Journal of Quality Engineering and Management

www.pqprc.ir

J. Qual. Eng. Manag. Vol. 16, No. 2 (2026) 171–185.

Paper Type: Original Article

Comparing Classical Statistical and Machine Learning Models in Survival Analysis

Kimia Samavati[1], Akram Kohansal[1,*], Hashem Talamkhani[1], Zahra Barikbin[2]

[1] Department of Statistics, Imam Khomeini International University, Qazvin, Iran; kimia.samavatiii@gmail.com; kohansal@sci.ikiu.ac.ir; hashemtalam1@gmail.com;
[2] Department of Applied Mathematics, Imam Khomeini International University, Qazvin, Iran; barikbin@sci.ikiu.ac.ir.

Citation:

Received: 16 September 2025 *Revised: 27 November 2025* *Accepted: 02 January 2026*	Samavati, K., Kohansal, A., Talamkhani, H., & Barikbin, Z. (2026). Comparing classical statistical and machine learning models in survival analysis. *Journal of Quality Engineering and Management*, *16*(2), 171-185.

Abstract

Purpose: The purpose of this paper is to evaluate the performance of classical survival analysis techniques and contemporary machine learning algorithms in predicting survival outcomes among patients diagnosed with Heart Failure (HF). The study aims to determine whether machine learning–based models provide improved predictive accuracy compared to conventional statistical approaches.

Methodology: The study applied Kaplan–Meier, log-rank, Cox, and Weibull models alongside machine learning techniques including decision tree, random forest, Inverse Probability of Censoring Weighting (IPCW) logistic regression, Support Vector Machine (SVM), and a survival neural network. Model performances were evaluated based on predictive accuracy.

Findings: The Cox model performed best among classical methods, while IPCW logistic regression and random forest showed the highest accuracy among machine learning models. The survival neural network performed similarly to Cox and random forest without significant improvement.

Originality/Value: The paper provides a concise comparison showing that machine learning models can outperform traditional survival analysis approaches in predicting heart failure outcomes.

Keywords: Survival analysis, Heart failure, Random forest, Logistic regression, Survival neural network (DeepSurv).

Corresponding Author: kohansal@sci.ikiu.ac.ir
10.48313/jqem.2026.573578.1599
Licensee System Analytics. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0).

مهندسی و مدیریت کیفیت

www.pqprc.ir

دوره ۱۶، شماره (۲)، (۱۴۰۵)، ۱۸۵-۱۷۱

نوع مقاله: پژوهشی

مقایسه مدل‌های کلاسیک آماری و یادگیری ماشین در تحلیل بقا

کیمیا سماواتی[۱]، اکرم کهن‌سال[۱،*]، هاشم تلمخانی[۱]، زهرا باریک‌بین[۲]

[۱] گروه آمار، دانشگاه بین المللی امام خمینی (ره)، قزوین، ایران.

[۲] گروه ریاضی کاربردی، دانشگاه بین المللی امام خمینی (ره)، قزوین، ایران.

چکیده

هدف: هدف این مقاله ارزیابی عملکرد روش‌های کلاسیک تحلیل بقای بیماری و الگوریتم‌های نوین یادگیری ماشین در پیشبینی پیامدهای بقای بیماران مبتلا به نارسایی قلبی است. این مقاله می‌کوشد مشخص کند آیا مدل‌های مبتنی بر یادگیری ماشین می‌توانند دقت پیشبینی بالاتری نسبت به رویکردهای آماری مرسوم ارایه دهند یا خیر.

روش‌شناسی پژوهش: در این مقاله از مدل‌های کاپلان–مایر، آزمون لگرنک، مدل خطرات متناسب کاکس و مدل پارامتریک وایبول در کنار تکنیک‌های یادگیری ماشین شامل درخت تصمیم، جنگل تصادفی، رگرسیون لجستیک وزن‌دهی‌شده بر اساس احتمال سانسور[۱]، ماشین بردار پشتیبان[۲] و یک شبکه عصبی بقا استفاده شد وعملکرد مدل‌ها بر اساس دقت پیشبینی ارزیابی گردید.

یافته‌ها: مدل کاکس در میان روش‌های کلاسیک بهترین عملکرد را نشان داد، در حالی که رگرسیون لجستیک IPCW و جنگل تصادفی بالاترین دقت را در میان مدل‌های یادگیری ماشین به دست آوردند. شبکه عصبی بقا نیز عملکردی مشابه مدل کاکس و جنگل تصادفی داشت، اما بهبود معناداری نسبت به آن‌ها ایجاد نکرد.

اصالت/ارزش‌افزوده علمی: این مقاله مقایسه‌ای مختصر و موثر ارایه می‌دهد که نشان می‌دهد مدل‌های یادگیری ماشین می‌توانند در پیشبینی نتایج بیماران مبتلا به نارسایی قلبی نسبت به روش‌های کلاسیک تحلیل بقا عملکرد برتری داشته باشند.

کلیدواژه‌ها: تحلیل بقا، نارسایی قلبی، جنگل تصادفی، رگرسیون لجستیک، شبکه عصبی مصنوعی (DeepSurv).

۱- مقدمه

تحلیل بقا شاخه‌ای تخصصی از علم آمار است که به مطالعه و مدل‌سازی مدت زمان تا وقوع یک رویداد خاص می‌پردازد. این رویداد می‌تواند مرگ بیمار، بهبود بیماری، خرابی قطعه صنعتی، ترک شغل یا هر اتفاقی باشد که زمان رخ دادن آن اهمیت دارد. ویژگی متمایز تحلیل بقا، تمرکز آن بر متغیر زمان و برخورد با داده‌های ناقص یا سانسورشده است به این معنا که در بسیاری از مطالعات، زمان دقیق وقوع رویداد برای برخی افراد مشخص

[1] Inverse Probability of Censoring Weighting (IPCW)

[2] Support Vector Machine (SVM)

نیست یا رویداد تا پایان دوره مطالعه رخ نمی‌دهد. به همین دلیل، تحلیل بقا ابزارها و مدل‌های خاصی برای مدیریت چنین داده‌هایی دارد. کاربردهای تحلیل بقا گسترده است. در پزشکی برای ارزیابی اثربخشی درمان‌ها، مقایسه طول عمر بیماران و پیش‌بینی بازگشت بیماری استفاده می‌شود. همچنین در مهندسی، برای پیش‌بینی زمان شکست تجهیزات و افزایش عمر مفید آن‌ها کاربرد دارد. در علوم اجتماعی نیز به مطالعه مدت زمان وقوع رویدادهایی مانند ازدواج، بیکاری یا تغییر شغل می‌پردازد و در حوزه بیمه و مالی برای پیش‌بینی مرگ بیمه‌گذاران و محاسبه حق بیمه بر اساس زمان وقوع رویدادهای بیمه‌ای به کار می‌رود [1]. علاوه بر تعیین مدت زمان تا وقوع رویداد، تحلیل بقا امکان شناسایی و بررسی عوامل موثر بر این مدت زمان را نیز فراهم می‌کند. به طور کلی، تحلیل بقا ابزاری کلیدی برای تحلیل داده‌های زمانی و سانسورشده است که پژوهشگران را قادر می‌سازد علاوه بر پیش‌بینی زمان وقوع رویداد، عوامل موثر را نیز شناسایی نمایند. در تحلیل بقا، متغیر تصادفی اصلی را با T نمایش می‌دهیم که نشان‌دهنده

مدت زمان تا وقوع یک رویداد خاص است [2]. این رویداد می‌تواند مرگ، خرابی یک قطعه، یا هر رخداد دیگری باشد. مقدار T همواره مثبت خواهد بود. تابع بقا احتمال این را بیان می‌کند که یک واحد مورد بررسی تا زمان t هنوز دچار رویداد مورد نظر (مثلا مرگ یا خرابی) نشده باشد. تعریف ریاضی آن به صورت

$$S(t) = P(T > t),$$

است و یکی از مهم‌ترین ابزارها در تحلیل بقا است و به ما این امکان را می‌دهد که احتمال زنده ماندن یا سالم ماندن یک واحد را تا زمان دلخواه محاسبه کنیم.

مدل پارامتری تحلیل بقا مدلی است که فرض می‌کند زمان بقا از یک توزیع مشخص و شناخته شده پیروی می‌کند. این مدل‌ها شامل توزیع‌های نمایی ، وایبول ، لگ-نرمال ، لگ-لاجستیک و غیره هستند. مدل پارامتری به صورت کلی می‌تواند تابع بقا را به شکل تابعی از پارامترهای توزیع بیان کند [3]. برای مثال، تابع بقا مدل نمایی به صورت

$$S(t) = exp(-\lambda t),$$

است. در این فرمول $\lambda > 0$ نرخ شکست ثابت است. مدل وایبول به صورت کلی‌تر است و تابع بقا آن به شکل

$$S(t) = exp(-\lambda t^{\gamma}),$$

است. که $\lambda > 0$ و $\gamma > 0$ پارامترهای مدل هستند. اگر $\gamma = 1$ مدل وایبول به مدل نمایی تبدیل می‌شود. تابع چگالی احتمال توزیع وایبول به صورت

$$f(t) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} exp\left[-\left(\frac{t}{\eta}\right)^{\beta}\right], \quad t \geq 0, \qquad (۱)$$

است و برای $t < 0$ خواهیم داشت:

$$f(t) = 0.$$

تابع توزیع تجمعی به صورت

$$F(t) = 1 - exp\left[-\left(\frac{t}{\eta}\right)^{\beta}\right], \quad t \geq 0, \qquad (۲)$$

است. تابع بقا برابر است با فرمول بعدی که قرارگرفته

$$S(t) = exp\left[-\left(\frac{t}{\eta}\right)^{\beta}\right], \quad t \geq 0, \qquad (۳)$$

است. تابع نرخ شکست به شکل

$$\lambda(t) = \frac{f(t)}{S(t)} = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1}, \quad t \geq 0, \qquad (۴)$$

است. امید ریاضی توزیع وایبول نیز به این شکل

$$E(T) = \eta\, \Gamma\left(1 + \frac{1}{\beta}\right),$$

است. واریانس آن نیز به صورت

$$Var(T) = \eta^2 \left[\Gamma\left(1+\frac{2}{\beta}\right) - \left(\Gamma\left(1+\frac{1}{\beta}\right)\right)^2 \right],$$

تعریف شده است که در آن $\Gamma(\cdot)$، تابع گاما است.

مدل‌های پارامتری در شرایطی بسیار مفید هستند که توزیع زمان بقا مشخص و قابل اعتماد باشد و بتوان پیش‌بینی‌های دقیق از توزیع بقا ارایه کرد، اما نیازمند فرضیات قوی هستند. در مقابل، مدل‌های نیمه‌پارامتری مانند مدل کاکس با فرضیات ساده‌تر، انعطاف و کارایی بیشتری در شرایط مختلف داده دارند و به‌طور گسترده در تحلیل‌های بقا استفاده می‌گردد. مدل‌های نیمه‌پارامتری در تحلیل بقا به طور خاص برای زمانی به کار می‌روند که نمی‌خواهیم توزیع کامل زمان بقا را به صورت صریح فرض کنیم، ولی در عین حال مایلیم رابطه‌ی بین متغیرهای توضیحی و ریسک بقا را بررسی کنیم [4]. مدل نیمه‌پارامتریک معروف و پرکاربرد در تحلیل بقا، مدل مخاطرات متناسب کاکس است. این مدل فرض می‌کند که نرخ شکست به صورت حاصل‌ضرب تابعی از زمان و تابعی از متغیرهای توضیحی است که به شکل

$$Var(T) = \eta^2 \left[\Gamma\left(1+\frac{2}{\beta}\right) - \left(\Gamma\left(1+\frac{1}{\beta}\right)\right)^2 \right], \qquad (۶)$$

نوشته شده است که در آن $\lambda_0(t)$ نرخ شکست پایه است که به صورت غیرپارامتری مدل می‌شود، و β بردار پارامترهای مربوط به متغیرهای توضیحی X است. مدل کاکس نیازی به فرض توزیع مشخص برای زمان بقا ندارد و به همین دلیل نیمه‌پارامتری نامیده شده است.

اگرچه در سال‌های اخیر مطالعات متعددی به مقایسه مدل‌های کلاسیک تحلیل بقا و روش‌های یادگیری ماشین پرداخته‌اند، بخش قابل‌توجهی از این مطالعات صرفا بر مقایسه قدرت پیش‌بینی مدل‌ها در یک مجموعه‌داده خاص متمرکز بوده و ابعاد مهمی مانند کالیبراسیون پیش‌بینی‌ها، تغییر عملکرد مدل در افق‌های زمانی مختلف، اثر میزان سانسورشدگی داده‌ها، پایداری نتایج، تبیین‌پذیری مدل‌ها و قابلیت استفاده عملی از آن‌ها را به‌طور هم‌زمان بررسی نکرده‌اند. وجه تمایز پژوهش حاضر، ارایه یک چارچوب مقایسه برای ارزیابی هم‌زمان مدل‌های کلاسیک و الگوریتم‌های یادگیری ماشین در تحلیل داده‌ها است.

۲- الگوریتم‌های یادگیری ماشین در تحلیل بقا

الگوریتم‌های یادگیری ماشین مجموعه‌ای از روش‌ها و مدل‌های ریاضی هستند که به کامپیوترها امکان می‌دهند از داده‌ها یاد بگیرند و بر اساس آن پیش‌بینی یا تصمیم‌گیری کنند [5]. یادگیری تحت نظارت نوعی از یادگیری ماشینی است که در آن مدل با استفاده از داده‌های برچسب‌دار آموزش می‌بیند تا بتواند الگوها را شناسایی کرده و برای داده‌های جدید پیش‌بینی یا طبقه‌بندی انجام دهد. این روش به گونه‌ای طراحی شده که ورودی‌ها به همراه خروجی‌های متناظر در اختیار الگوریتم قرار می‌گیرد و هدف، یادگیری رابطه میان ورودی و خروجی می‌باشد. برخی از الگوریتم‌های مهم یادگیری تحت نظارت عبارت‌اند از، رگرسیون خطی، رگرسیون لجستیک، درخت تصمیم، ماشین بردار پشتیبان، بیز، جنگل تصادفی و الگوریتم‌های بوستینگ. این خانواده از الگوریتم‌ها پایه بسیاری از کاربردهای یادگیری ماشینی در تحلیل داده، پیش‌بینی و تصمیم‌گیری خودکار هستند و بسته به ماهیت مساله، داده و هدف مدل‌سازی انتخاب می‌گردد [6].

درخت تصمیم یکی از الگوریتم‌های پایه و محبوب در حوزه یادگیری ماشین است که به دلیل سادگی و قابلیت تفسیر بالایش، کاربرد فراوانی در مسایل طبقه‌بندی، رگرسیون و تحلیل داده‌ها دارد. ساختار درخت تصمیم متشکل از گره‌هایی است که در هر گره یک تصمیم مبتنی بر یک ویژگی از داده‌ها گرفته می‌شود و شاخه‌ها نمایانگر نتایج این تصمیمات هستند. این روند به طور بازگشتی ادامه پیدا می‌کند تا به گره‌های برگ برسد که نوع خروجی یا پیش‌بینی نهایی مدل را تعیین می‌کنند. فرآیند ساخت درخت از گره ریشه آغاز می‌شود که شامل کل داده‌های ورودی است. در هر مرحله، ویژگی‌ای که بیشترین توانایی در تفکیک داده‌ها دارد انتخاب می‌شود. این انتخاب معمولا بر اساس معیارهایی مانند کاهش آنتروپی یا بیشینه بهره‌وری اطلاعات صورت می‌گیرد [7]. سپس داده‌ها بر اساس مقدارهای ویژگی انتخاب شده تقسیم و به زیرمجموعه‌هایی با ویژگی‌های همگن‌تر تبدیل می‌شوند. این روند تقسیم‌بندی بازگشتی تا زمانی که معیار توقف برقرار شود (مانند رسیدن به حداقل نمونه در گره یا بیشینه عمق درخت) ادامه

می‌یابد. این فرآیند باعث می‌شود تا در نهایت ساختار درخت تصمیم بهینه با قابلیت تعمیم مناسب ساخته شود. آنتروپی $H(D)$ برای مجموعه داده D به صورت

$$H(D) = -\sum_{i=1}^{n} p_i \log_2(p_i), \quad (۷)$$

محاسبه می‌شود که p_i نسبت نمونه‌های متعلق به کلاس i در داده‌ها است. آنتروپی نمایانگر میزان عدم‌قطعیت یا تنوع در مجموعه داده است. درخت تصمیم‌گیری به ویژه، آنتروپی برای اندازه‌گیری خلوص یا عدم‌خلوص مجموعه داده‌ها به کار می‌رود تا بهترین تقسیم‌بندی برای جداسازی داده‌ها انتخاب شود [4]. بهره‌وری اطلاعات برای تقسیم داده با استفاده از ویژگی A به شکل

$$IG(D, A) = H(D) - \sum_{v \in (A)} \frac{|D_v|}{|D|} H(D_v), \quad (۸)$$

تعریف می‌شود که در آن D_v زیرمجموعه‌ای از داده‌ها است که ویژگی A مقدار v را دارد. بهره‌وری اطلاعات بیانگر میزان کاهش آنتروپی پس از تقسیم داده است و معیار اصلی برای انتخاب ویژگی در هر گره به شمار می‌رود.

الگوریتم جنگل تصادفی یکی از روش‌های قدرتمند و پرکاربرد در حوزه یادگیری ماشین است که به عنوان یک مدل یادگیری ترکیبی شناخته می‌شود. این الگوریتم با آموزش مجموعه‌ای از درختان تصمیم به صورت تصادفی و ترکیب نتایج آنان، دقت و تعمیم‌پذیری مدل را به‌طور قابل‌ملاحظه‌ای افزایش می‌دهد و مقاومت بالایی در برابر پدیده‌ی بیش‌برازش دارد. تصادفی بر اساس دو مفهوم کلیدی بنا شده است که شامل نمونه‌برداری با جایگزینی و انتخاب تصادفی ویژگی‌ها در هر گره تقسیم درخت است. از مجموعه داده اصلی، به تعداد B درخت نمونه‌های تصادفی با جایگزینی انتخاب می‌شود. هر یک از این نمونه‌ها، پایه آموزش یک درخت تصمیم خواهد بود. ساخت درخت‌های تصمیم تصادفی به این شکل است که در هر گره تقسیم درخت، به جای استفاده از تمامی ویژگی‌ها (ویژگی‌های داده)، یک زیرمجموعه تصادفی از ویژگی‌ها انتخاب می‌شود. سپس بهترین نقطه تقسیم بر اساس معیارهای خلوص (مانند شاخص جینی یا آنتروپی) در این زیرمجموعه محاسبه و اعمال می‌گردد. این کار منجر به افزایش تنوع بین درخت‌ها و کاهش همبستگی بین پیش‌بینی‌های آن‌ها می‌شود. سپس به سراغ ترکیب نتایج پیش‌بینی‌ها خواهیم رفت [8]. در مسایل طبقه‌بندی پاسخ نهایی بر اساس رای اکثریت درختان تعیین می‌شود و همچنین در مسایل رگرسیون میانگین خروجی‌های تمامی درختان، پیش‌بینی نهایی را تشکیل می‌دهد. پیش‌بینی نهایی مدل برای نمونه x به صورت

$$\hat{f}(x) = \frac{1}{B} \sum_{b=1}^{B} T_b(x), \quad (۹)$$

بیان می‌شود که در آن B تعداد کل درختان و $T_b(x)$ پیش‌بینی درخت شماره b برای نمونه x است. برای ساخت درخت تصمیم بهینه، باید در هر گره، ویژگی و نقطه تقسیم مناسبی انتخاب شود که باعث افزایش خلوص داده‌ها در گره‌های فرزند شود. شاخص جینی میزان عدم خلوص را به صورت

$$Gini = 1 - \sum_{k=1}^{K} p_k^2, \quad (۸)$$

محاسبه می‌کند که p_k احتمال حضور نمونه‌ای از کلاس k در آن نود و K تعداد کلاس‌ها است. مقدار کمتر شاخص جینی نشان‌دهنده خلوص بیشتر است.

حداقل کردن آنتروپی به معنای افزایش نظم و یکنواختی نمونه‌هااست. الگوریتم در هر گره سعی می‌کند ویژگی و نقطه‌ای را انتخاب کند که یکی از این معیارها را به حداقل برساند و در نتیجه تفکیک داده‌ها بین کلاس‌ها بهینه شود. با میانگین‌گیری نتایج تعداد زیادی درخت از داده‌های نمونه‌برداری شده تصادفی، جنگل تصادفی واریانس مدل را کاهش داده و از بیش‌برازش جلوگیری می‌کند. همچنین انعطاف‌پذیری بالا در ابعاد بزرگ داده که انتخاب تصادفی ویژگی‌ها در هر گره، باعث می‌شود مدل بتواند بدون افت عملکرد در داده‌های با فضای ویژگی بسیار بزرگ عمل کند. یکی دیگراز مزیت ها مدیریت داده‌های ناقص است. جنگل تصادفی قادر به مواجهه با داده‌های گمشده و پراکنده است و می‌تواند بدون حذف داده‌ها به آموزش خود ادامه دهد. همچنین امکان ارزیابی اهمیت ویژگی‌ها را خواهدداشت که با توجه به تاثیر هر ویژگی در کاهش معیار خلوص در گره‌های تقسیم،

می‌توان به صورت کمی میزان اهمیت ویژگی‌ها را تعیین کرد که این امر در انتخاب ویژگی و تفسیر مدل بسیار مفید است. این الگوریتم توانایی پایداری و قابلیت تعمیم بالا را دارد و به دلیل تنوع درختان و اصول بوت‌استرپ و انتخاب تصادفی ویژگی‌ها، جنگل تصادفی در مواجهه با داده‌های جدید و نویزی عملکرد پایدار و قابل توجهی دارد. الگوریتم جنگل تصادفی با استفاده ترکیبی از روش نمونه‌برداری و انتخاب تصادفی ویژگی‌ها، با ساخت چندین درخت تصمیم متنوع و ترکیب پیش‌بینی‌های آن‌ها، یکی از قدرتمندترین الگوریتم‌های یادگیری نظارت‌شده است که توانمندی زیادی در بهبود دقت پیش‌بینی و افزایش تعمیم‌پذیری مدل دارد. معیارهای دقیق و استاندارد شده‌ای مانند شاخص جینی و آنتروپی در طراحی هر درخت موجب ساختار بهینه و دقیق‌تر مدل می‌شوند [3].

ماشین بردار پشتیبان یک الگوریتم یادگیری ماشین با ناظر است که برای طبقه‌بندی و رگرسیون استفاده می‌شود، اما بیشتر در مسایل طبقه‌بندی دودویی کاربرد دارد. هدف اصلی این الگوریتم پیدا کردن یک خط یا ابرصفحه بهینه است که داده‌ها را به دو دسته مختلف به بهترین شکل ممکن جدا می کند. فرض کنید داده‌ها به صورت نقاطی در فضای چندبعدی نمایش داده شده است ماشین بردار پشتیبان سعی می‌کند یک ابرصفحه پیدا کند که دو دسته داده را از هم جدا کند به‌طوری که فاصله این ابرصفحه تا نزدیک‌ترین نقاط هر دسته بیشینه باشد. این فاصله به نام «حاشیه» شناخته می‌شود و هدف ماشین بردار پشتیبان بزرگ کردن این حاشیه است تا طبقه‌بندی دقیق‌تر و مقاوم‌تری را بتوانیم ایجاد کنیم. تنها نقاطی که نزدیک‌ترین به ابرصفحه هستند در تعیین موقعیت ابرصفحه نقش دارند و بقیه داده‌ها تاثیر مستقیم ندارند و اگر داده‌ها به صورت خطی قابل جدا شدن نباشند، ماشین بردار پشتیبان با استفاده از تکنیک‌های مختلف مانند ترفند کرنل داده‌ها را به فضای با ابعاد بالاتر منتقل می‌کند تا در آن فضا بتواند ابرصفحه جداکننده را پیدا کند. هدف ماشین بردار پشتیبان پیدا کردن یک مرز تصمیم بهینه است که بتواند داده‌های متعلق به دو یا چند کلاس را با بیشترین فاصله از هم جدا کند. این مرز باید طوری باشد که نزدیک‌ترین داده‌ها از هر کلاس که به آن‌ها بردار پشتیبان گویند. بیشترین فاصله را از مرز داشته باشند در حالت خطی، مرز تصمیم به صورت

$$w \cdot x - b = 0,$$

تعریف می‌شودکه به آن فرمول مرز بهینه گویند. هدف ماشین‌بردارپشتیبان بیشینه کردن مارجین است که معادل کمینه‌سازی $\| w \|$ (نرم بردار وزن‌ها) است.

$$\min_{w,b} \quad \frac{1}{2} \| w \|^2.$$

$$y_i(\mathrm{w} \cdot \mathrm{x}_i - b) \geq 1 \qquad \forall i.$$

بردار وزن به صورت ترکیبی خطی از بردارهای پشتیبان به‌صورت

$$w = \sum_{i=1}^{n} \alpha_i y_i x_i, \qquad (۱۰)$$

تعریف می‌شود. ابرصفحه بهینه، داده‌ها را با بیشترین فاصله از نزدیک‌ترین نقاط هر کلاس جدا می‌کند و باعث افزایش توان تعمیم مدل روی داده‌های جدید می‌شود. بردارهای پشتیبان، نقاط کلیدی هستند که مرز تصمیم را تعیین می‌کنند و مدل فقط به این نقاط وابسته است، نه کل داده‌ها. فرمول بهینه‌سازی تضمین می‌کند که مدل تا حد امکان ساده و با مارجین بالا باشد تا از بیش‌برازش جلوگیری شود [9].

ماشین بردار پشتیبان داده‌ها را به فضای دیگری می‌برد تا بتواند این جداسازی را انجام دهد. نقاطی که نزدیک مرز هستند و روی تصمیم‌گیری تاثیر می‌گذارند، بردار پشتیبان نامیده می‌شوند [8]. بردارپشتیبان فقط به این نقاط حساس است و اگر یکی از آن‌ها جابه‌جا شود، مرز تغییر می‌کند. این الگوریتم در مسایل پیچیده و داده‌های با ابعاد بالا هم کارایی خوبی دارد. ماشین بردارپشتیبان الگوریتمی است که بهترین خط یا صفحه را برای جدا کردن گروه‌های مختلف داده پیدا می‌کند، طوری که فاصله این خط یا صفحه از نزدیک‌ترین داده‌های هر گروه حداکثر باشد و مدل روی داده‌های جدید عملکرد خوبی داشته باشد. ماشین بردار پشتیبان به دلیل دقت بالا و توانایی کار با داده‌های پیچیده، در حوزه‌های مختلف کاربرد دارد. در پزشکی برای تشخیص و پیش‌بینی بیماری‌ها مانند سرطان استفاده می‌شود. در بیوانفورماتیک برای شناسایی ساختارهای پروتئینی و توالی‌های ژنتیکی استفاده شده است. برای تشخیص چهره و حالت‌های صورت در سیستم‌های امنیتی و نرم‌افزارهای پردازش تصویر نیز مورد استفاده قرار می‌گیرد. یکی دیگر از موارد مورد استفاده تشخیص تقلب و امنیت اطلاعات می‌باشد وهمچنین در پردازش زبان طبیعی و طبقه‌بندی متن در دسته‌بندی ایمیل‌ها، اخبار و اسناد و بسیاری از موارد دیگر می توانیم از ماشین بردار پشتیبان استفاده کنیم. می‌توان بطور ساده این گونه تفسیر کرد ماشین بردار

پشتیبان را مانند کشیدن یک خط فرضی بین دو گروه از اشیا در یک صفحه تصور کرد، به‌طوری که این خط تا حد امکان از هر دو گروه دور باشد. این فاصله زیاد باعث می‌شود که وقتی یک داده جدید وارد می‌شود، راحت‌تر و با اطمینان بیشتری بتوان آن را به یکی از دو گروه نسبت داد. نکته مهم این است که فقط چند نقطه نزدیک به این خط (بردارهای پشتیبان) تعیین می‌کنند که خط کجا کشیده شود، نه همه داده‌ها و در نهایت به طور خلاصه، ماشین بردار پشتیبان الگوریتمی است که با پیدا کردن بهترین مرز جداکننده بین دسته‌ها، داده‌ها را طبقه‌بندی می‌کند و به دلیل دقت و انعطاف‌پذیری بالا، در بسیاری از زمینه‌ها کاربرد خواهد داشت [9].

رگرسیون لجستیک یکی از روش‌های پرکاربرد در یادگیری ماشین و آمار است که برای مسایل طبقه‌بندی دودویی به کار می‌رود. در این روش، هدف پیش‌بینی احتمال تعلق هر مشاهده به یکی از دو کلاس است؛ برای مثال بیمار یا سالم، وقوع یا عدم‌وقوع رویداد، و بله یا خیر. خروجی این مدل مقداری بین صفر و یک است که به‌عنوان احتمال وقوع کلاس مورد نظر تفسیر می‌شود. رگرسیون لجستیک رابطه بین متغیرهای پیش‌بینی‌کننده و احتمال وقوع پیامد را مدل‌سازی می‌کند. این الگوریتم با استفاده از تابع لجستیک، مقادیر حاصل از ترکیب خطی متغیرها را به احتمال تبدیل می‌کند. به این ترتیب، مدل می‌تواند بر اساس ویژگی‌های ورودی، احتمال وقوع یک پیامد دودویی را برآورد کند.یکی از مزایای مهم رگرسیون لجستیک، سادگی، سرعت اجرا و تفسیرپذیری مناسب آن است. ضرایب مدل نشان می‌دهند که هر متغیر چه اثری بر احتمال وقوع پیامد دارد؛ به همین دلیل این روش در مطالعات پزشکی، اپیدمیولوژی و علوم داده کاربرد گسترده‌ای دارد. همچنین این الگوریتم برای داده‌هایی که رابطه نسبتا مشخصی میان متغیرهای مستقل و پیامد دارند، عملکرد مناسبی ارایه می‌دهد. با این حال، رگرسیون لجستیک در مدل‌سازی روابط بسیار پیچیده و غیرخطی ممکن است محدودیت داشته باشد. علاوه بر این، وجود داده‌های گمشده، هم‌خطی بین متغیرها یا نامتوازن بودن کلاس‌ها می‌تواند بر عملکرد آن اثر بگذارد. با وجود این محدودیت‌ها، رگرسیون لجستیک همچنان یکی از روش‌های پایه و معتبر برای پیش‌بینی پیامدهای دودویی به شمار می‌رود. در تحلیل بقا نیز از مفاهیم مرتبط با رگرسیون لجستیک در برخی رویکردهای طبقه‌بندی خطر و پیش‌بینی وقوع رویداد در افق‌های زمانی مشخص استفاده می‌شود. به طور کلی، این الگوریتم به دلیل سادگی، تفسیرپذیری و کارایی مناسب، یکی از ابزارهای مهم در تحلیل داده‌های پزشکی و پیش‌بینی پیامدها محسوب می‌شود [10]. این مدل به وسیله تابع لجستیک احتمال وقوع یا عدم وقوع یک رویداد دودویی را مدل می‌کند. به طور خلاصه، خروجی دودویی در یادگیری ماشین به ما کمک می‌کند تا داده‌ها را به دو دسته مختلف تقسیم کرده و پیش‌بینی‌های دقیقی ارایه کنیم. نحوه کارکردن الگوریتم رگرسیون لجستیک به صورت زیراست. مدل ابتدا مانند رگرسیون خطی، یک ترکیب خطی از ویژگی‌ها را محاسبه می‌کند

$$z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n,$$

که در آن β_0 عرض از مبدا و $\beta_1, \beta_2, \ldots, \beta_n$ ضرایب مدل هستند. برای اینکه خروجی مدل بین ۰ و ۱ باشد، مقدار خطی به تابع لجستیک (یا سیگموئید) داده شود.

$$y = \frac{1}{1 + e^{-z}},$$

که در آن e عدد اویلر است و z همان مقدار خطی مرحله قبل است. مقدار y به عنوان احتمال تعلق نمونه به کلاس مثبت تفسیر می‌شود. اگر این

$$y = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n)}}, \qquad (۱۵)$$

احتمال از یک آستانه (معمولا ۰.۵) بیشتر باشد، نمونه به کلاس مثبت وگرنه به کلاس منفی تعلق می‌گیرد. فرمول کلی مدل لجستیک به شکل

است y .احتمال تعلق نمونه به کلاس مثبت (مقداری بین ۰ و ۱) β_0 عرض از مبدا $\beta_1, \ldots, \beta_n$ ضرایب مدل و $x_1, \ldots, x_n$ ویژگی‌های نمونه یا به صورت تبدیل لوجیت هم نوشته شود که به شکل

$$ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n, \qquad (۱۶)$$

است که در آن p همان احتمال تعلق به کلاس مثبت است. رگرسیون لجستیک در حوزه‌های مختلفی مانند پزشکی و قتصاد همچنین بازاریابی و بسیاری زمینه‌های دیگر کاربرد دارد. بطور کلی رگرسیون لجستیک یک مدل طبقه‌بندی است که با ترکیب خطی ویژگی‌ها و تبدیل آن با تابع سیگموئید، احتمال تعلق نمونه به یک کلاس را پیش‌بینی می‌کند و با یک آستانه، تصمیم نهایی را می‌گیرد. رگرسیون لجستیک رتبه ای یک روش آماری در پردازش داده برای پیشبینی ارزش یک متغیر بر اساس متغیرهای وابسته استفاده می شود. این روش برای به‌دست‌آوردن خطی که بهترین برازش را با

داده‌ها دارد و کمترین خطای ممکن را ایجاد می‌کند، بهینه‌ترین یا کارآمدترین روش است. به بیان دیگر، رگرسیون لجستیک رتبه‌ای روشی است برای یافتن مدلی خطی یا تابع تصمیم که به‌صورت زیر عمل می‌کند. ابتدا برای هر نمونه‌ی داده، خطای بین مقدار پیش‌بینی‌شده و مقدار واقعی محاسبه می‌شود. سپس برای تمام نمونه‌ها، مجموع قدر مطلق خطاها (یا در برخی موارد مجموع مربعات خطاها) محاسبه می‌شود تا میزان کلی خطا به دست آید. برای کاهش مجموع مربعات خطاها، پارامترهای خط یا همان ضرایب اریبی به‌روزرسانی می‌شوند تا بهترین برازش با داده‌ها به‌دست آید. این روش یکی از شناخته‌شده‌ترین روش‌ها در رگرسیون خطی است و در حالتی به‌کار می‌رود که بین متغیرهای مستقل و متغیر وابسته رابطه‌ای خطی وجود داشته باشد. این روش با به‌روزرسانی پارامترهای مدل، بهترین خط برازش را برای داده‌ها پیدا کرده و به پیش‌بینی مقدار متغیر وابسته با کمترین خطای ممکن کمک می‌کند.

درخت رگرسیون لجستیک ترکیبی از درخت تصمیم و رگرسیون لجستیک است که برای پیش‌بینی پیامدهای دودویی به کار می‌رود. در این روش، داده‌ها ابتدا بر اساس ویژگی‌های مهم به چند بخش نسبتا همگن تقسیم می‌شوند و سپس در هر بخش یک مدل رگرسیون لجستیک برازش می‌شود. این ساختار باعث می‌شود مدل بتواند روابط پیچیده‌تر بین متغیرهای ورودی و خروجی را بهتر شناسایی کند. یکی از مزایای اصلی این روش، افزایش انعطاف‌پذیری در مقایسه با رگرسیون لجستیک ساده است. از آن‌جا که هر بخش از داده‌ها به طور جداگانه مدل‌سازی می‌شود، امکان بررسی الگوهای متفاوت در زیرگروه‌های مختلف فراهم می‌گردد. همچنین، ترکیب ساختار درختی با رگرسیون لجستیک می‌تواند تفسیر نتایج را نیز آسان‌تر کند، زیرا هم نحوه تقسیم داده‌ها مشخص است و هم اثر متغیرها در هر بخش قابل‌بررسی است.با وجود این مزایا، درخت رگرسیون لجستیک در مواجهه با داده‌های ناقص یا حجم بالای متغیرها ممکن است به پیش‌پردازش بیشتری نیاز داشته باشد. همچنین در صورت کنترل نکردن پیچیدگی درخت، احتمال بیش‌برازش نیز وجود دارد. با این حال، این روش در داده‌های پیچیده و ناهمگن می‌تواند عملکرد مناسبی داشته باشد و به عنوان مدلی قابل تفسیر برای پیش‌بینی نتایج دودویی مورد استفاده قرار گیرد.

۳- شبکه‌ی عصبی و کاربردهای آن در تحلیل بقا

شبکه‌های عصبی مصنوعی از جمله روش‌های مهم یادگیری ماشین هستند که با الهام از ساختار مغز انسان طراحی شده‌اند. این شبکه‌ها از واحدهای پردازشی ساده‌ای به نام نورون مصنوعی تشکیل شده‌اند که به‌صورت لایه‌ای به یکدیگر متصل هستند و با دریافت داده‌های ورودی، الگوها و روابط موجود در داده‌ها را یاد می‌گیرند. هر نورون پس از دریافت ورودی‌ها و اعمال وزن‌های مربوطه، با استفاده از یک تابع فعال‌سازی خروجی تولید می‌کند. در فرایند آموزش، وزن‌های شبکه به گونه‌ای تنظیم می‌شوند که خطای پیش‌بینی کاهش یابد. یکی از مهم‌ترین ویژگی‌های شبکه‌های عصبی، توانایی آن‌ها در شناسایی روابط پیچیده و غیرخطی میان متغیرها است. به همین دلیل، این روش‌ها در حوزه‌های مختلفی مانند تشخیص الگو، پردازش تصویر، پردازش زبان طبیعی و علوم پزشکی کاربرد گسترده‌ای دارند. در مسایل پزشکی، شبکه‌های عصبی می‌توانند برای پیش‌بینی پیامد بیماری، طبقه‌بندی بیماران و تحلیل داده‌های پیچیده مورد استفاده قرار گیرند. شبکه‌های عصبی را می‌توان بر اساس نوع یادگیری به سه دسته کلی تقسیم کرد. در یادگیری نظارت‌شده، شبکه با استفاده از داده‌های برچسب‌دار آموزش می‌بیند و هدف آن پیش‌بینی خروجی صحیح برای داده‌های جدید است. در یادگیری بدون نظارت، شبکه بدون استفاده از برچسب‌ها به کشف الگوها و ساختارهای پنهان در داده‌ها می‌پردازد. در یادگیری تقویتی نیز مدل از طریق تعامل با محیط و دریافت پاداش یا تنبیه، راهبرد مناسب را یاد می‌گیرد.در تحلیل بقا، شبکه‌های عصبی به دلیل توانایی در مدل‌سازی روابط پیچیده بین متغیرهای پیش‌بینی‌کننده و زمان وقوع رویداد، موردتوجه قرار گرفته‌اند. این روش‌ها می‌توانند در داده‌هایی که دارای الگوهای غیرخطی و تعاملات پیچیده هستند، عملکرد مناسبی داشته باشند و به پیش‌بینی دقیق‌تر بقا کمک کنند. از این رو، شبکه‌های عصبی به عنوان یکی از روش‌های نوین و کارآمد در تحلیل بقا شناخته می‌شوند.

۱. انتشار رو به جلو که در آن داده‌های ورودی از لایه‌های شبکه عبور می‌کنند. در هر نورون

$$z = \sum (w_i \times x_i) + b,$$

$$a = f(z),$$

برقرار است. که در آن w_i وزن‌ها، x_i ورودی‌ها، b اریبی و f تابع فعال‌سازی است. در خروجی نهایی شبکه تولید می‌شود.

۲. محاسبه خطا که در آن تفاوت بین خروجی پیش‌بینی شده و مقدار واقعی اندازه‌گیری می‌شود.

۳. گرادیان خطا نسبت به وزن‌ها محاسبه می‌شود. پس‌انتشار خطا که محاسبات از لایه خروجی به سمت لایه ورودی انجام می‌شود. از قانون زنجیره‌ای مشتق‌گیری استفاده می‌شود.

۴. به‌روزرسانی وزن‌ها، وزن‌ها با استفاده از بهینه‌سازها به‌روزرسانی می‌شوند. فرمول به‌روزرسانی به شکل

$$w_{new} = w_{old} - \eta \times \frac{\partial L}{\partial w},$$

است که در آن η نرخ یادگیری است.

شبکه‌های عصبی مصنوعی، به‌ویژه پرسپترون‌های چندلایه[1]، به دلیل انعطاف‌پذیری بالا در مدل‌سازی روابط پیچیده و غیرخطی میان متغیرها، جایگزینی کارآمد برای مدل‌های کلاسیک (مانند مدل مخاطرات تناسبی کاکس) در تحلیل بقا محسوب می‌شوند. برخلاف مدل‌های کلاسیک که اغلب بر فرضیات سخت‌گیرآن‌های مانند خطی بودن تکیه دارند، شبکه‌های عصبی قادرند الگوهای پنهان و تعاملات غیرخطی میان داده‌های بالینی، دموگرافیک و پارامترهای آزمایشگاهی را بدون نیاز به فرضیات پیش‌فرض استخراج کنند [11]. معماری این شبکه‌ها معمولا شامل یک لایه ورودی (برای متغیرهای پیش‌بین)، یک یا چند لایه پنهان با نورون‌های متعدد و توابع فعال‌سازی غیرخطی برای استخراج ویژگی‌های پیچیده، و یک لایه خروجی برای برآورد احتمال بقا یا تابع خطر است. در برخی کاربردهای خاص، از شبکه‌های عصبی بازگشتی[2] نیز استفاده می‌شود که به دلیل قابلیت حفظ اطلاعات در طول زمان، برای داده‌های زمانی و متوالی در تحلیل بقا بسیار مناسب هستند. در این مدل‌ها، برخلاف تابع خطر صریح در مدل کاکس، پیش‌بینی بر اساس وزن‌های آموزش‌دیده و توابع غیرخطی انجام می‌شود [12]. به طور کلی، خروجی شبکه عصبی به شکل

$$g(x, \theta),$$

است. در آن x تابعی غیرخطی از ورودی‌ها و θ پارامترهای شبکه است که از طریق آموزش با داده‌های بقا بهینه می‌شود. در مدل‌های شبکه عصبی برای بقا، معمولا از توابع فعال‌سازی غیرخطی برای یادگیری وزن‌ها استفاده می‌شود. معماری شبکه عصبی در تحلیل بقا معمولا بر پایه شبکه پرسپترون چندلایه است که ورودی‌ها را دریافت کرده و با استفاده از وزن‌های آموزش‌دیده و توابع فعال‌سازی غیرخطی، احتمال بقا یا خطر مرگ را پیش‌بینی می‌کند. این شبکه‌ها برخلاف مدل‌های کلاسیک رگرسیونی، فرمول مشخص و خطی ندارند و تابع پیش‌بینی آن‌ها به صورت غیرخطی و با پارامترهای آموزش‌دیده تعریف می‌شود. طراحی معماری مناسب شامل تعیین تعداد لایه‌ها و نورون‌ها، انتخاب توابع فعال‌سازی و الگوریتم آموزش است که به صورت تجربی و با توجه به داده‌ها انجام می‌شود. شبکه‌های عصبی در تحلیل بقا نه تنها ابزارهای قدرتمندی برای پیش‌بینی هستند، بلکه امکان کشف الگوهای جدید و ارایه بینش عمیق‌تر درباره رفتار سیستم‌ها را فراهم می‌کنند. این ویژگی‌ها باعث شده است که این روش‌ها به یکی از ابزارهای اصلی در تحقیقات پزشکی و سایر حوزه‌ها تبدیل شوند و مورد استفاده قرار گیرند [11].

۴- مقایسه‌ی مدل‌های کلاسیک آماری و مدل‌های یادگیری ماشین

در داده‌های بقا، همه‌ی بیماران تا وقوع رویداد مورد نظر (مانند مرگ، عود بیماری یا شکست درمان) پیگیری نمی‌شوند. برخی از بیماران ممکن است پیش از وقوع رویداد از مطالعه خارج شوند یا مدت مطالعه پایان یابد. به این حالت سانسور گفته می‌شود. نادیده گرفتن سانسور در مدل‌های آماری باعث ایجاد سوگیری در برآوردها می‌شود، زیرا داده‌های کامل‌تر وزن بیشتری در تحلیل دارند. برای رفع این سوگیری از روش وزن احتمال معکوس سانسور استفاده می‌شود. ایده‌ی اصلی روش وزن احتمال معکوس سانسور آن است که به جای حذف داده‌های سانسور شده، به هر مشاهده وزنی متناسب با احتمال سانسور نشدن تا زمان مشاهده اختصاص داده شود. به بیان دیگر، اگر احتمال اینکه یک فرد تا زمان مشخصی سانسور نشود کم باشد، ولی واقعا اطلاعات کامل از او در دسترس باشد، باید وزن بیشتری در مدل بگیرد [8].

[1] Multilayer Perceptron (MLP)

[2] Recurrent Neural Network (RNN)

فرض کنید برای فرد i داریم T_i زمان واقعی وقوع رویداد؛ C_i زمان سانسور؛ $\tilde{T}_i = min(T_i, C_i)$ زمان مشاهده‌شده؛ $\delta_i = I(T_i \leq C_i)$ شاخص وقوع رویداد (۱ اگر رویداد رخ داده و ۰ در صورت سانسور). تابع بقای سانسور، یعنی احتمال سانسور نشدن تا زمان t، به صورت

$$S_C(t) = P(C_i > t), \quad (۱۷)$$

تعریف می‌شود. که معمولا با استفاده از برآوردگر کاپلان‌مایر برآورد می‌شود. وزن معکوس سانسور برای هر مشاهده به صورت

$$w_i = \frac{\delta_i}{\hat{S}_C(\tilde{T}_i)}, \quad (۱۸)$$

محاسبه می‌شود. در این رابطه اگر فرد سانسور نشده باشد ($\delta_i = 1$)، وزن او برابر با معکوس احتمال سانسور نشدن تا آن زمان است؛ اگر فرد سانسور شده باشد ($\delta_i = 0$)، وزن او صفر در نظر گرفته می‌شود. این وزن‌دهی باعث می‌شود داده‌های مشاهده‌شده (غیرسانسور) به‌گونه‌ای رفتار کنند که نماینده‌ی کل جمعیت باشند، گویی برای همه‌ی افراد زمان رویداد مشاهده شده است. بدین‌ترتیب، مدل‌های آماری یا یادگیری ماشین که با داده‌های وزن‌دار آموزش داده می‌شوند، از سوگیری ناشی از سانسور مصون می‌مانند. به عنوان مثال اگر احتمال سانسور نشدن تا روز ۲۰۰ برابر ۰.۳ باشد و بیماری واقعا تا آن زمان پیگیری کامل داشته باشد ($\delta_i = 1$)، وزن داده‌ی او برابر $w_i = \frac{1}{0.3} \approx 3.33$، خواهد بود. یعنی داده‌ی این فرد به اندازه‌ی ۳.۳۳ فرد مشابه سانسورشده در تحلیل اثر می‌گذارد. رفع سوگیری ناشی از سانسور، استفاده از همه‌ی داده‌ها و بهبود دقت مدل را فراهم می‌نماید اما در نمونه‌های کوچک یا در صورت سانسور زیاد، وزن‌ها ممکن است ناپایدار شوند.

روش بگینگ یکی از رویکردهای یادگیری ترکیبی است که برای کاهش واریانس مدل و افزایش پایداری پیش‌بینی‌ها به‌کار می‌رود. از داده‌ی اصلی D با اندازه‌ی n، تعداد B نمونه‌ی تصادفی با جایگزینی استخراج می‌شود؛ برای هر نمونه‌ی $b = 1,2,\dots,B$، یک مدل یادگیری ماشین آموزش داده می‌شود؛ پیش‌بینی نهایی مدل برابر میانگین پیش‌بینی‌های مدل‌های مختلف است که به شکل

$$\hat{f}(x) = \frac{1}{B}\sum_{b=1}^{B} f_b(x), \quad (۱۹)$$

نمایش داده می‌شود. بررسی شاخص‌های توصیفی متغیرهای کمی (**جدول ۲**) تصویری جامع از وضعیت بالینی جمعیت مورد مطالعه ارایه می‌دهد. زمان پیگیری بیماران در دامنه ۴ تا ۲۸۵ روز با میانگین ۱۳۰.۲۶ روز گزارش شد که نشان‌دهنده ناهمگونی در طول مدت مشاهده و ماهیت داده‌های بقا است. میانگین سنی بیماران ۶۰.۸۳ سال و میانگین کسر جهشی قلب[1] ۳۸.۰۸% به‌دست آمد؛ این یافته‌ها نشان می‌دهند که اکثریت جامعه مورد مطالعه را افراد میانسال و سالمندی تشکیل می‌دهند که دچار نارسایی قلبی با کاهش قابل‌توجه عملکرد سیستولی هستند. از منظر آزمایشگاهی، متغیرهایی نظیر کراتینین سرم با میانگین ۱.۳۹ (دارای چولگی به راست) و آنزیم *CPK* با انحراف معیار بسیار بالای ۹۷۰.۲۹، بیشترین پراکندگی را نشان دادند که حاکی از تفاوت در شدت آسیب‌های قلبی و کلیوی در میان بیماران است. در مقابل، متغیرهایی مانند سدیم سرم با میانگین ۱۳۶.۶۳، توزیع متعادل‌تری داشتند. شناسایی این ویژگی‌های دموگرافیک و بالینی، زیربنای لازم را برای ورود به مراحل مدل‌سازی و تفسیر تحلیل‌های استنباطی فراهم نمود.به‌منظور مقابله با چالش داده‌های ناقص ناشی از سانسور (که ۶۸% از کل داده‌های ۲۹۹ بیمار را شامل می‌شد)، از ترکیب روش وزن‌دهی معکوس احتمال سانسور[2] و تکنیک بگینگ[3] استفاده شد. رویکرد *IPCW* با اختصاص وزن‌های بیشتر به بیمارانی که اطلاعات کامل‌تری داشتند، سوگیری ناشی از سانسور را کاهش داد، در حالی که الگوریتم بگینگ از طریق نمونه‌گیری بوت‌استرپ و میانگین‌گیری از پیش‌بینی‌ها، منجر به پایداری مدل و جلوگیری از بیش‌برازش در این حجم نمونه محدود گردید. برای ارزیابی توان تعمیم‌پذیری، داده‌ها با نسبت ۸۰ به ۲۰ به دو بخش آموزش و آزمون تقسیم شدند و فرآیند انتخاب متغیرهای بهینه با روش حذف بازگشتی ویژگی‌ها[4] انجام گرفت. نتایج این بخش نشان داد که در اکثر مدل‌ها، متغیرهای مدت پیگیری، کسر جهشی بطن چپ و کراتینین خون به عنوان حیاتی‌ترین عوامل پیش‌بین مرگ شناسایی شدند که این همگرایی میان مدل‌های مختلف، اعتبار یافته‌ها را دوچندان می‌کند.در مرحله ارزیابی نهایی، مقایسه میان مدل‌های کلاسیک آماری و الگوریتم‌های یادگیری ماشین نشان داد که هیچ الگوی واحدی برای برتری مطلق یک مدل وجود ندارد، با این حال برخی مدل‌ها عملکرد پایداری از خود نشان دادند. مدل

[1] EF
[2] IPCW
[3] Bagging
[4] RFE

کاکس به دلیل تفسیرپذیری بالا و امکان سنجش مستقیم اثر متغیرها بر خطر مرگ، همچنان ابزاری قدرتمند در محیط‌های بالینی محسوب می‌شود. در مقابل، الگوریتم جنگل تصادفی با کسب بالاترین شاخص C[1] (حدود ۰.۷۲) و *AUC* وابسته به زمان (حدود ۰.۷۱)، اندکی بهتر از مدل کاکس عمل کرد که این برتری ناشی از توانایی این الگوریتم در شناسایی روابط غیرخطی و تعاملات پیچیده میان متغیرهاست. همچنین رگرسیون لجستیک عملکرد بسیار مطلوبی به‌ویژه در افق‌های زمانی میان‌مدت (۲۵۰ روز) نشان داد. در سوی دیگر، مدل‌های پارامتری مانند وایبول به دلیل عدم انطباق با توزیع واقعی داده‌ها و درخت تصمیم به علت افت دقت در بازه‌های زمانی میانی، ضعیف‌ترین نتایج را ارایه دادند. شبکه عصبی بقا نیز علی‌رغم ظرفیت بالا، به دلیل حجم محدود داده‌ها برتری معناداری نسبت به مدل‌های ساده‌تر ایجاد نکرد. در مجموع، این پژوهش نشان داد که برای دستیابی به بهترین نتیجه در پیش‌بینی بقا، انتخاب مدل باید بر اساس اولویت بین «تفسیرپذیری بالینی» (مدل کاکس) و «بیشینه‌سازی دقت پیش‌بینی» صورت گیرد. **جدول ۱** به توصیف متغییرهای کیفی بیمارن پرداخته است.

جدول ۱- توصیف متغیرهای کیفی بیماران.

Table 1- Description of patients' qualitative variables.

ویژگی	تعداد (کل=۲۹۹)	کدگذاری
وضعیت رویداد	203	۰ زنده/ سانسور
	96	۱ فوت
جنسیت	105	۰ زن
	194	۱ مرد
سیگار	203	۰ خیر
	96	۱ بله
دیابت	174	۰ خیر
	125	۱ بله
فشار خون	194	۰ خیر
	105	۱ بله
آنمی	170	۰ خیر
	129	۱ بله

جدول ۲- توصیف متغیرهای کمی بیماران.

Table 2- Description of quantitative variables of patients.

ویژگی	انحراف معیار	میانه	میانگین	دامنه
زمان پیگیری	77.61	115	130.26	[4,285]
سن	11.89	60	60.83	[40,95]
کسر جهشی قلب	11.83	38	38.08	[14,80]
سدیم	4.41	137	136.63	[113,148]
کراتینین	1.03	1.1	1.39	[0.5,9.4]
پلاکت‌ها	97804.24	262000.0	263358.03	[25100,850000]
CPK	970.29	250	581.84	[23,7861]

جدول ۲ نشان می‌دهد که بیماران مطالعه عمدتا میانسال تا سالمند بوده‌اند و به‌طور میانگین حدود ۱۳۰ روز پیگیری شده‌اند. میانگین سنی بیماران ۶۰٫۸۳ سال و میانگین *EF* برابر ۳۸٫۰۸% بوده که بیانگر وجود نارسایی قلبی با کاهش قابل توجه عملکرد سیستولی در بیشتر بیماران است. از نظر یافته‌های آزمایشگاهی، سدیم بیماران در بیشتر موارد نزدیک به محدوده طبیعی بوده، اما برخی موارد هیپوناترمی نیز دیده شده است. کراتینین

[1] شاخص همبستگی

میانگین نسبتا بالاتری داشته که می‌تواند نشان‌دهنده وجود درجاتی از اختلال عملکرد کلیوی در بخشی از بیماران باشد. همچنین تعداد پلاکت‌ها پراکندگی نسبتا زیادی داشته، اما توزیع آن تا حدی متعادل بوده است. در میان متغیرها، *CPK* بیشترین پراکندگی و چولگی را نشان داده و بیانگر ناهمگونی زیاد وضعیت بیماران از نظر آسیب عضلانی یا قلبی بوده است. در مجموع، این جدول نشان می‌دهد که جمعیت مورد مطالعه شامل بیمارانی با نارسایی قلبی و *EF* پایین بوده و برخی شاخص‌های آزمایشگاهی مانند کراتینین و *CPK* تنوع بالایی داشته‌اند که می‌تواند در تحلیل‌های بعدی اهمیت زیادی داشته باشد. **جدول ۳** به توزیع رویدادهای فوت سانسور در داده‌ها پرداخته است.

جدول ۳- توزیع رویدادهای فوت و سانسور در داده‌ها.

Table 3- Distribution of death and censoring events in the data.

بازه زمانی (روز)	فوت	فوت (%)	سانسور	سانسور (%)
[0,50]	49	16.35	6	2
[50,100]	22	7.35	49	16.35
[100,150]	9	3.01	46	15.38
[150,200]	12	4.01	28	9.36
[200,250]	4	1.33	56	18.72
[250,300]	0	0	18	6.02

برای جلوگیری از بیش‌برازش و ارزیابی توان تعمیم مدل‌ها، داده‌ها به دو بخش آموزش و آزمون تقسیم شدند. در این پژوهش، مدل‌ها بر روی داده‌های آموزش برازش داده شدند و عملکرد آن‌ها با استفاده از شاخص *C* و *AUC* وابسته به زمان بر روی داده‌های آزمون ارزیابی شد. البته در متن اولیه نسبت تقسیم داده‌ها در برخی بخش‌ها ۹۰ به ۱۰ و در برخی دیگر ۸۰ به ۲۰ ذکر شده بود؛ بنابراین، لازم است این نسبت در نسخه نهایی یکسان‌سازی شود.برای انتخاب متغیرهای موثر، از روش حذف بازگشتی ویژگی‌ها استفاده شد. در این روش، ابتدا یک مدل پایه مانند رگرسیون لجستیک یا جنگل تصادفی بر کل متغیرها برازش داده می‌شود و سپس اهمیت هر متغیر بر اساس معیار اختصاصی همان مدل محاسبه می‌گردد. در ادامه، متغیرهای کم‌اهمیت حذف شده و مدل دوباره برازش داده می‌شود. این فرایند به‌صورت تکراری ادامه می‌یابد تا مجموعه‌ای بهینه از متغیرهای مهم به دست آید. نتایج نشان داد که در رگرسیون لجستیک، متغیرهای مدت پیگیری، کسر جهشی قلب و کراتینین خون بیشترین اهمیت را داشتند، در حالی که در جنگل تصادفی، مدت پیگیری و کراتینین خون مهم‌ترین متغیرها بودند و *CPK* نیز رتبه بالاتری کسب کرد [8].

ارزیابی عملکرد این مدل‌ها با استفاده از *AUC* وابسته به زمان و روش وزن‌دهی معکوس احتمال سانسور انجام شد. جنگل تصادفی به دلیل توانایی در شناسایی روابط غیرخطی، مقاومت در برابر نویز و امکان شناسایی متغیرهای مهم، عملکرد مناسبی نشان داد و شاخص *C* آن حدود ۰/۷۲ و *AUC* آن حدود ۰/۷۱ گزارش شد. همچنین این مدل عملکردی نزدیک به مدل کاکس داشت و متغیرهای *EF* و کراتینین را به عنوان مهم‌ترین عوامل خطر معرفی کرد. درخت تصمیم با وجود سادگی و تفسیرپذیری بالا، نسبت به سایر مدل‌ها عملکرد ضعیف‌تری داشت [8]. این مدل در افق‌های کوتاه‌مدت عملکرد نسبتا مناسبی نشان داد، اما در افق‌های میانی با افت دقت همراه بود که بیانگر ضعف آن در مدل‌سازی الگوهای پیچیده‌تر است. از این رو، اگرچه درخت تصمیم از نظر آموزشی و تفسیری مفید است، برای پیش‌بینی دقیق‌تر بقا چندان مطلوب نیست.رگرسیون لجستیک نیز عملکرد بسیار مناسبی ارایه داد و از نظر میانگین *AUC* یکی از بهترین مدل‌ها بود. این مدل به‌ویژه در افق ۲۵۰ روز به مقدار بالای *AUC* دست یافت و نشان داد که در پیش‌بینی کوتاه‌مدت و بلندمدت عملکرد پایداری دارد. سادگی، تفسیرپذیری و دقت مناسب، این مدل را به گزینه‌ای مطلوب برای کاربردهای بالینی تبدیل می‌کند.ماشین بردار پشتیبان نیز به دلیل توانایی در مدل‌سازی روابط غیرخطی بررسی شد، اما در داده‌های کوچک و دارای سانسور، عملکرد آن محدودتر بود. در مجموع، یافته‌ها نشان داد که رگرسیون لجستیک و جنگل تصادفی نسبت به سایر مدل‌ها نتایج مناسب‌تری در پیش‌بینی بقای بیماران مبتلا به نارسایی قلبی ارایه می‌دهند [8].

الگوریتم ماشین بردار پشتیبان در کوتاه‌مدت و بلندمدت عملکرد متوسطی دارد، اما در میانه‌ی افق‌ها دقت آن کاهش جدی نشان می‌دهد. میانگین *AUC* تقریبا ۰.۷۳ است که پایین‌تر از مدل‌های لجستیک و جنگل تصادفی است. به نظر می‌رسد داده‌های این مطالعه برای الگوریتم ماشین بردار پشتیبان با کرنل به اندازه کافی مناسب نبوده‌اند یا نیاز به تنظیم دقیق‌تر پارامترها دارند. مدل ماشین بردار پشتیبان با وجود توانایی در مدل‌سازی روابط

غیرخطی، در این داده‌ها نسبت به سایر الگوریتم‌ها ضعیف‌تر عمل کرده است. افت شدید در افق‌های میانی، محدودیت این الگوریتم را در کاربردهای بالینی برای پیش‌بینی بقا نشان می‌دهد.

جدول ۴- مقایسه شاخص AUC(t) الگوریتم‌ها در افق‌های زمانی مختلف.

Table 4- Distribution of death and censoring events in the data.

الگوریتم	میانگین AUC	۲۵۰ روز	۲۰۰ روز	۱۵۰ روز	۹۰ روز
رگرسیون لجستیک	0.786	0.852	0.725	0.755	0.812
جنگل تصادفی	0.784	0.824	0.713	0.758	0.840
تجمیع (رگرسیون لجستیک)	0.783	0.852	0.721	0.749	0.809
ماشین بردار پشتیبان	0.730	0.790	0.639	0.710	0.782
درخت تصمیم	0.726	0.770	0.654	0.687	0.791

جدول ۴ مقایسه شاخص *AUC(t)* الگوریتم‌های مختلف را در افق‌های زمانی ۹۰، ۱۵۰، ۲۰۰ و ۲۵۰ روز نشان می‌دهد. این شاخص بیانگر توانایی مدل‌ها در تفکیک صحیح بین افراد دارای رخداد و بدون رخداد در طول زمان است؛ به‌طوری‌که هرچه مقدار *AUC* به عدد ۱ نزدیک‌تر باشد، عملکرد مدل در پیش‌بینی بهتر خواهد بود. بر اساس نتایج این جدول، تفاوت‌هایی معنادار میان الگوریتم‌ها در افق‌های زمانی مختلف مشاهده می‌شود و این موضوع نشان می‌دهد که کارایی مدل‌ها وابسته به بازه زمانی پیش‌بینی است. در افق کوتاه‌مدت ۹۰ روز، الگوریتم جنگل تصادفی با مقدار *AUC* برابر با ۰/۸۴۰ بهترین عملکرد را در میان تمامی روش‌ها داشته است. پس از آن، رگرسیون لجستیک با مقدار ۰/۸۱۲ و روش تجمیعی مبتنی بر رگرسیون لجستیک با مقدار ۰/۸۰۹ قرار دارند. این یافته نشان می‌دهد که جنگل تصادفی در پیش‌بینی‌های کوتاه‌مدت، توانایی بیشتری در شناسایی الگوهای پیچیده و روابط غیرخطی داده‌ها دارد. در مقابل، ماشین بردار پشتیبان و درخت تصمیم به‌ترتیب با مقادیر ۰/۷۸۲ و ۰/۷۹۱ عملکرد ضعیف‌تری نسبت به دو مدل برتر نشان داده‌اند، هرچند همچنان در محدوده قابل‌قبول قرار دارند. در افق زمانی ۱۵۰ روز، مقادیر *AUC* در اکثر الگوریتم‌ها کاهش یافته است. جنگل تصادفی با مقدار ۰/۷۵۸ همچنان بالاترین عملکرد را حفظ کرده، اما اختلاف آن با رگرسیون لجستیک که مقدار ۰/۷۵۵ را به دست آورده، بسیار ناچیز است. روش تجمیعی رگرسیون لجستیک نیز با مقدار ۰/۷۴۹ عملکردی نزدیک به این دو مدل داشته است. این نزدیکی مقادیر نشان می‌دهد که در افق میان‌مدت، برتری مدل‌ها نسبت به یکدیگر کمتر می‌شود و چند الگوریتم می‌توانند دقت تقریبا مشابهی ارایه دهند.

در این بازه نیز ماشین بردار پشتیبان و درخت تصمیم افت بیشتری را تجربه کرده‌اند و به‌ترتیب به مقادیر ۰/۷۱۰ و ۰/۶۸۷ رسیده‌اند.در افق ۲۰۰ روز، روند کاهش عملکرد مدل‌ها واضح‌تر می‌شود. رگرسیون لجستیک با مقدار ۰/۷۲۵ بالاترین *AUC* را در این بازه به دست آورده و پس از آن روش تجمیعی رگرسیون لجستیک با مقدار ۰/۷۲۱ قرار گرفته است. جنگل تصادفی نیز با وجود عملکرد مناسب در کوتاه‌مدت، در این افق به مقدار ۰/۷۱۳ کاهش یافته است. این نتایج نشان می‌دهد که هرچند جنگل تصادفی در پیش‌بینی‌های زودهنگام بسیار قوی عمل می‌کند، اما در افق‌های میانی تا حدودی از پایداری آن کاسته می‌شود. در همین بازه، افت شدیدتر در ماشین بردار پشتیبان و درخت تصمیم مشاهده می‌شود؛ به‌طوری‌که مقادیر آن‌ها به‌ترتیب ۰/۶۳۹ و ۰/۶۵۴ گزارش شده است. این افت نشان می‌دهد که این دو مدل برای پیش‌بینی بقا در بازه‌های میان‌مدت، توان تفکیک پایین‌تری نسبت به سایر روش‌ها دارند.در افق بلندمدت ۲۵۰ روز، الگوی نتایج تا حدودی تغییر می‌کند. در این بازه، رگرسیون لجستیک و روش تجمیعی آن هر دو با مقدار *AUC* برابر با ۰/۸۵۲ بهترین عملکرد را ثبت کرده‌اند. این نتیجه از اهمیت زیادی برخوردار است، زیرا نشان می‌دهد مدل‌های مبتنی بر رگرسیون لجستیک در پیش‌بینی بلندمدت از ثبات و توان تفکیک بالایی برخوردارند. جنگل تصادفی با مقدار ۰/۸۲۴ همچنان عملکرد خوبی دارد، اما نسبت به دو مدل فوق در رتبه پایین‌تری قرار می‌گیرد. از سوی دیگر، ماشین بردار پشتیبان و درخت تصمیم به‌ترتیب با مقادیر ۰/۷۹۰ و ۰/۷۷۰ همچنان عملکرد پایین‌تری نسبت به مدل‌های برتر نشان می‌دهند، هرچند در مقایسه با افق ۲۰۰ روز، اندکی بهبود در آن‌ها دیده می‌شود.با بررسی میانگین *AUC* در تمامی افق‌های زمانی، رگرسیون لجستیک با میانگین ۰/۷۸۶ بالاترین مقدار را به‌دست آورده است.

جنگل تصادفی با میانگین ۰/۷۸۴ و روش تجمیعی رگرسیون لجستیک با میانگین ۰/۷۸۳ اختلاف بسیار اندکی با آن دارند. این اختلاف کم بیانگر آن است که هر سه مدل از توان پیش‌بینی مناسبی برخوردارند و در مجموع عملکرد قابل‌قبولی ارایه می‌دهند. با این حال، رگرسیون لجستیک به دلیل ثبات بیشتر در بازه‌های زمانی مختلف و دستیابی به بهترین نتیجه در افق بلندمدت، می‌تواند به‌عنوان یکی از مناسب‌ترین گزینه‌ها برای پیش‌بینی بقا معرفی شود. در مقابل، درخت تصمیم با میانگین ۰/۷۲۶ و ماشین بردار پشتیبان با میانگین ۰/۷۳۰ پایین‌ترین عملکرد کلی را ثبت کرده‌اند که

نشان‌دهنده محدودیت این مدل‌ها در استخراج الگوهای دقیق از داده‌های مورد مطالعه است.به‌طور کلی، نتایج جدول نشان می‌دهد که جنگل تصادفی برای پیش‌بینی کوتاه‌مدت مناسب‌تر است، در حالی که رگرسیون لجستیک و نسخه تجمیعی آن در پیش‌بینی‌های بلندمدت عملکرد برتری دارند. از نظر کاربردی، اگر هدف پیش‌بینی سریع و کوتاه‌مدت باشد، جنگل تصادفی می‌تواند گزینه مطلوب‌تری باشد؛ اما برای تصمیم‌گیری‌های بالینی، تحلیل بقا و پیش‌بینی در افق‌های زمانی طولانی‌تر، رگرسیون لجستیک و به‌ویژه نسخه تجمیعی آن انتخاب‌های قابل اتکاتری محسوب می‌شوند. در مقابل، درخت تصمیم و ماشین بردار پشتیبان در این مطالعه نتوانسته‌اند دقتی هم‌تراز با مدل‌های برتر ارایه دهند و بنابراین، اولویت کمتری برای استفاده در این داده‌ها دارند.

۵- نتایج

عملکرد مدل‌های کلاسیک تحلیل بقا و الگوریتم‌های یادگیری ماشین مقایسه شد تا بهترین مدل برای پیش‌بینی بقا در بیماران نارسایی قلبی مشخص شود. معیارهای اصلی ارزیابی شامل شاخص C و $AUC(t)$ وابسته به زمان بودند. مطابق **جدول ۵** شاخص C توانایی مدل را در رتبه‌بندی صحیح بیماران بر اساس خطر مرگ اندازه‌گیری می‌کند (مقدار نزدیک به ۱ نشان‌دهنده دقت بالاتر است).

جدول ۵- مقایسه مدل‌ها بر اساس شاخص C.

Table 5- Comparison of models based on C index.

مدل	شاخص C
کاپلانمایر	–
مدل کاکس	0.70
وایبول	0.62
درخت تصمیم	0.66
جنگل تصادفی	0.72
ماشین بردار پشتیبان	0.68

نتایج نشان می‌دهد که جنگل تصادفی (0.72) و مدل کاکس (0.70) بالاترین شاخص C را دارند، در حالی که مدل پارامتری وایبول با شاخص C برابر با 0.62 ضعیف‌ترین عملکرد را نشان داده است. شاخص $AUC(t)$ توانایی مدل را در تمایز بیماران با بقا کوتاه‌مدت و بلندمدت نشان می‌دهد. مشابه شاخص C مدل‌های کاکس و جنگل تصادفی بهترین عملکرد را نشان دادند. باتوجه به **جدول ۶** مدل درخت تصمیم و ماشین بردار پشتیبان در حد متوسط عمل کردند و مدل وایبول ضعیف‌ترین بود.

جدول ۶- مقایسه مدل‌ها بر اساس شاخص AUC(t).

Table 6- Comparison of models based on AUC(t) index.

مدل	AUC(t)
مدل کاکس	0.70
وایبول	0.62
درخت تصمیم	0.64
جنگل تصادفی	0.71
ماشین بردار پشتیبان	0.66

مدل کاکس، با وجود سادگی و تفسیرپذیری، همچنان عملکرد خوبی در پیش‌بینی بقا دارد. جنگل تصادفی، از نظر دقت پیش‌بینی کمی بهتر از کاکس عمل کرد، به‌خصوص در افق‌های کوتاه‌مدت. رگرسیون لجستیک با وزن احتمال معکوس سانسور و لجستیک بگینگ نیز عملکرد نزدیک به کاکس و جنگل تصادفی داشتند و حتی در افق ۲۵۰ روز AUC بالاتر از ۰٫۸۵ نسبت به کاکس نشان دادند. ماشین بردار پشتیبان و درخت تصمیم عملکرد متوسط داشتند و در افق‌های میانی افت چشمگیر نشان دادند. مدل پارامتری وایبول کمترین دقت را ارایه داد و برای پیش‌بینی بقا مناسب نبود .مدل کاکس و جنگل تصادفی بهترین عملکرد را در پیش‌بینی بقا ارایه دادند. الگوریتم‌های لجستیک و لجستیک بگینگ نیز عملکردی هم‌تراز با این مدل‌ها داشتند و در بلندمدت حتی بهتر عمل کردند. مدل‌های کاپلان-مایر و لاگ-رنک ابزارهای توصیفی و مقایسه‌ای مفیدی هستند، اما برای

پیش‌بینی فردی مناسب نیستند. EF پایین، کراتینین بالا و سن بیشتر در همه مدل‌ها به‌عنوان اصلی‌ترین عوامل خطر مرگ شناسایی شدند. بنابراین، اگر سادگی و تفسیر بالینی اهمیت دارد مدل کاکس بهترین انتخاب است. و اگر دقت پیش‌بینی در اولویت است مدل‌های جنگل تصادفی و لجستیک با وزن احتمال معکوس سانسور یا بگینگ گزینه‌های بهینه هستند.

به منظور تکمیل مقایسه میان مدل‌های کلاسیک و الگوریتم‌های یادگیری ماشین، یک شبکه عصبی ساده با رویکرد شبکه عصبی بقا نیز به‌صورت اکتشافی مورد بررسی قرار گرفت. این مدل در واقع نسخه‌ای از رگرسیون کاکس است که به‌جای ترکیب خطی متغیرها، از یک لایه پنهان شبکه عصبی برای استخراج روابط غیرخطی استفاده می‌کند. شبکه عصبی با همان متغیرهای اصلی مطالعه شامل EF (کسر جهشی قلب)، سطح کراتینین، سن و سطح سدیم آموزش داده شد. معیارهای ارزیابی همانند سایر مدل‌ها، شاخص‌های C و شاخص وابسته به زمان بودند.نتایج به دست آمده نشان داد که عملکرد شبکه عصبی از نظر شاخص C و شاخص وابسته به زمان نزدیک به مدل کاکس و الگوریتم جنگل تصادفی بود. با وجود توانایی شبکه‌های عصبی در مدل‌سازی روابط غیرخطی، در این مجموعه‌داده‌ی کوچک بهبود چشمگیری نسبت به مدل‌های کلاسیک و الگوریتم‌های ساده‌تر مشاهده نشد. بنابراین، همان‌طور که انتظار می‌رودمطابق **جدول ۷**، مزیت شبکه‌های عصبی در داده‌های بزرگ‌تر و پیچیده‌تر آشکار می‌شود، اما در این مطالعه مدل‌های کاکس، لجستیک و جنگل تصادفی همچنان بهترین گزینه‌ها باقی ماندند.

جدول ۷- مقایسه تقریبی عملکرد مدل‌ها شامل شبکه عصبی بقا.

Table 7- Approximate comparison of the performance of models including the survival neural network.

مدل	میانگین AUC(t)	شاخص تقریبی C
مدل کاکس	≈ 0.70	≈ 0.70
جنگل تصادفی	≈ 0.71	≈ 0.72
رگرسیون لجستیک	≈ 0.79	≈ 0.78
بگینگ لجستیک	≈ 0.78	≈ 0.78
شبکه عصبی بقا	$0.71 - 0.73$	$0.70 - 0.72$

نتایج این پژوهش نشان داد که هیچ مدل واحدی در تحلیل بقا بر همه روش‌ها برتری مطلق ندارد، اما در میان مدل‌های بررسی‌شده، مدل کاکس و الگوریتم جنگل تصادفی بهترین عملکرد را در پیش‌بینی بقای بیماران نارسایی قلبی نشان دادند. مدل کاکس با وجود سادگی، تفسیرپذیری بالا و کاربرد بالینی مناسب، همچنان ابزاری قابل اعتماد برای تحلیل بقا بود. در مقابل، جنگل تصادفی به دلیل توانایی در شناسایی روابط غیرخطی و تعاملات پیچیده، اندکی عملکرد بهتری داشت و بالاترین شاخص C را کسب کرد. در سوی دیگر، مدل وایبول ضعیف‌ترین عملکرد را نشان داد که احتمالا ناشی از عدم‌انطباق فرض توزیع پارامتری آن با داده‌های واقعی بود. درخت تصمیم و ماشین بردار پشتیبان نیز عملکردی متوسط داشتند و از نظر دقت و پایداری به سطح کاکس و جنگل تصادفی نرسیدند. همچنین رگرسیون لجستیک با وزن احتمال معکوس سانسور و بگینگ لجستیک عملکرد مطلوبی داشتند و در برخی افق‌های زمانی حتی از مدل کاکس بهتر ظاهر شدند؛ بنابراین، این روش‌ها نیز می‌توانند گزینه‌های مناسبی برای پیش‌بینی بقا باشند. در بررسی اکتشافی، شبکه عصبی بقا عملکردی نزدیک به کاکس و جنگل تصادفی داشت، اما برتری آشکاری نسبت به آن‌ها نشان نداد؛ موضوعی که احتمالا به حجم محدود نمونه و تعداد کم متغیرها مربوط بود. همچنین EF پایین، کراتینین بالا و سن بیشتر در اغلب مدل‌ها به‌عنوان مهم‌ترین عوامل خطر مرگ شناسایی شدند که نشان‌دهنده پایداری و اعتبار این متغیرها در پیش‌بینی بقاست. در مجموع، یافته‌ها نشان دادند که مدل‌های کلاسیک آماری، به‌ویژه کاکس، همچنان از نظر تفسیرپذیری و کاربرد بالینی ارزشمند هستند، در حالی که برخی روش‌های یادگیری ماشین، به‌خصوص جنگل تصادفی و مدل‌های لجستیکی، می‌توانند دقت پیش‌بینی بیشتری فراهم کنند. بنابراین، انتخاب مدل مناسب باید بر اساس هدف پژوهش، ماهیت داده‌ها و اولویت میان تفسیرپذیری یا دقت پیش‌بینی انجام شود. با این حال، نتایج باید با توجه به محدودیت‌هایی مانند حجم نمونه کم، تعداد محدود متغیرها و تک‌مرکزی بودن داده‌ها تفسیر شوند.

تشکر و قدردانی

نویسندگان بدین‌وسیله از سردبیر محترم و داوران گرامی نشریه به‌دلیل ارایه نظرات علمی دقیق و سازنده که نقش موثری در بهبود کیفیت و انسجام این مقاله داشته است، صمیمانه قدردانی می‌کنند.

منابع مالی

این پژوهش هیچ‌گونه حمایت مالی مشخصی از سوی سازمآن‌ها، نهادهای دولتی، موسسات خصوصی یا غیرانتفاعی دریافت نکرده است.

تعارض با منافع

ما نویسندگان این مقاله، بدینوسیله اعلام می‌داریم که هیچ‌گونه تعارض منافع شخصی یا سازمانی در ارتباط با انجام و انتشار این پژوهش وجود ندارد. همچنین تایید می‌کنیم که نسخه نهایی مقاله را مطالعه و تایید کرده‌ایم، این مقاله حاصل کار اصیل ماست، پیش‌تر منتشر نشده و هم‌اکنون نیز تحت بررسی یا چاپ در نشریه دیگری نیست.

منابع

[1] Miller, J. R. (1981). *Survival analysis*. John Wiley & Sons. https://books.google.com/books?id=cttQAAAAMAAJ
[2] Lucijanic, M., & Petrovecki, M. (2012). Analysis of censored data. *Biochemia Medica*, *22*(2), 151–155. https://hrcak.srce.hr/83219
[3] Deo, S. V., Deo, V., & Sundaram, V. (2021). Survival analysis—Part 2: Cox proportional hazards model. *Indian Journal of Thoracic and Cardiovascular Surgery*, *37*(2), 229–233. https://doi.org/10.1007/s12055-020-01108-7
[4] Weisberg, S. (2005). *Applied linear regression*. McGraw-Hill. https://doi.org/10.1002/0471704091.
[5] Vallarino, D. (2024). Machine learning algorithms for survival analysis: Advantages, disadvantages, and examples. *International Journal of Artificial Intelligence and Machine Learning*, *4*(1), 10–21. https://doi.org/10.51483/IJAIML.4.1.2024.10-21
[6] Khodaverdi, M. (2018). *Survival prediction in heart failure patients using machine learning approaches* [Thesis].
[7] Srujana, B., Verma, D., & Naqvi, S. (2024). Machine learning vs. survival analysis models: A study on right censored heart failure data. *Communications in Statistics - Simulation and Computation*, *53*(4), 1899–1916. https://doi.org/10.1080/03610918.2022.2060510
[8] Nichols, J. A., Herbert Chan, H. W., & Baker, M. A. B. (2019). Machine learning: Applications of artificial intelligence to imaging and diagnosis. *Biophysical Reviews*, *11*(1), 111–118. https://doi.org/10.1007/s12551-018-0449-9
[9] Ahmad, T., Munir, A., Bhatti, S. H., Aftab, M., & Raza, M. A. (2017). Survival analysis of heart failure patients: A case study. *PloS one*, *12*(7), e0181001. https://doi.org/10.1371/journal.pone.0181001
[10] Dong, G., Mao, L., Huang, B., Gamalo-Siebers, M., Wang, J., Yu, G., & Hoaglin, D. C. (2020). The inverse-probability-of-censoring weighting (IPCW) adjusted win ratio statistic: An unbiased estimator in the presence of independent censoring. *Taylor & Francis*, *30*(05), 882–899. 10.1080/10543406.2020.1757692
[11] Katzman, J. L., Shaham, U., Cloninger, A., Bates, J., Jiang, T., & Kluger, Y. (2018). DeepSurv: Personalized treatment recommender system using a Cox proportional hazards deep neural network. *BMC Medical Research Methodology*, *18*(1), 24. https://doi.org/10.1186/s12874-018-0482-1
[12] Fotso, S. (2018). *Deep neural networks for survival analysis based on a multi-task framework*. https://doi.org/10.48550/arXiv.1801.05512