخیص عدم پاسخ

دست است. در این تحقیق، از ابزارهای نرم‌افزاری متلب، رپیدمایر و زبان برنامه‌نویسی پایتون استفاده شده است. نتایج ارزیابی شامل معیارهای دقت پاسخ به درمان CRT، صحت، جامعیت و نمره‌ی F1 طبقه‌ها بر اساس روش تحقیق گزارش خواهد شد. همچنین، معیار میانگین هارمونیک نیز برای ارزیابی و بحث در خصوص حل این نوع مجموعه داده مدنظر قرار خواهد گرفت. نتایج حاصل از این پژوهش با نتایج طبقه‌بندی انجام شده توسط نژاده و همکاران [۲۵]، بر روی مجموعه دادگان درمان CRT مورد مقایسه قرار خواهد گرفت. در عملیات پیش‌پردازش، فرآیند حذف ویژگی‌های غیرموثر از بردار اولیه را با استفاده از تعیین بهره اطلاعاتی هر ویژگی صورت می‌پذیرد.

در این مرحله از تحلیل، پنج ویژگی بالینی شامل نارسایی احتقانی قلب^{۴۲}، سوفل سیستولیک^{۴۳}، سوفل دیاستولیک^{۴۴}، درد قفسه سینه ناشی از فعالیت^{۴۵} و آنژین آستانه پایین^{۴۶} دارای بهره اطلاعاتی صفر بودند و در نتیجه قدرت تفکیک‌پذیری این ویژگی‌ها در طبقه‌بندی بیماران پاسخ‌دهنده و غیرپاسخ‌دهنده به درمان CRT بدون اثر ارزیابی شد. این بدان معناست که این ویژگی‌ها هیچ تأثیری در عملیات تشخیصی ندارند و تنها باعث افزایش زمان پردازش می‌شوند. بنابراین، پیشنهاد می‌شود که این ویژگی‌ها از بردار ویژگی حذف شوند تا کارایی سیستم بهبود یابد و زمان پردازش کاهش یابد. عملیات پیش‌پردازشی انجام شده کمک می‌کند تا تنها ویژگی‌های مؤثر در بردار ویژگی اولیه باقی بمانند، که در نتیجه سرعت تشخیص را افزایش می‌دهد. در ادامه چالش تحقیق برای توجیه مراحل ارزیابی نشان داده می‌شود. شکل ۲ تعداد نمونه‌های مربوط به دو طبقه "پاسخ به درمان" و "عدم پاسخ به درمان" در نتایج CRT را نشان می‌دهد. محور افقی طبقات مختلف (پاسخ به درمان و عدم پاسخ به درمان) را نشان می‌دهد، و محور عمودی تعداد نمونه‌ها را در هر طبقه نمایش می‌دهد. طبق داده‌های موجود، از مجموعه Heshmat_CRT، ۱۷۲ نمونه (معادل ۸۲/۳٪) به درمان پاسخ

42 CHFT

43 Systolic_Murmur

44 Diastolic_Murmur

45 Exertional_CP

46 LowTH_Ang

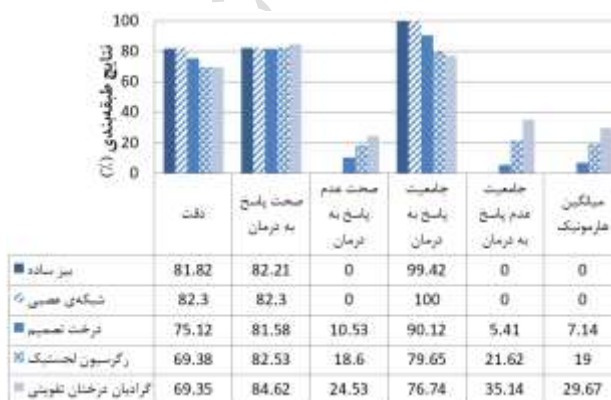
47 Naïve Bayes (NB)

سال، مقدار ۲ و برای سایر شرایط مقدار ۳ اختصاص داده می‌شود یا ویژگی Bp2، بر اساس مقدار فشار خون به سه دسته کمتر از ۹۰ (مقدار ۱)، بین ۹۰ تا ۱۴۰ (مقدار ۲)، و بیشتر از ۱۴۰ (مقدار ۳) تقسیم می‌شوند. گسسته‌سازی این ویژگی‌ها به مدل‌های ML کمک می‌کند تا با داده‌های ساده‌تر و مؤثرتر کار کنند، که در نتیجه بهبود عملکرد تشخیصی را در پی خواهد داشت.

جدول ۱- ویژگی‌های گسسته‌سازی شده جهت تقویت نتایج طبقه‌بندی

ویژگی	توضیحات
Ag2	اگر جنسیت مرد و سن کمتر یا مساوی ۴۵ سال باشد یا اگر جنسیت زن و سن کمتر یا مساوی ۵۵ سال باشد، مقدار ۲ و در غیر این صورت ۳ است.
Bp2	اگر فشار خون کمتر از ۹۰ باشد مقدار ۱، اگر کمتر یا مساوی ۱۴۰ باشد مقدار ۲، و در غیر این صورت ۳ است.
BUN2	اگر نیتروژن اوره خون کمتر از ۷ باشد مقدار ۱، اگر کمتر یا مساوی ۲۰ باشد مقدار ۲، و در غیر این صورت ۳ است.
Cr2	اگر کراتینین کمتر از ۰٫۷ باشد مقدار ۱، اگر کمتر یا مساوی ۱٫۵ باشد مقدار ۲، و در غیر این صورت ۳ است.
EF2	اگر کسر جهشی بطن چپ کمتر یا مساوی ۵۰ باشد مقدار ۱، و در غیر این صورت ۲ است.
ESR2	اگر جنسیت مرد و سرعت رسوب گلوبول کمتر یا مساوی نصف سن باشد یا اگر جنسیت زن و سرعت رسوب گلوبول کمتر یا مساوی نصف سن به علاوه ۵ باشد، مقدار ۲ و در غیر این صورت ۳ است.
FBS2	اگر قند خون ناشتا کمتر از ۷۰ باشد مقدار ۱، اگر کمتر یا مساوی ۱۰۵ باشد مقدار ۲، و در غیر این صورت ۳ است.
HB2	اگر جنسیت مرد و هموگلوبین کمتر از ۱۴ باشد یا اگر جنسیت زن و هموگلوبین کمتر از ۱۲٫۵ باشد، مقدار ۱ است؛ اگر جنسیت مرد و هموگلوبین کمتر یا مساوی ۱۷ باشد یا اگر جنسیت زن و هموگلوبین کمتر یا مساوی ۱۵ باشد، مقدار ۲ و در غیر این صورت ۳ است.
HDL2	اگر کلسترول پرچگال کمتر از ۳۵ باشد مقدار ۱، و در غیر این صورت ۲ است.
K2	اگر پتاسیم کمتر از ۳٫۸ باشد مقدار ۱، اگر کمتر یا مساوی ۵٫۶ باشد مقدار ۲، و در غیر این صورت ۳ است.
LDL2	اگر کلسترول کم‌چگال کمتر یا مساوی ۱۳۰ باشد مقدار ۲، و در غیر این صورت ۳ است.
Na2	اگر سدیم کمتر از ۱۳۶ باشد مقدار ۱، اگر کمتر یا مساوی ۱۴۶ باشد مقدار ۲، و در غیر این صورت ۳ است.
PLT2	اگر پلاکت کمتر از ۱۵۰ هزار باشد مقدار ۱، اگر کمتر یا مساوی ۴۵۰ هزار باشد مقدار ۲، و در غیر این صورت ۳ است.
PR2	اگر فاصله PR کمتر از ۶۰ باشد مقدار ۱، اگر کمتر یا مساوی ۱۰۰ باشد مقدار ۲، و در غیر این صورت ۳ است.

به درمان نتیجه مشابهی با NB دارد. در مقابل، الگوریتم GBT توانسته است به دقت ۶۹/۳۵٪ و صحت ۲۴/۵۳٪ در تشخیص عدم پاسخ به درمان دست یابد که بالاتر از سایر الگوریتم‌ها است. این الگوریتم همچنین دارای میانگین هارمونیک ۲۹/۶۷٪ است که بالاترین مقدار میان سایر الگوریتم‌ها است و نشان‌دهنده تعادل نسبی بین صحت و جامعیت است. نتایج حاصل شده به وضوح نشان می‌دهند که هر الگوریتم مزایا و معایب خاص خود را دارد اما انتخاب الگوریتم مناسب باید بر اساس تشخیص بهتر طبقه‌بندی اقلیت صورت گیرد. از اینرو الگوریتم GBT نسبت به سایر الگوریتم‌ها عملکرد بهتری را ارائه می‌کند هر چند دقت کلی تشخیص آن پایین است.



شکل ۳- مقایسه‌ی نتایج طبقه‌بندی الگوریتم‌های رایج ML در تشخیص وضعیت پاسخ به درمان CRT

در ادامه لیستی از ویژگی‌های گسسته‌سازی شده ارائه می‌گردد که برای تقویت نتایج طبقه‌بندی در سیستم تشخیص درمان CRT مورد استفاده قرار می‌گیرند. گسسته‌سازی ویژگی‌ها فرآیندی است که طی آن مقادیر پیوسته ویژگی‌ها به دسته‌های گسسته تبدیل می‌شوند تا پیچیدگی مدل کاهش یابد و عملکرد آن بهبود یابد. در جدول ۱، ۱۷ ویژگی گسسته‌سازی شده به همراه شرایط تعریف شده برای هر کدام ارائه شده است. هر ویژگی بر اساس شرایط مختلفی از جمله جنسیت، سن، و مقادیر عددی سایر پارامترها، به دسته‌های مختلف تقسیم شده است. این فرآیند گسسته‌سازی بر اساس روش پژوهشی نژاده و همکاران انجام شده و منجر به استخراج ویژگی‌های بهینه برای بهبود عملکرد مدل‌های طبقه‌بندی می‌شود. برای مثال ویژگی Ag2، بر اساس جنسیت و سن تعریف شده است. برای مردان با سن کمتر یا مساوی ۴۵ سال و زنان با سن کمتر یا مساوی ۵۵

RWMA2	اگر ناهنجاری حرکت دیواره ناحیه‌ای ۰ باشد مقدار ۲، و در غیر این صورت ۳ است.
TG2	اگر تری گلیسرید کمتر یا مساوی ۲۰۰ باشد مقدار ۲، و در غیر این صورت ۳ است.
WBC2	اگر تعداد گلبول‌های سفید کمتر از ۴۰۰۰ باشد مقدار ۱، اگر کمتر یا مساوی ۱۱۰۰۰ باشد مقدار ۲، و در غیر این صورت ۳ است.

استفاده از این ویژگی‌های گسسته‌سازی شده می‌تواند دقت، صحت و سایر معیارهای مرتبط با طبقه‌بندی را افزایش دهد و در نتیجه بهبود چشمگیری در نتایج حاصل از مدل‌های ML مشاهده شود. با اضافه نمودن هفده ویژگی استخراج شده از متغیرهای پیوسته، بردار ویژگی شامل ۷۷ متغیر جهت طبقه‌بندی می‌شود.

همانطور که در بخش ۳-۱ مقاله ذکر شد، پیش‌پردازش داده‌ها یکی از مراحل مهم در توسعه مدل‌های یادگیری ماشین است و تأثیر مستقیمی بر کیفیت و دقت نتایج دارد. در این پژوهش، مجموعه داده Heshmat_CRT در ابتدا شامل ۲۰۹ نمونه بیمار با ۶۰ ویژگی بود و در مراحل مختلف پیش‌پردازش به ۲۰۶ نمونه با ۷۷ ویژگی تبدیل شد. در ادامه، تمام مراحل عملیاتی که با جزئیات کامل انجام شد، توضیح داده می‌شود.

بررسی و حذف داده‌های نویزی: در مرحله نخست، مجموعه داده از نظر مقادیر گمشده و داده‌های نویزی بررسی گردید. در حالت ایده‌آل، هیچ مقدار گمشده‌ای وجود نداشت زیرا داده‌ها به طور دقیق توسط پزشکان بیمارستان حشمت رشت جمع‌آوری شده بودند. با این وجود، تجزیه و تحلیل آماری نشان داد که سه نمونه در ویژگی‌های آزمایشگاهی خود مقادیر پرت^{۴۸} داشتند که بیش از حد انحراف معیار از میانگین فاصله داشتند. این نمونه‌ها، پس از مشورت با پزشکان و تأیید خطا در ثبت یا اندازه‌گیری، از مجموعه داده‌ها حذف شدند و بنابراین تعداد نهایی نمونه‌ها به ۲۰۶ کاهش یافت.

حذف ویژگی‌های غیرموثر با استفاده از بهره اطلاعاتی: برای شناسایی و حذف ویژگی‌های بدون قدرت تشخیصی، بهره اطلاعاتی تمام ۶۰ ویژگی اولیه نسبت به برچسب پاسخ به درمان CRT محاسبه گردید. بهره اطلاعاتی معیاری است که میزان

کاهش آنتروپی یا عدم قطعیت در طبقه‌بندی پس از استفاده از یک ویژگی خاص را نشان می‌دهد و با فرمول $IG(T,a) = H(T) - H(T|a)$ محاسبه می‌شود که در آن $H(T)$ آنتروپی مجموعه داده اصلی و $H(T|a)$ آنتروپی شرطی پس از تقسیم بر اساس ویژگی a است. نتایج نشان داد که پنج ویژگی بالینی شامل نارسایی احتقانی قلب (CHFT) با $IG=0.000$ ، سופل سیستولیک با $IG=0.000$ ، سوفل دیاستولیک با $IG=0.000$ ، درد قفسه سینه ناشی از فعالیت با $IG=0.000$ و آنژین آستانه پایین با $IG=0.001$ دارای بهره اطلاعاتی بسیار پایین (کمتر یا مساوی ۰,۰۰۱) هستند و عملاً هیچ توانایی تمایزی بین بیماران پاسخ‌دهنده و غیرپاسخ‌دهنده به درمان ندارند. این ویژگی‌ها از بردار ویژگی حذف شدند و تعداد ویژگی‌ها به ۵۵ مورد کاهش یافت.

گسسته‌سازی ویژگی‌های پیوسته بر اساس معیارهای بالینی: پس از حذف ویژگی‌های غیرموثر، ۱۷ ویژگی پیوسته شناسایی شدند که بر اساس دستورالعمل‌های بالینی استاندارد دارای محدوده‌های تشخیصی مشخص هستند. این ویژگی‌ها با استفاده از روش گسسته‌سازی با عرض یکسان (Equal-Width Binning) به سه دسته تبدیل شدند. تعداد سه دسته بر اساس فرمول Sturges یعنی $k = 1 + \log_2(n)$ که برای $n=206$ مقدار $k \approx 8.68$ را می‌دهد انتخاب شد، اما به دلیل محدوده‌های بالینی مشخص و برای حفظ تفسیرپذیری پزشکی، از سه دسته استفاده گردید که با نظر دو متخصص قلب و عروق تأیید شد. به‌عنوان مثال، ویژگی فشار خون (BP) به سه دسته کمتر از ۹۰ (هیپوتانسیون)، بین ۹۰ تا ۱۴۰ (نرمال) و بیشتر از ۱۴۰ (هیپرتانسیون) تقسیم شد که کاملاً مطابق با راهنماهای بالینی ACC/AHA 2021 برای نارسایی قلبی است. همچنین ویژگی کسر جهشی بطن چپ (EF) به دو دسته کمتر یا مساوی ۵۰ درصد (نارسایی قلبی با کسر جهشی کاهش یافته) و بیشتر از ۵۰ درصد (عملکرد نسبتاً نرمال) تقسیم شد. ویژگی نیتروژن اوره خون (BUN) به سه دسته کمتر از ۷، بین ۷ تا ۲۰ (نرمال) و بیشتر از ۲۰ (نارسایی کلیوی احتمالی) گسسته‌سازی شد. جدول ۱ تمام ۱۷ ویژگی گسسته‌سازی شده و محدوده‌های انتخابی را که بر اساس منابع معتبر پزشکی تعیین شده‌اند نشان می‌دهد. این فرآیند گسسته‌سازی نه‌تنها به کاهش پیچیدگی محاسباتی کمک می‌کند بلکه تفسیرپذیری پزشکی

مدل را نیز افزایش می‌دهد و همان‌طور که مطالعات نشان داده‌اند می‌تواند کیفیت طبقه‌بندی را به‌طور قابل توجهی بهبود بخشد.

نرمال‌سازی ویژگی‌های پیوسته باقیمانده: پس از گسسته‌سازی ۱۷ ویژگی، ۳۸ ویژگی پیوسته دیگر باقی ماندند که شامل پارامترهای آزمایشگاهی مانند کراتینین، هموگلوبین، پلاکت، الکترولیت‌ها و پارامترهای الکتروکاردیوگرافی بودند. معیار تصمیم‌گیری برای انتخاب بین گسسته‌سازی و نرمال‌سازی این بود که ویژگی‌هایی که دارای محدوده‌های بالینی استاندارد و آستانه‌های تشخیصی مشخص بودند گسسته‌سازی شدند، در حالی که ویژگی‌هایی که دارای طیف پیوسته از مقادیر بودند و محدوده استاندارد واحدی نداشتند یا اطلاعات کمی دقیق‌تر در آن‌ها اهمیت داشت نرمال‌سازی شدند تا از دست رفتن اطلاعات جلوگیری شود. این ۳۸ ویژگی با استفاده از روش استانداردسازی امتیاز Z ^{۴۹} نرمال‌سازی شدند که در آن هر ویژگی به فرم $Z = (x - \mu) / \sigma$ تبدیل می‌شود، به‌طوری که μ میانگین و σ انحراف معیار ویژگی در مجموعه داده آموزشی است. این تبدیل تضمین می‌کند که تمام ویژگی‌های پیوسته دارای میانگین صفر و انحراف معیار یک هستند و در نتیجه در مقیاس یکسانی قرار می‌گیرند، بنابراین هیچ ویژگی به دلیل داشتن محدوده عددی بزرگ‌تر به‌طور نامتناسب بر فرآیند یادگیری مدل تأثیر نخواهد گذاشت.

کدگذاری ویژگی‌های دسته‌ای: مجموعه داده شامل ۷ ویژگی دسته‌ای غیررتبه‌ای بود که شامل جنسیت، نوع دیابت، نوع قلبی، وضعیت سیگار، بیماری راه هوایی، سگته مغزی و نوع دستگاه CRT بودند. این ویژگی‌ها با استفاده از روش کدگذاری وان-هات^{۵۰} به ویژگی‌های باینری تبدیل شدند. در این روش، هر ویژگی دسته‌ای با k دسته به k ویژگی باینری جدید تبدیل می‌شود که در هر نمونه تنها یکی از این ویژگی‌های باینری مقدار ۱ و بقیه مقدار ۰ دارند. این ۷ ویژگی دسته‌ای در مجموع به ۱۹ ویژگی باینری تبدیل شدند. علاوه بر این، ۱۰ ویژگی دسته‌ای ترتیبی نیز در مجموعه داده وجود داشت که شامل طبقه

عملکردی قلب (NYHA Class) و شدت برخی علائم بود که به‌صورت عددی کدگذاری شدند، به‌طوری که ترتیب طبیعی دسته‌ها در آن‌ها حفظ شود.

خلاصه فرآیند پیش‌پردازش: با انجام تمام مراحل فوق، بردار ویژگی از ۶۰ ویژگی اولیه به‌صورت زیر تبدیل شد: ابتدا ۵ ویژگی غیرموثر حذف و به ۵۵ ویژگی رسیدیم، سپس ۱۷ ویژگی پیوسته گسسته‌سازی شدند، ۳۸ ویژگی نرمال‌سازی شدند، ۷ ویژگی دسته‌ای غیررتبه‌ای حذف و به‌جای آن‌ها ۱۹ ویژگی باینری اضافه شد، و ۱۰ ویژگی دسته‌ای ترتیبی کدگذاری عددی شدند. در نتیجه، بردار ویژگی نهایی قبل از اعمال روش CSA-SMOTE شامل ۷۷ ویژگی شد که عبارتند از: ۱۷ ویژگی گسسته‌شده + ۳۸ ویژگی نرمال‌شده + ۱۹ ویژگی باینری حاصل از کدگذاری وان-هات و ۷ ویژگی دسته‌ای اصلی که به جای ۱۰ ویژگی ترتیبی کدگذاری شده جایگزین شدند، بصورت مجموع ۷۷ ویژگی حاصل شد. این فرآیند پیش‌پردازش جامع زمینه را برای مراحل بعدی شامل متعادل‌سازی کلاس‌ها با روش CSA-SMOTE و انتخاب ویژگی با الگوریتم ژنتیک فراهم کرد که منجر به آموزش مدل‌های یادگیری گروهی با دقت و قابلیت اعتماد بالا باشد.

جهت کاهش چالش تحقیق، متعادل‌سازی طبقات با روش SMOTE بهینه توصیه شده است. هدف از این عملیات بهبود وضعیت عدم‌تعادل طبقات از طریق ایجاد و بهینه‌سازی نمونه‌های مصنوعی برای کلاس اقلیت است. اگرچه عملیات گسسته‌سازی بردار ویژگی به بهبود تشخیص کمک کرده است، اما همچنان چالش عدم‌تعادل طبقات باقی مانده است. برای رفع این مشکل، روش پیشنهادی از تکنیک SMOTE استفاده می‌کند. در این روش، ابتدا نمونه‌های مصنوعی از کلاس اقلیت تولید می‌شوند و سپس این نمونه‌ها با استفاده از الگوریتم CSA بهینه‌سازی می‌شوند.

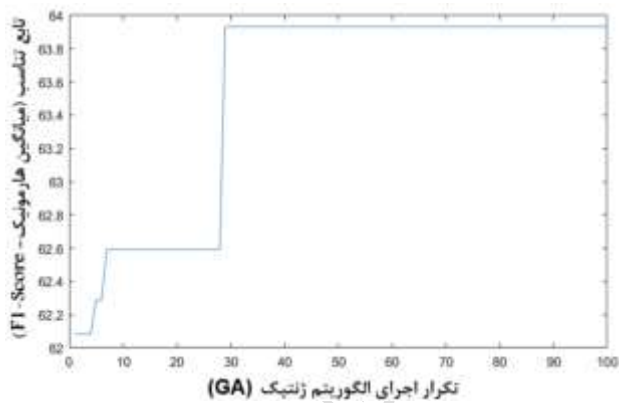
جدول ۲، پارامترهای بکارگرفته شده برای الگوریتم CSA نشان می‌دهد. اندازه جمعیت کلاغ (N) به صورت پویا و برابر با تعداد نمونه‌های مصنوعی تولید شده توسط SMOTE تعیین شد. با توجه به عدم تعادل اولیه داده‌ها (۳۷ نمونه اقلیت در مقابل ۱۷۲ نمونه اکثریت)، الگوریتم SMOTE ابتدا ۱۳۵ نمونه مصنوعی

⁴⁹ Z-score Standardization

⁵⁰ One-Hot Encoding

می‌شود (در این مرحله، از الگوریتم GBT با پارامترهای پیش‌فرض استفاده شده است) و میانگین هارمونیک حاصل به عنوان شاخص کیفیت آن نمونه محاسبه می‌شود. این رویکرد تضمین می‌کند که نمونه‌های مصنوعی نه تنها از نظر آماری مشابه کلاس اقلیت هستند، بلکه در بهبود قدرت تشخیصی مدل نیز مؤثر هستند.

تابع تناسب، که بهبود عملکرد مدل را بر اساس معیار میانگین هارمونیک ارزیابی می‌کند، در هر نسل به‌روزرسانی می‌شود. شکل ۴ نشان می‌دهد که میانگین هارمونیک، که معیاری برای ارزیابی کارایی نمونه‌های تولید شده است، در طول ۱۰۰ نسل به تدریج افزایش یافته و بهبود یافته است. این نمودار نشان می‌دهد که تا پیش از نسل ۳۰، فرآیند بهینه‌سازی به یک وضعیت نسبی بهینگی رسیده است، که نشان‌دهنده تثبیت کیفیت نمونه‌های مصنوعی تولید شده است.



شکل ۴- نمودار نسلی نتایج تابع تناسب برای پیش‌نمونه‌برداری بهینه از کلاس اقلیت

به عبارت دیگر، پس از نسل ۳۰، تغییرات کمتری در میانگین هارمونیک مشاهده می‌شود، که حاکی از رسیدن به یک نقطه تعادل در فرآیند بهینه‌سازی است. این نمودار بیانگر موفقیت روش پیشنهادی در بهبود تناسب نمونه‌های کلاس اقلیت است و به کاهش عدم‌تعادل بین طبقات کمک می‌کند.

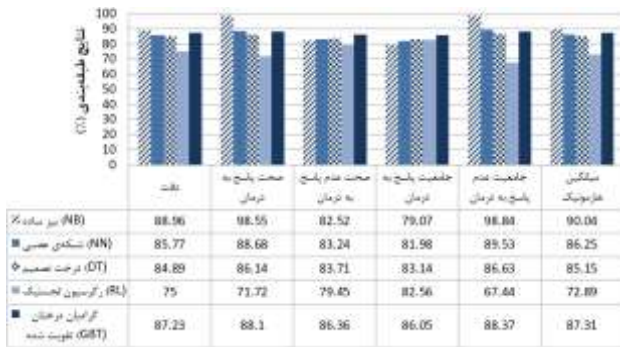
مقایسه اطلاعات مرتبط با نمونه‌ها قبل و بعد از اجرای فرآیند پیش‌نمونه‌برداری اقلیت تعمیم‌یافته (پیش‌نمونه‌برداری بهینه) در جدول ۳ آورده شده است. در حالت اولیه، تعداد نمونه‌های طبقه "عدم پاسخ به درمان" ۳۷ مورد و تعداد نمونه‌های طبقه "پاسخ به درمان" ۱۷۲ مورد است، که نشان‌دهنده عدم تعادل قابل توجهی بین طبقات است و باعث می‌شود مدل‌های ML تمایل بیشتری

ایجاد کرد، بنابراین $N=135$ در نظر گرفته شد. حداکثر تعداد تکرارها (tmax) روی ۱۰۰۰ تنظیم شد، با این حال، برای جلوگیری از هدر رفتن منابع محاسباتی، یک مکانیسم توقف برای خاتمه زودهنگام پیاده‌سازی شد که در صورت عدم بهبود میانگین هارمونیک برای ۱۰۰ نسل متوالی ($gwi=100$) منجر به توقف فرآیند می‌شود. در آزمایش‌های انجام شده، الگوریتم معمولاً پس از ۲۰۰ تا ۳۰۰ تکرار همگرا می‌شود. احتمال آگاهی 0.1 (AP) بود که به خوبی بین بهره‌برداری از راه‌حل‌های موجود خوب و کاوش در فضای جستجو تعادل برقرار می‌کرد. طول پرواز (fl) نیز برابر با ۱ در نظر گرفته شد تا حرکت کلاخ‌ها در فضای جستجو با سرعت متوسطی انجام شود. برای اطمینان از اعتبار نمونه‌های تولید شده، هر موقعیت کلاخ جدید در محدوده $[minX_1, maxX_1]$ که نشان‌دهنده حداقل و حداکثر هر ویژگی در داده‌های اصلی است، محدود شد. علاوه بر این، برای ویژگی‌های گسسته عملیات گرد کردن مقطعی اعمال شد تا مقادیر تولید شده با ماهیت ویژگی‌های گسسته مطابقت داشته باشند.

جدول ۲- پارامترهای اجرایی الگوریتم CSA برای بهینه‌سازی SMOTE

پارامتر	مقدار	توضیحات
اندازه جمعیت کلاخ‌ها (N)	متغیر (برابر با تعداد نمونه‌های مصنوعی تولید شده)	تعداد کلاخ‌ها برابر با تعداد نمونه‌های جدید تولید شده توسط SMOTE است
حداکثر تعداد تکرار (tmax)	1000	حداکثر تعداد نسل‌های بهینه‌سازی
احتمال آگاهی (AP)	0.1	احتمال جهش تصادفی به جای پیروی از بهترین موقعیت
طول پرواز (fl)	1	ضریب گام حرکت در فضای جستجو
شرط توقف زودهنگام	فعال	توقف در صورت عدم بهبود به مدت ۱۰۰ نسل متوالی
نسل‌های بدون بهبود (gwi)	100	تعداد نسل‌های مجاز بدون بهبود قبل از توقف
محدوده جستجو	$[minX_1, maxX_1]$	محدود به بازه مقادیر ویژگی‌های داده اصلی

تابع تناسب (هدف) در الگوریتم CSA برای ارزیابی کیفیت هر نمونه مصنوعی بر اساس میانگین هارمونیک طبقه‌بندی استفاده شد. در هر تکرار، مجموعه داده شامل داده‌های اصلی و نمونه‌های جدید توسط یک طبقه‌بندی‌کننده پایه آموزش داده



شکل ۵- مقایسه‌ی نتایج طبقه‌بندی الگوریتم‌های رایج ML پس از مرحله‌ی بیش‌نمونه‌برداری بهینه

همچنین مقایسه‌ای از نتایج طبقه‌بندی الگوریتم GBT را قبل و بعد از اجرای فرآیند بیش‌نمونه‌برداری بهینه در شکل ۶ ارائه شده که هدف این مقایسه، ارزیابی میزان بهبودی است که پس از اعمال بیش‌نمونه‌برداری بهینه روی بردار ویژگی اولیه به دست آمده است. در مرحله اولیه، الگوریتم GBT با بردار ویژگی اولیه عملکرد نسبتاً ضعیفی در شناسایی طبقه "عدم پاسخ به درمان" داشته است. دقت کلی این الگوریتم ۶۹,۳۵ درصد بوده و میانگین هارمونیک تنها ۲۹,۶۷ درصد را نشان می‌دهد. به طور خاص، صحت شناسایی طبقه "عدم پاسخ به درمان" ۲۴,۵۳ درصد بوده است که بیانگر توانایی پایین الگوریتم در شناسایی دقیق این طبقه است. پس از اجرای فرآیند بیش‌نمونه‌برداری بهینه، نتایج به‌طور قابل توجهی بهبود یافته است. دقت کلی الگوریتم به ۸۷,۲۳ درصد افزایش یافته و میانگین هارمونیک به ۸۷,۳۱ درصد رسیده است. همچنین، صحت شناسایی طبقه "عدم پاسخ به درمان" به ۸۶,۳۶ درصد ارتقا یافته است. این نتایج نشان‌دهنده تأثیر مثبت و قابل توجه فرآیند بیش‌نمونه‌برداری بهینه به‌ویژه در شناسایی طبقه اقلیت در بهبود عملکرد الگوریتم GBT است. در مجموع اعمال بیش‌نمونه‌برداری بهینه نه تنها توانسته است دقت کلی مدل را افزایش دهد، بلکه تعادل بهتری را بین شناسایی طبقات "پاسخ به درمان" و "عدم پاسخ به درمان" ایجاد کرده است، که نتیجه‌ی آن بهبود چشمگیر در کارایی طبقه‌بندی است.

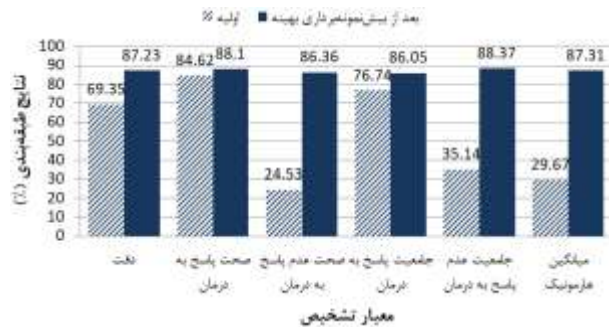
به تشخیص طبقه اکثریت (پاسخ به درمان) داشته باشند. پس از اجرای فرآیند بیش‌نمونه‌برداری بهینه، تعداد نمونه‌های طبقه "عدم پاسخ به درمان" به ۱۷۲ افزایش یافته و برابر با تعداد نمونه‌های طبقه "پاسخ به درمان" شده است. این افزایش تعداد نمونه‌ها از طریق تولید و بهینه‌سازی نمونه‌های مصنوعی با استفاده از روش‌های پیشرفته انجام شده و نتیجه‌ی آن، یک مجموعه داده متعادل‌تر است که می‌تواند منجر به بهبود عملکرد مدل‌های ML در تشخیص هر دو طبقه شود. مقایسه‌ی نتایج طبقه‌بندی الگوریتم‌های مختلف ML را پس از اجرای مرحله‌ی بیش‌نمونه‌برداری بهینه در شکل ۵ نشان داده شده است. در این مرحله، بهینه‌سازی از طریق ادغام دو روش SMOTE و الگوریتم CSA انجام گرفته که نقش مهمی در بهبود عملکرد کلی الگوریتم‌های مورد بررسی دارد.

جدول ۳- مقایسه‌ی نمونه‌ها قبل و بعد از بیش‌نمونه‌برداری بهینه

نمونه	فراوانی طبقه عدم‌پاسخ به درمان	فراوانی طبقه پاسخ به درمان	بردار ویژگی اولیه
209	37	172	اولیه
344	172	172	بعد از بیش‌نمونه‌برداری بهینه

هدف این آزمایش، نمایش تأثیر بهبودهای حاصل از این فرآیند بهینه‌سازی بر عملکرد الگوریتم‌های ML است. نتایج نشان می‌دهد که الگوریتم NB و GBT بهترین عملکرد را از خود نشان داده‌اند. الگوریتم NB، با پارامترهای مناسب، به دقت ۸۸/۹۶٪ و میانگین هارمونیک ۹۰/۰۴٪ دست یافته است، که نشان‌دهنده توانایی بالای آن در تشخیص دقیق نمونه‌های مربوط به هر دو طبقه است. در مقابل، الگوریتم GBT نیز با دقت ۸۷/۲۳٪ و میانگین هارمونیک ۸۷/۳۱٪، پس از NB توانسته نتایج مطلوبی را گزارش نماید و همچنین اثبات می‌گردد که الگوریتم GBT، با وجود پارامترهای متنوع، عملکردی قوی و پایدار ارائه دهد. به طور کلی، این شکل بیانگر آن است که بهینه‌سازی انجام شده به کمک روش‌های پیشرفته، منجر به افزایش دقت و بهبود عملکرد الگوریتم‌های مختلف ML شده است، به ویژه در تشخیص نمونه‌های اقلیت که چالش اصلی در این تحقیق بوده است.

مقدار پارامتر	پارامتر طبقه‌بند
9	حداکثر تعداد نخ برای اجرا
10	حداکثر عمق درخت
10	حداقل سطرهای مرتبط به برچسب
20	تعداد سطل‌بندی برای گسسته‌سازی
0.01	نرخ یادگیری



شکل ۶- مقایسه‌ی نتایج طبقه‌بندی GBT برای قبل و بعد از

بیش‌نمونه‌برداری بهینه

به منظور دستیابی به نتایج دقیق‌تر در فرآیند طبقه‌بندی، پارامترهای الگوریتم GBT تنظیم شده‌اند تا عملکرد بهتری از الگوریتم حاصل شود. جدول ۴ تنظیمات به‌روزرسانی شده پارامترهای الگوریتم GBT را برای بهبود بهره‌وری در تشخیص درمان CRT نشان می‌دهد. تنظیمات شامل افزایش تعداد تخمین‌زننده‌ها (درخت‌ها) به ۶۵، که تعداد مدل‌های ترکیب‌شده در الگوریتم را مشخص می‌کند و بهبود قدرت پیش‌بینی را به همراه دارد. همچنین، حداکثر تعداد نخ برای اجرا به ۹ تنظیم شده است، که نشان‌دهنده تعداد فرآیندهای موازی برای تسریع عملیات محاسباتی است. حداکثر عمق درخت به ۱۰ محدود شده است، که به کنترل پیچیدگی مدل و جلوگیری از بیش‌برازش کمک می‌کند. پارامتر حداقل سطرهای منتسب به برچسب به ۱۰ تنظیم شده که به تعیین حداقل تعداد نمونه‌ها در یک گره کمک می‌کند و از ایجاد گره‌های بسیار کوچک جلوگیری می‌کند. علاوه بر این، تعداد سطل‌بندی برای گسسته‌سازی به ۲۰ تنظیم شده است تا گسسته‌سازی ویژگی‌ها با دقت بیشتری انجام شود. نرخ یادگیری نیز به ۰/۰۱ تعیین شده که بیانگر سرعت آموزش مدل است و تأثیر هر درخت جدید در به‌روزرسانی مدل نهایی را کنترل می‌کند. این تنظیمات به‌روزرسانی شده به طور جامع به بهینه‌سازی الگوریتم کمک کرده و به بهبود دقت و کارایی در تشخیص وضعیت پاسخ به درمان CRT منجر شده‌اند.

جدول ۴- تنظیم پارامترهای الگوریتم GBT برای بهره‌وری تشخیص

درمان CRT

مقدار پارامتر	پارامتر طبقه‌بند
65	تعداد تخمین‌زننده (درخت)

می‌توان مقایسه‌ی نتایج طبقه‌بندی الگوریتم GBT را قبل و بعد از بهینه‌سازی پارامترهای آن نشان داد، که تأثیر قابل توجه تنظیم دقیق پارامترها بر عملکرد مدل در تشخیص پاسخ به درمان CRT را گزارش می‌کند. بر اساس شکل ۷، در مرحله‌ی اولیه، پس از اجرای فرآیند بیش‌نمونه‌برداری بهینه، مدل به دقت ۸۷/۲۳٪ دست یافته بود، با صحت ۸۸/۱٪ برای طبقه‌ی "پاسخ به درمان" و ۸۶/۳۶٪ برای طبقه‌ی "عدم پاسخ به درمان". جامعیت مدل نیز به ترتیب برای این دو طبقه ۸۶/۰۵٪ و ۸۸/۳۷٪ بود که منجر به میانگین هارمونیک ۸۷/۳۱٪ می‌شد. با انجام بهینه‌سازی پارامتری، عملکرد مدل به طور محسوسی بهبود یافت؛ دقت کلی به ۸۹/۸۳٪ افزایش پیدا کرد که نشان‌دهنده‌ی ارتقای قابل توجه در توانایی مدل برای تشخیص صحیح نمونه‌ها است. به‌طور خاص، صحت طبقه‌ی "پاسخ به درمان" به ۹۶/۶٪ رسید، که بیانگر افزایش دقت در شناسایی بیماران پاسخ‌دهنده به درمان است، در حالی که صحت طبقه‌ی "عدم پاسخ به درمان" به ۸۴/۷۷٪ رسید، که نشان می‌دهد مدل همچنان توانایی بالایی در تشخیص صحیح این گروه دارد. جامعیت برای طبقه‌ی "پاسخ به درمان" ۸۲/۵۶٪ و برای طبقه‌ی "عدم پاسخ به درمان" ۹۷/۰۹٪ بود، که تعادل بهتری بین نرخ تشخیص درست و نرخ خطا را فراهم می‌کند. این بهبودها در نهایت منجر به افزایش میانگین هارمونیک به ۹۰/۵۷٪ شد، که نمایانگر کارایی و تعادل بهتر مدل در تشخیص هر دو طبقه است. این نتایج تأیید می‌کنند که بهینه‌سازی پارامترهای الگوریتم GBT می‌تواند به شکل قابل توجهی دقت و قابلیت اعتماد مدل را در تشخیص وضعیت پاسخ به درمان CRT ارتقا دهد، و در نتیجه به تصمیم‌گیری‌های بالینی مؤثرتر و دقیق‌تری منجر شود.

ویژگی	ویژگی
Airway_disease, BUN, BUN2, CABG, CR, Cr2, CRF, CVA, Device_Type, DM, ESR, EX-Smoker, FBS2, FH, Function_Class, HB, HR, HR2, LDL, LDL2, Lung_rales, LVH, Na2, Neut, Nonanginal, PLT2, Poor_R_Progression, PR_Interval2, QRS_Duration, QT_Interval, QT_Interval2, Rhythm, Sex, ST_Depression, ST_Elevation, TG, TG2, Thyroid_Disease, Total_chol2, Troponin, WBC, Weak_Peripheral_Pulse	انتخاب شده 41
Age, Age2, AI, Atypical, BP, BP2, Current_Smoker, DLP, Dyspnea, Edema, ESR2, FBS, Fragment, HB2, HDL, HDL2, HTN, K, K2, LV_EF, LV_EF2, Lymph, Lymph2, MR, Na, Neut2, Number_of_fragment, PLT, PR_Interval, Q_Wave, QRS_Duration2, T_inversion, Total_chol, Typical_Chest_Pain, WBC2	حذف شده 36

روش پیشنهادی یادگیری گروه را با بکارگیری الگوریتم‌های GBT و RF برای دستیابی به نتایج مطلوب طبقه‌بندی توصیه می‌کند. برای دستیابی عملکرد بهینه، پارامترهای الگوریتم‌های ML (RF و GBT) با دقت تنظیم شدند. جدول ۶ مقادیر پارامترهای اصلی این الگوریتم‌ها را در شرایط بهینه آنها نشان می‌دهد که بر اساس آزمایش‌ها و ارزیابی‌های متعدد با معیار میانگین هارمونیک به دست آمده‌اند. این تنظیمات به طور خاص برای مجموعه داده CRT بهینه شده‌اند و منجر به بهبود قابل توجه معیارهای عملکرد شده‌اند. در روش یادگیری گروهی، از استراتژی رأی‌گیری نرم^{۵۱} برای ترکیب نتایج الگوریتم‌های GBT و RF با وزن‌های برابر استفاده شد.

جدول ۶- پارامترهای مدل طبقه‌بندی بهینه شده برای شناسایی درمان CRT

پارامتر	روش گروهی	الگوریتم RF	الگوریتم GBT
تعداد تخمین‌زننده/درخت	-	100	65
حداکثر عمق درخت	-	15	10
حداقل نمونه در گره	-	5	10
نرخ یادگیری	-	-	0.01
تعداد سطل‌بندی	-	-	20
معیار تقسیم	-	Gain_ratio	-
استراتژی رأی‌گیری	Soft Voting	Confidence vote	-
حداکثر تعداد نخ	-	-	9



شکل ۷- مقایسه‌ی نتایج طبقه‌بندی قبل و بعد از بهینه‌سازی پارامترهای

الگوریتم GBT

مرحله‌ی بعد، فرآیند انتخاب ویژگی برای استخراج بردار بهبودیافته است. جدول ۵ نتایج فرآیند انتخاب ویژگی را نشان می‌دهد که پس از اعمال GA بر روی مجموعه داده‌های مورد بررسی، ویژگی‌های مؤثر برای تشخیص بهینه درمان CRT (همزمان‌سازی مجدد قلبی) را مشخص می‌کند. در این فرآیند، GA با استفاده از پارامترهایی نظیر سائز جمعیت برابر با ۲۵، تعداد نسل‌های ۱۰۰ و حداکثر ۱۰ نسل بدون بهبود به‌عنوان شرایط خروج از الگوریتم، به انتخاب ویژگی‌های بهینه پرداخته است. ترکیب ۹۰٪ و جهش ۱۰٪ نیز به‌منظور تولید نسل‌های جدید و کاوش در فضای ویژگی‌ها استفاده شده‌اند. تابع هدف این الگوریتم، میانگین هارمونیک طبقه‌بندی بوده است که با استفاده از الگوریتم GBT ارزیابی شده است. نتایج نشان می‌دهند که از میان ویژگی‌های موجود، ۴۱ ویژگی به‌عنوان ویژگی‌های انتخاب شده حفظ گردیده‌اند. برخی این ویژگی‌ها شامل Airway_disease, BUN, CABG, CR, CRF, CVA و مقابل، ۳۶ ویژگی دیگر که شامل Age2, ESR2, HB2, HDL2 و WBC2 می‌باشند، به دلیل عدم تأثیرگذاری قابل توجه در بهبود معیار میانگین هارمونیک مدل، حذف گردیده‌اند. این فرآیند انتخاب و حذف ویژگی‌ها منجر به بهینه‌سازی مدل و کاهش پیچیدگی محاسباتی آن شده است، که در نهایت به افزایش میانگین هارمونیک و کارایی مدل در تشخیص وضعیت پاسخ به درمان CRT کمک می‌کند. در بخش پایانی، طبقه‌بندی برای تشخیص CRT، عملکرد پیش‌بینی روش پیشنهادی را نشان می‌دهد.

جدول ۵- ویژگی‌های گزینش و حذف شده پس از انتخاب ویژگی

وضعیت	تعداد	نام ویژگی
-------	-------	-----------

⁵¹ Soft Voting

						بردار اولیه
87.31	88.37	86.05	86.36	88.1	87.2 3	GBT بعد از بیش نمونه برداری بهینه
90.57	97.09	82.56	84.77	96.6	89.8 3	GBT بعد از بهینه سازی پارامتری
87.99	85.47	91.28	90.74	86.2 6	88.3 8	GBT پس از انتخاب ویژگی
88.05	80.23	98.26	97.87	83.2 5	89.2 5	RF پس از انتخاب ویژگی
89.07	84.88	94.19	93.59	86.1 7	89.5 5	یادگیری گروهی درختی پس از انتخاب ویژگی

در ارزیابی‌ها، برای اطمینان از قابلیت اطمینان روش و تعمیم‌پذیری یافته‌ها، طیف وسیعی از آزمون‌های آماری و روش‌های اعتبارسنجی به کار گرفته شد.

- اعتبارسنجی متقابل کافولد (KCV)⁵²: برای سنجش عملکرد مدل و جلوگیری از بیش‌برازش، از روش اعتبارسنجی متقابل ۱۰-فولد (بخشی) استفاده شد. در این روش، مجموعه داده‌ها به ۱۰ زیرمجموعه تقریباً مساوی تقسیم شدند و در هر مرحله، ۹ زیرمجموعه برای آموزش و ۱ زیرمجموعه برای آزمایش مدل استفاده شد. در ۱۰ باری که این فرآیند تکرار شد، همه زیرمجموعه‌ها به طور متناوب به عنوان مجموعه آزمون به کار گرفته شدند. میانگین و انحراف معیار معیارهای ارزیابی در تکرارها برای اندازه‌گیری پایداری و قابلیت اطمینان مدل محاسبه شد.

- آزمون T زوجی برای مقایسه روش‌ها: برای بررسی معناداری آماری تفاوت‌های بین عملکرد روش پیشنهادی و حالت پایه، از روش آزمون t زوجی⁵³ استفاده شد. این آزمون در سطح اطمینان ۹۵٪ ($\alpha=0.05$) انجام شد تا نشان دهد که آیا بهبود در معیارهای ارزیابی از نظر آماری معنادار بوده است یا خیر. نتایج نشان داد که تفاوت بین میانگین هارمونیک روش پیشنهادی و الگوریتم طبقه‌بندی مبتنی بر GBT (قبل از نمونه‌برداری بیش از حد) با مقدار p کمتر از ۰,۰۰۱ از نظر آماری بسیار معنی‌دار بود. محاسبه فاصله اطمینان: برای هر یک از معیارهای ارزیابی دقت،

جدول ۷، نتایج مقایسه‌ای طبقه‌بندی‌های انجام‌شده در مراحل مختلف تشخیص CRT را به تصویر می‌کشد و بهترین نتایج حاصل از یادگیرنده‌های گروهی را برجسته می‌کند. در این جدول، مشاهده می‌شود که طبقه‌بندی با الگوریتم GBT پس از بهینه‌سازی پارامتری و پیش از انتخاب ویژگی‌ها، بالاترین دقت (۸۹,۸۳٪)، صحت پاسخ به درمان (۹۶,۶٪)، جامعیت عدم پاسخ به درمان (۹۷,۰۹٪) و میانگین هارمونیک (۹۰,۵۷٪) را گزارش کرده است. این نتایج نشان می‌دهد که بهینه‌سازی پارامتری نقش مهمی در بهبود عملکرد این الگوریتم داشته است. از سوی دیگر، طبقه‌بندی با الگوریتم RF پس از انتخاب ویژگی‌ها در دو معیار صحت عدم پاسخ به درمان (۹۷,۸۷٪) و جامعیت پاسخ به درمان (۹۸,۲۶٪) عملکرد مطلوبی را نشان داده است، که بیانگر توانایی این الگوریتم در تشخیص دقیق عدم پاسخ به درمان می‌باشد. در نهایت، نتایج طبقه‌بندی با یادگیری گروهی درختی پس از انتخاب ویژگی‌ها، با میانگین هارمونیک ۸۹,۰۷٪ و دقت ۸۹,۵۵٪، بهترین عملکرد کلی را در بین روش‌های ارزیابی‌شده به نمایش می‌گذارد. یادگیری گروهی در اینجا شامل ادغام الگوریتم‌های RF و GBT است که توانسته‌اند تعادل مناسبی بین صحت، دقت، و جامعیت در تشخیص پاسخ به درمان و عدم پاسخ به درمان برقرار کنند. در مجموع، این جدول نشان می‌دهد که استفاده از یادگیرنده‌های گروهی، به‌ویژه پس از مرحله انتخاب ویژگی، منجر به بهبود قابل‌توجهی در کارایی مدل‌ها شده است. همچنین، مقایسه نتایج حاکی از آن است که GBT پس از بهینه‌سازی پارامتری، با میانگین هارمونیک ۹۰,۵۷٪ و ۷۷ ویژگی، نسبت به یادگیری گروهی درختی پس از انتخاب ویژگی، با میانگین هارمونیک ۸۹,۰۷٪ و ۴۱ ویژگی، عملکرد برتری دارد. این نتایج بر اهمیت بهینه‌سازی پارامتری و انتخاب ویژگی‌ها در بهبود دقت و عملکرد مدل‌های ML در تشخیص پاسخ به درمان CRT تأکید می‌کند.

جدول ۷- مقایسه‌ی منتخب نتایج بدست آمده در مراحل مختلف تشخیص

درمان CRT

بردار ویژگی	دقت	صحت پاسخ به درمان	صحت عدم پاسخ به درمان	جامعیت پاسخ به درمان	جامعیت عدم پاسخ به درمان	میانگین هارمونیک
GBT روی	69.3 5	84.6 2	24.53	76.74	35.14	29.67

⁵² K-Fold Cross Validation (KCV)

⁵³ Paired t-test

استفاده از ۵۶ ویژگی از مجموعه دادگان Heshmat_CRT حاصل شده‌اند که بار محاسباتی بالایی را نیز به همراه دارد. در مقابل، روش پیشنهادی با استفاده از ۴۱ ویژگی منتخب و با بار محاسباتی متوسط، موفق به دستیابی به نتایج قابل توجهی شده است.

جدول ۸- مقایسه‌ی نتایج روش پیشنهادی با منتخب پیشینه‌ی تحقیق بر

روی مجموعه دادگان Heshmat_CRT

روش پیشنهادی	نژاده و همکاران [۲۵]	روش تحقیق
89.55	91.85	دقت
93.59	90.57	صحت عدم پاسخ به درمان
94.19	89.53	جامعیت پاسخ به درمان
89.07	92	میانگین هارمونیک
41	56	تعداد ویژگی
متوسط	بالا	بار محاسباتی

در این روش، هرچند معیارهای دقت و میانگین هارمونیک اندکی کمتر از تحقیق نژاده و همکاران بوده (دقت ۸۹٫۵۵٪ و میانگین هارمونیک ۸۹٫۰۷٪)، اما بهبودهای چشمگیری در صحت عدم پاسخ به درمان (۹۳٫۵۹٪) و جامعیت پاسخ به درمان (۹۴٫۱۹٪) مشاهده می‌شود. این نتایج نشان می‌دهد که روش پیشنهادی با استفاده از بردار ویژگی‌های کاهش‌یافته نه تنها بار محاسباتی را به‌طور قابل توجهی کاهش داده، بلکه با بهبود معیارهای کلیدی، عملکرد بهتری در تشخیص پاسخ به درمان ارائه می‌دهد. این امر بر اهمیت و کارایی روش بهینه متعادل‌سازی طبقات و بهره‌گیری از الگوریتم‌های یادگیری گروهی در بهبود عملکرد مدل‌های تشخیصی تاکید می‌کند. البته می‌توان به پژوهش دیگر نژاده و همکاران [۲۴]، اشاره نمود که با استفاده از یک رویکرد خوشه‌بندی جدید، بار محاسباتی را برای عملیات پیش‌پردازش کاهش دادند. هر چند نتایج گزارش شده آنها روی طبقه‌بندی‌های غیر بهینه ارائه گردید اما آنها جامعیت طبقه‌ی عدم پاسخ به درمان ۸۴٫۲۵٪ گزارش کردند که تنها ۰٫۶۳٪ (کمتر از ۱٪)، از روش پیشنهادی کمتر است، که با توجه به اینکه گزارش آنها بر اساس طبقه‌بندی NN از نوع پرسپترون ساده است یعنی می‌توان نتایج بهتری در صورت بهبود طبقه‌بند در نظر گرفت. هر چند برای سایر معیارهای طبقه‌بندی

صحت، حساسیت، ویژگی و میانگین هارمونیک، فاصله اطمینان ۹۵٪ محاسبه شد. فاصله اطمینان برای میانگین هارمونیک روش پیشنهادی [۸۷٫۲۳٪، ۹۰٫۹۱٪] به دست آمد که نشان دهنده پایداری بالای روش در تکرارهای مختلف است. نزدیکی این فاصله اطمینان، واریانس کم و قابلیت اعتماد بالای نتایج را نشان می‌دهد.

ارزیابی با استراتژی نمونه‌گیری منصفانه^۴: برای حفظ نسبت کلاس‌ها در مجموعه‌های آموزشی و آزمایشی، از روش نمونه‌گیری منصفانه استفاده شد. این رویکرد تضمین می‌کند که در هر بار اعتبارسنجی KCV، نسبت پاسخ‌دهندگان به درمان و غیرپاسخ‌دهندگان به درمان مشابه با مجموعه داده‌های کلی، آن نمونه است که امکان ارزیابی دقیق‌تر عملکرد مدل را فراهم می‌کند.

تحلیل حساسیت پارامترها: تحلیل حساسیت برای پارامترهای کلیدی الگوریتم CSA، که شامل اندازه جمعیت کلاغ‌ها، احتمال آگاهی و طول پرواز بود، انجام شد. نتایج نشان داد که روش پیشنهادی در برابر تغییرات جزئی در این پارامترها مقاوم است و تغییرات در میانگین هارمونیک در سطح ۲٪± باقی می‌ماند که نشان‌دهنده پایداری روش است.

نتایج حاصل از این آزمون‌های آماری تأیید می‌کند که بهبودهای مشاهده‌شده در روش پیشنهادی نه تنها از نظر عملی قابل توجه هستند، بلکه از نظر آماری نیز معنادار و قابل اعتماد هستند. این ارزیابی‌های جامع، اعتبار علمی یافته‌های تحقیق حاضر را افزایش داده و پتانسیل تعمیم روش پیشنهادی را به کاربردهای میدانی بزرگ تضمین می‌کند.

پس از گزارش نتایج طبقه‌بندی روش پیشنهادی، مقایسه‌ی نتایج بدست آمده با منتخب کار گذشته ارائه شده است. جدول ۸ به مقایسه نتایج روش پیشنهادی با نتایج به‌دست‌آمده از تحقیق نژاده و همکاران [۲۵]، بر روی مجموعه دادگان Heshmat_CRT می‌پردازد و در آن بهترین نتایج به‌دست‌آمده از یادگیرنده‌های گروهی مورد بررسی قرار گرفته است. هرچند نتایج تحقیق نژاده و همکاران با دقت ۹۱٫۸۵٪ و میانگین هارمونیک ۹۲٪ بسیار مطلوب به نظر می‌رسد، اما این نتایج با

آنها نیاز به بهینه‌سازی است.

این پژوهش با وجود دستاوردهای قابل توجه، با محدودیت‌هایی مواجه است که باید در تفسیر نتایج مدنظر قرار گیرند. یکی از محدودیت‌های اصلی، استفاده از یک مجموعه داده واحد Heshmat_CRT با تعداد نمونه‌های نسبتاً محدود است که شامل ۲۰۹ نمونه اولیه بوده و این حجم ممکن است بر قابلیت تعمیم‌پذیری مدل به جمعیت‌های بیمار متنوع‌تر تأثیر بگذارد. این حجم از داده‌ها ممکن است بر توانایی مدل در تعمیم به گروه‌های مختلف بیماران تأثیر بگذارد. همچنین، عدم توازن شدید بین کلاس‌ها در داده‌های اولیه، که شامل ۸۲٫۳٪ پاسخ‌دهنده و ۱۷٫۷٪ غیرپاسخ‌دهنده است، چالشی اساسی ایجاد می‌کند؛ هرچند این موضوع با استفاده از روش CSA-SMOTE تا حدود قابل توجهی قابل مدیریت است. اما نمونه‌های مصنوعی تولید شده ممکن است به طور کامل پیچیدگی‌های ذاتی داده‌های بالینی را منعکس نکنند. از دیدگاه محاسباتی، فرآیند بهینه‌سازی چند مرحله‌ای شامل اجرای الگوریتم CSA با ۱۰۰ نسل، GA با جمعیتی از ۲۵ و ۱۰۰ نسل، و تنظیم دقیق پارامترهای متعدد GBT نیاز به زمان پردازش قابل توجهی داشت و بار محاسباتی متوسطی ایجاد می‌کرد که می‌توانست در محیط‌های بالینی با منابع محدود چالش برانگیز باشد. علاوه بر این، عملکرد مدل به شدت به تنظیمات پارامترها وابسته است و یافتن ترکیب بهینه این پارامترها برای مجموعه داده‌های دیگر نیاز به تنظیم مجدد دارد. محدودیت دیگر این است که این مدل بر روی مجموعه داده‌های مستقل از بیمارستان‌ها یا کشورهای مختلف اعتبارسنجی متقابل نشده است، که ارزیابی جامع از قابلیت‌های تعمیم‌پذیری را دشوار می‌کند، به طوری که نتایج فقط با سه مطالعه قبلی [۲۵، ۲۴، ۳] بر روی همان مجموعه داده مقایسه شده است. اگرچه استفاده از تکنیک‌های یادگیری جمعی و ترکیب چندین الگوریتم نیز می‌تواند دقت را بهبود بخشد، اما پیچیدگی مدل و بنابراین تفسیرپذیری تصمیمات پزشکان را افزایش می‌دهد که این موضوع حتی ممکن است بر پذیرش بالینی سیستم تأثیر منفی بگذارد. در نهایت، معیار پاسخگویی CRT بر اساس وضعیت پاسخ به درمان بیمار پس از عمل تعریف شده است، در حالی که پیامدهای درمانی طولانی مدت

و تغییرات زمانی در پاسخ بیمار در این مدل در نظر گرفته نشده است. این محدودیت‌ها فرصت‌هایی را برای تحقیقات آینده فراهم می‌کنند که شامل جمع‌آوری داده‌های چند منبعی، توسعه روش‌های محاسباتی کارآمدتر، بهبود قابلیت تفسیر مدل‌های پیشنهادی و بررسی پیامدهای طولی درمان می‌شود.

این پژوهش با ارائه ۴۱ ویژگی و نتایجی نزدیک به پژوهش [۲۵]، جایگاه ویژه‌ای دارد. از لحاظ پردازشی عملیات طبقه‌بندی با تعداد ویژگی کمتر مدل‌سازی را انجام می‌دهد. البته رویکرد کم‌هزینه خوشه‌بندی خودکار [۲۴] و روش متعادل‌سازی طبقات پیشنهادی در کم هزینه نمودن بار محاسباتی در عین حال پاسخ مطلوب می‌تواند به عنوان پژوهش‌های آتی امیدبخش باشد.

۵. نتیجه‌گیری

در این پژوهش، با بهره‌گیری از تکنیک‌های بهینه‌سازی در ML و تمرکز بر برطرف‌سازی چالش عدم تعادل طبقات، عملکرد تشخیص پاسخ به درمان CRT به شکل چشمگیری ارتقاء یافت؛ ابتدا آشکار شد که روش‌های پایه مانند NB و NN گرچه دقت مناسبی داشتند، اما به دلیل اتکا بر طبقه اکثریت، نتایجی محدود ارائه می‌کردند، از این رو با ادغام روش SMOTE بهینه‌شده توسط الگوریتم جستجوی فراابتکاری، کیفیت نمونه‌سازی طبقات بهبود یافته و مدل GBT توانست ضمن افزایش معنی‌دار معیارهای عملکردی، قابلیت تشخیص هر دو گروه پاسخ‌دهنده و غیرپاسخ‌دهنده را به تعادل برساند. با استخراج ۱۷ ویژگی جدید و ارزیابی ۷۷ متغیر، مسئله عدم تعادل طبقات شناسایی شد و با به‌کارگیری الگوریتم فراابتکاری CSA برای بهینه‌سازی روش SMOTE، دقت مدل به‌طور قابل توجهی ارتقا یافت، به طوری که دقت کلی به ۸۷٫۲۳٪ و میانگین هارمونیک به ۸۷٫۳۱٪ رسید. این بهینه‌سازی به ویژه در بهبود تشخیص عدم پاسخ به درمان نقش مهمی داشت. سپس با استفاده از GA و تابع هدف GBT، ۳۶ ویژگی حذف و بردار ویژگی به ۴۱ متغیر کاهش یافت که موجب ساده‌سازی مدل و حفظ تعمیم‌پذیری آن گردید. در مرحله نهایی، مدل‌های یادگیری گروهی درختی، پس از انتخاب ویژگی، با دقت ۸۹٫۵۵٪ و میانگین هارمونیک ۸۹٫۰۷٪ عملکردی رقابتی نسبت به سایر روش‌ها از جمله GBT بهینه‌سازی شده (با دقت ۸۹٫۸۳٪ و میانگین هارمونیک ۹۰٫۵۷٪)

- pp. 4065-4076, 2017, doi: <https://doi.org/10.1109/TNNLS.2017.2751612>.
- [7] K. E. Bennin, J. Keung, P. Phannachitta, A. Monden, and S. Mensah, "Mahakil: Diversity based oversampling approach to alleviate the class imbalance issue in software defect prediction," *IEEE Transactions on Software Engineering*, vol. 44, no. 6, pp. 534-550, 2017, doi: <https://doi.org/10.1145/3180155.3182520>.
- [8] K. Cheng, C. Zhang, H. Yu, X. Yang, H. Zou, and S. Gao, "Grouped SMOTE with noise filtering mechanism for classifying imbalanced data," *IEEE Access*, vol. 7, pp. 170668-170681, 2019, doi: <https://doi.org/10.1109/ACCESS.2019.2955086>.
- [9] A. S. Hussein, T. Li, C. W. Yohannese, and K. Bashir, "A-SMOTE: A new preprocessing approach for highly imbalanced datasets by improving SMOTE," *International Journal of Computational Intelligence Systems*, vol. 12, no. 2, pp. 1412-1422, 2019, doi: <https://doi.org/10.2991/ijcis.d.191114.002>.
- [10] R. N. Laveti, A. A. Mane, and S. N. Pal, "Dynamic Stacked Ensemble with Entropy based Undersampling for the Detection of Fraudulent Transactions," in *2021 6th International Conference for Convergence in Technology (I2CT)*, 2021: IEEE, pp. 1-7, doi: <https://doi.org/10.1109/I2CT51068.2021.9417896>.
- [11] A. Sharma, P. K. Singh, and R. Chandra, "SMOTified-GAN for class imbalanced pattern classification problems," *Ieee Access*, vol. 10, pp. 30655-30665, 2022, doi: <https://doi.org/10.1109/ACCESS.2022.3158977>.
- [12] D. Sisodia and D. S. Sisodia, "Quad division prototype selection-based k-nearest neighbor classifier for click fraud detection from highly skewed user click dataset," *Engineering Science and Technology, an International Journal*, vol. 28, p. 101011, 2022, doi: <https://doi.org/10.1016/j.jestch.2021.05.015>.
- [13] S. Janbhasha, S. N. Bhavanam, and K. Harshita, "GAN-Based Data Imbalance Techniques for ECG Synthesis to Enhance Classification Using Deep Learning Techniques and Evaluation," presented at the 2023 Third International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT), 2023, doi: <https://doi.org/10.1109/I2CT51068.2021.9417897>.

از خود نشان دادند، در حالی که مزایایی نظیر سادگی مدل و صحت بالاتر در تشخیص عدم پاسخ به درمان را نیز ارائه دادند. یافته‌های حاکی از آن است که ترکیب بهینه‌سازی بیش‌نمونه‌برداری و انتخاب ویژگی، مدلی منسجم و قابل اعتماد برای پشتیبانی از تصمیم‌گیری در CRT فراهم می‌آورد و با گسترش آینده‌نگرانه به داده‌های بالینی و مارکرهای زیستی، افق نوینی در بهبود نتایج بیماران نارسایی قلبی می‌گشاید.

۶. منابع

- [1] O. O. Awe, J. B. Ojumu, G. A. Ayanwoye, J. S. Ojumoola, and R. Dias, "Machine Learning Approaches for Handling Imbalances in Health Data Classification," in *Sustainable Statistical and Data Science Methods and Practices: Reports from LISA 2020 Global Network, Ghana, 2022*, O. O. Awe and E. A. Vance Eds.: Springer Nature Switzerland, 2023, pp. 375-391, doi: https://doi.org/10.1007/978-3-031-41352-0_19.
- [2] S. Huda *et al.*, "An ensemble oversampling model for class imbalance problem in software defect prediction," *IEEE access*, vol. 6, pp. 24184-24195, 2018, doi: <https://doi.org/10.1109/ACCESS.2018.2817572>.
- [3] M. Nejadeh, P. Bayat, J. Kheirkhah, and H. Moladoust, "Evaluation of Pattern Recognition Techniques in Response to Cardiac Resynchronization Therapy (CRT)," *Information Systems & Telecommunication*, p. 197, 2020. [Online]. Available: <http://jist.acecr.org/Article/15549/FullText>.
- [4] Y. Dong, H. Xiao, and Y. Dong, "SA-CGAN: An oversampling method based on single attribute guided conditional GAN for multi-class imbalanced learning," *Neurocomputing*, vol. 472, pp. 326-337, 2022, doi: <https://doi.org/10.1016/j.neucom.2021.04.135>.
- [5] S. S. Rawat and A. K. Mishra, "Review of Methods for Handling Class-Imbalanced in Classification Problems," *arXiv preprint arXiv:2211.05456*, 2022, doi: <https://doi.org/10.48550/arXiv.2211.05456>.
- [6] J. Mathew, C. K. Pang, M. Luo, and W. H. Leong, "Classification of imbalanced data by oversampling in kernel space of support vector machines," *IEEE transactions on neural networks and learning systems*, vol. 29, no. 9,

- [22] M. Soleimani and A. S. Mirshahzadeh, "Multi-class classification of imbalanced intelligent data using deep neural network," *EAI Endorsed Transactions on AI and Robotics*, vol. 2, 2023, doi: <https://doi.org/10.4108/airo.3486>.
- [23] Z. Sharifi Mehrjard, H. Momeni, and H. Adabi Ardekani, "A review of machine learning algorithms to diagnose autism using the EEG signal," *Soft Computing Journal*, vol. 13, no. 1, pp. 2-19, 2024, doi: <https://doi.org/10.22052/scj.2023.248522.1110>.
- [24] M. Nejadeh, P. Bayat, J. Kheirkhah, and H. Moladoust, "Reinforcement of Feature Vector Learning for Predicting Cardiac Resynchronization Therapy Using Automatic Weighted Clustering," *International Journal of Engineering*, vol. 38, no. 9, pp. 2038-2060, 2025, doi: <https://doi.org/10.5829/ije.2025.38.09c.03>.
- [25] M. Nejadeh, P. Bayat, J. Kheirkhah, and H. Moladoust, "Predicting the response to cardiac resynchronization therapy (CRT) using the deep learning approach," *Biocybernetics and Biomedical Engineering*, vol. 41, no. 2, pp. 758-778, 2021/04/01/ 2021, doi: <https://doi.org/10.1016/j.bbe.2021.05.002>.
- [26] F. Salahi, "Improve opinion mining by combining gray wolf algorithm and support vector machine," *Soft Computing Journal*, vol. 13, no. 1, pp. 96-107, 2024, doi: <https://doi.org/10.22052/scj.2025.248730.1122>.
- [27] H. Shafiei, V. Rafe, and M. Amiri, "Optimization of vehicle routing based on the combination of ant colony and particle swarm algorithms with the heuristic function of the cosine of angles," *Soft Computing Journal*, vol. 12, no. 2, pp. 146-164, 2024. [Online]. Available: <https://doi.org/10.22052/scj.2023.248702.1118>.
- [28] C. Sheng and H. Yu, "An optimized prediction algorithm based on XGBoost," in *2022 International Conference on Networking and Network Applications (NaNA)*, 2022: IEEE, pp. 1-6, doi: <https://doi.org/10.1109/NANA56854.2022.00082>.
- [29] M. Sarchahi and E. Mahdipour, "Using ensemble deep learning to improve the accuracy of CT-Scan lung image detection of COVID-19 patients," *Soft Computing Journal*, vol. 13, no. 1, pp. 20-39, 2024, doi: <https://doi.org/10.22052/scj.2023.253142.1158>.
- [14] T. Widmann, "Classifying Imbalanced Data: The Relevance of Accuracy and Feature Importance," presented at the Proceedings of the 57th Hawaii International Conference on System Sciences, Hawaii, 2024. [Online]. Available: https://aisel.aisnet.org/hicss-57/da/trustworthy_ai/3/.
- [15] J. Duan, X. Yang, S. Gao, and H. Yu, "A partition-based problem transformation algorithm for classifying imbalanced multi-label data," *Engineering Applications of Artificial Intelligence*, vol. 128, p. 107506, 2024/02/01/ 2024, doi: <https://doi.org/10.1016/j.engappai.2023.107506>.
- [16] M. Woźniak, M. Wiecek, and J. Siłka, "BiLSTM deep neural network model for imbalanced medical data of IoT systems," *Future Generation Computer Systems*, vol. 141, pp. 489-499, 2023/04/01/ 2023, doi: <https://doi.org/10.1016/j.future.2022.12.004>.
- [17] F. Thabtah, S. Hammoud, F. Kamalov, and A. Gonsalves, "Data imbalance in classification: Experimental evaluation," *Information Sciences*, vol. 513, pp. 429-441, 2020/03/01/ 2020, doi: <https://doi.org/10.1016/j.ins.2019.11.004>.
- [18] Z. Xu, D. Shen, T. Nie, Y. Kou, N. Yin, and X. Han, "A cluster-based oversampling algorithm combining SMOTE and k-means for imbalanced medical data," *Information Sciences*, vol. 572, pp. 574-589, 2021/09/01/ 2021, doi: <https://doi.org/10.1016/j.ins.2021.02.056>.
- [19] W. Nazar *et al.*, "Artificial intelligence models in prediction of response to cardiac resynchronization therapy: a systematic review," *Heart Failure Reviews*, vol. 29, no. 1, pp. 133-150, 2024, doi: <https://doi.org/10.1007/s10741-023-10357-8>.
- [20] U. C. Nguyễn, K. Vernooy, and F. W. Prinzen, "Quest for the ideal assessment of electrical ventricular dyssynchrony in cardiac resynchronization therapy," *Journal of Molecular and Cellular Cardiology Plus*, vol. 7, p. 100061, 2024, doi: <https://doi.org/10.1016/j.jmccpl.2024.100061>.
- [21] A. M. Halim, M. Dwifabri, and F. Nhita, "Handling Imbalanced Data Sets Using SMOTE and ADASYN to Improve Classification Performance of Ecoli Data Sets," *Building of Informatics, Technology and Science (BITS)*, vol. 5, no. 1, pp. 246-253, 2023, doi: <https://doi.org/10.47065/bits.v5i1.3647>.