ORIGINAL ARTICLE

# *Intelligent and fast recognition of heart disease based on synergy of linear neural network and logistic regression model*

Mojtaba Heravi[1],
Saeed Setayeshi[2]

[1] MSc, Department of Knowledge Engineering and Decision Sciences, Department of Computer, School of Management and Computer Engineering, University of Economic Sciences, Tehran, Iran
[2] Associate Professor, Department of Medical Radiation Engineering, School of Physics and Nuclear Engineering, Amirkabir University of Technology, Tehran, Iran

(Received September 29, 2013; Accepted April 13, 2014)

## *Abstract*

**Background and purpose:** Diseases have been the greatest threat for human being along the history. Heart disease (HD) has gained special attention in medical studies. Recently studying on classification and diagnosis of HD as a key topic and a lot of researches have been done in order to increase precise and reduce error in this type of decisions. With development of intelligent learning systems, these systems have played a great role in reducing the error of decision support systems (DSS).

**Materials and methods:** In this study, a simple hybrid model of logistic regression and single-layer perceptron neural network was presented which was trained with four-different learning rules (separately). The model for improving the classification and patterns recognition of HD has been used on clinical data of 270 patients from the Cleveland Clinic (UCI website). This method has been used in statistical data normalization and detection of noisy data, network training with only 20% of the data exist was performed. The model has been implemented in MATLAB.

**Results:** The mean-error of the proposed model on the total dataset was 11.11%, which was achieved a significant improvement compared to recent similar methods. In addition, the results showed that the proposed approach was very capable in dealing with noise in the data.

**Conclusion:** The results clearly showed that the linear proposed technique had large impact on reducing the error in the classification and identification of patients more accurately in a shorter time than conventional methods and complex nonlinear. The method can help physicians for early detection of disease or as a DSS.

***Keywords:*** Heart diseases diagnosis, biomedical engineering, pattern recognition, machine learning, artificial neural network, single layer perceptron, logistic regression

**J Mazand Univ Med Sci 2014; 24(112): 78-87 (Persian).**

[ Downloaded from jmums.mazums.ac.ir on 2026-11-02 ]

پژوهشی

مجله دانشگاه علوم پزشکی مازندران
دوره بیست و چهارم شماره ۱۱۲ اردیبهشت سال ۱۳۹۳ (۸۷-۷۸)

# تشخیص هوشمند و سریع بیماری قلبی بر اساس هم‌افزایی شبکه‌های عصبی خطی و روش رگرسیون منطقی

مجتبی هروی[۱]
سعید ستایشی[۲]

## چکیده

**سابقه و هدف:** در طول تاریخ، بیماری‌ها بزرگ‌ترین تهدید برای بشر به‌شمار می‌روند. در این میان بیماری‌های قلبی از توجه بیشتری در مطالعات پزشکی برخوردارند. در سال‌های اخیر، دسته‌بندی و تشخیص امراض قلبی به عنوان یک مبحث کلیدی بررسی شده و تحقیقات بسیاری در زمینه افزایش‌دقت و کاهش‌خطا در این‌گونه تصمیم‌گیری‌ها صورت گرفته است. با ایجاد سیستم‌های هوشمند یادگیر، این سیستم‌ها در پشتیبانی از تصمیمات و کاهش‌خطا نقش بزرگی را ایفا کرده‌اند.

**مواد و روش‌ها:** در این تحقیق، یک مدل ترکیبی ساده از رگرسیون منطقی و شبکه عصبی پرسپترون تک لایه ارایه شده‌است که با چهار قانون یادگیری مختلف (به صورت مجزا) آموزش می‌بیند. این مدل برای بهبود دسته‌بندی و بازشناسی الگوهای بیماری قلب، روی اطلاعات بالینی ۲۷۰ بیمار از کلینیک کلیولند (Cleveland Clinic) (سایت UCI) استفاده شد. این روش از نرمال سازی آماری بهره می‌گیرد و پس از تشخیص داده‌های مخدوش، آموزش شبکه فقط با ۲۰ درصد از داده‌های موجود انجام می‌شود. پیاده‌سازی مدل در نرم‌افزار Matlab صورت گرفته است.

**یافته‌ها:** میانگین خطای مشاهده شده مدل پیشنهادی روی کل مجموعه داده‌ها ۱۱/۱۱ درصد به دست آمد که بهبود قابل ملاحظه‌ای را نسبت به روش‌های مشابه اخیر نشان می‌دهد. همچنین، یافته‌ها نشان داد که مدل پیشنهادی در برخورد با اختلالات موجود در داده‌ها بسیار توانمند عمل می‌کند.

**استنتاج:** تکنیک خطی مطرح شده تأثیر زیادی بر کاهش‌خطا در دسته‌بندی و شناسایی افراد بیمار نسبت به روش‌های متداول و غیر خطی پیچیده با دقتی بیشتر و در زمانی کمتر داشته است. این روش برای تشخیص زود هنگام بیماری و یا به عنوان یک سیستم پشتیبان تصمیم می‌تواند به پزشک یاری رساند.

**واژه‌های کلیدی:** تشخیص بیماری‌های قلبی، مهندسی پزشکی، بازشناسی الگو، یادگیری ماشین، شبکه عصبی مصنوعی، پرسپترون تک لایه، رگرسیون منطقی

## مقدمه

در طول تاریخ حیات بشری، بیماری‌ها بزرگ‌ترین تهدید برای بشر به‌شمار رفته‌اند که تأثیرات مخربی نیز روی زندگی انسان‌ها داشته‌اند. در این میان، بیماری قلب از توجه بیشتری در تحقیقات پزشکی برخوردار بوده است. برطبق بیانیه مرکز آمار و بیولوژی سرطان استرالیا و همچنین بر اساس گزارش دفتر آمار و اطلاعات جمعیتی و مهاجرت ایران، بیماری‌های قلبی یکی از دو علل مهم مرگ و میر در دنیا به شمار می‌روند (۳-۱). در ایران نیز مطابق نمودار شماره ۱، بیشترین تعداد مرگ و میر را بیماری‌های قلبی- عروقی تشکیل داده‌اند.

پژوهش‌های اخیر نشان داده‌اند که بیماری‌های قلبی تحت یک وضعیت خاص اتفاق نمی‌افتند، بلکه شرایط زیادی در به وجود آمدن این گونه بیماری‌ها نقش دارند و همچنان دلایل اساسی بیماری‌های قلبی مبهم می‌باشند (۷-۴).

**مولف مسئول: مجتبی هروی** – تهران: خیابان سمیه، دانشگاه علوم اقتصادی، دانشکده مهندسی کامپیوتر و مدیریت، گروه کامپیوتر E-mail: mojtaba.heravi@ues.ac.ir

۱. کارشناس ارشد مهندسی دانش و علوم تصمیم، گروه کامپیوتر، دانشکده مهندسی کامپیوتر و مدیریت، دانشگاه علوم اقتصادی تهران، تهران، و عضو باشگاه پژوهشگران جوان و نخبگان، دانشگاه آزاد اسلامی، واحد قزوین، ایران

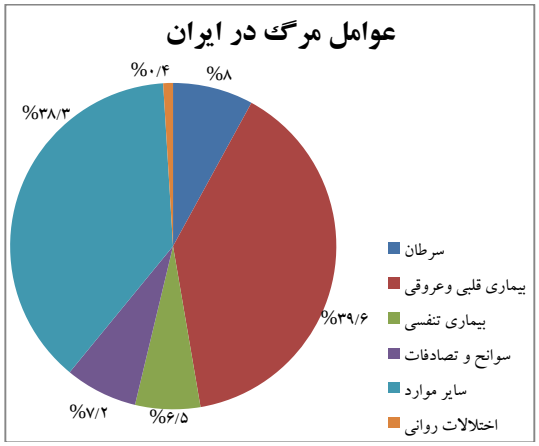
۲. دانشیار، گروه مهندسی پرتو پزشکی، دانشکده مهندسی هسته‌ای و فیزیک، دانشگاه صنعتی امیرکبیر، تهران، ایران

تاریخ دریافت: ۱۳۹۲/۷/۷ تاریخ ارجاع جهت اصلاحات: ۱۳۹۲/۱۱/۹ تاریخ تصویب: ۱۳۹۳/۱/۲۴



[ Downloaded from jmums.mazums.ac.ir on 2026-11-02 ]

در علوم پزشکی، چون شناسایی فرد بیمار از روی علایم بالینی صورت می‌پذیرد (به طور متداول)، در نتیجه نیاز به افراد خبره (پزشک معالج) در این زمینه حایز اهمیت است. همچنین افراد خبره همیشه و همه جا در دسترس نیستند و امکان بروز اشتباه در فرد خبره (هنگام زیاد بودن علایم یا متشابه بودن آن‌ها با انواع دیگر بیماری) بیشتر از یک سیستم یادگیر با حافظه نگهداری پیشینه بیمار است. سرعت و دقت پردازش اطلاعات در یک سیستم یادگیر نیز خیلی بیشتر از یک فرد خبره در شرایط حساس و پرتنش می‌باشد. این موارد به روشنی بر ضرورت وجود ماشینی تأکید می‌کند که بتواند در مواقع حساس -که با کوچک‌ترین اشتباه پزشک ممکن است بیماری به درستی تشخیص داده نشود و جان بیمار به خطر بیفتد- در تصمیم‌گیری به پزشک کمک نماید.




**نمودار شماره ۱: مقایسه عوامل متفاوت منجر به مرگ در ایران (۲۰۱۲)**

در سال‌های اخیر، از مفاهیم هوش محاسباتی ماشین در یافتن روابط بین بیماری‌های مختلف و ویژگی‌های بیماران استفاده شده است (۱۲-۸). به طور مثال Eberhart و همکاران، یک شبکه عصبی تطبیقی را برای دسته‌بندی الگوهای الکتروکاردیوگراف چند کاناله به‌کار برده‌اند (۱۳). Cios و همکاران از یک شبکه عصبی برای یافتن بیماری‌های قلبی توسط تصاویر اکوکاردیوگرافی استفاده کرده‌اند (۱۴).

Palaniappan و Awang یک نمونه عارضه‌یابی کامپیوتری (CAD یا Computer-aided diagnostics) با نام سیستم پیش‌بینی عارضه‌یابی قلب هوشمند (IHDPS یا Intelligent heart disease prediction system) به کمک روش‌های داده کاوی ارایه داده‌اند (۱۵). Avci و Turkoglu یک سیستم هوشمند بر مبنای استفاده از دیدگاه ماشین بردار پشتیبان ژنتیکی (GSVM یا Genetic-support vector machine) برای دسته‌بندی سیگنال‌های داپلر دریچه‌های قلب ارایه داده‌اند (۱۶).

در این مطالعه، با توجه به مهم و شایع بودن بیماری قلب، هدف این بود که با استفاده از یک روش غیر تهاجمی، مدلی از ترکیب رگرسیون منطقی دو مقداره و شبکه عصبی پرسپترون تک لایه به عنوان یک سیستم پشتیبان تصمیم هوشمند یا تصمیم‌گیر مستقل با پاسخگویی سریع و دقت بالا تولید شود که بتواند از روی اطلاعات و علایم بالینی فرد مشکوک به بیماری قلبی، ابتلا یا عدم ابتلای او را تشخیص دهد. همچنین در مقابل روش‌های تهاجمی از خطر کمتری نیز برخوردار باشد.

## مواد و روش‌ها

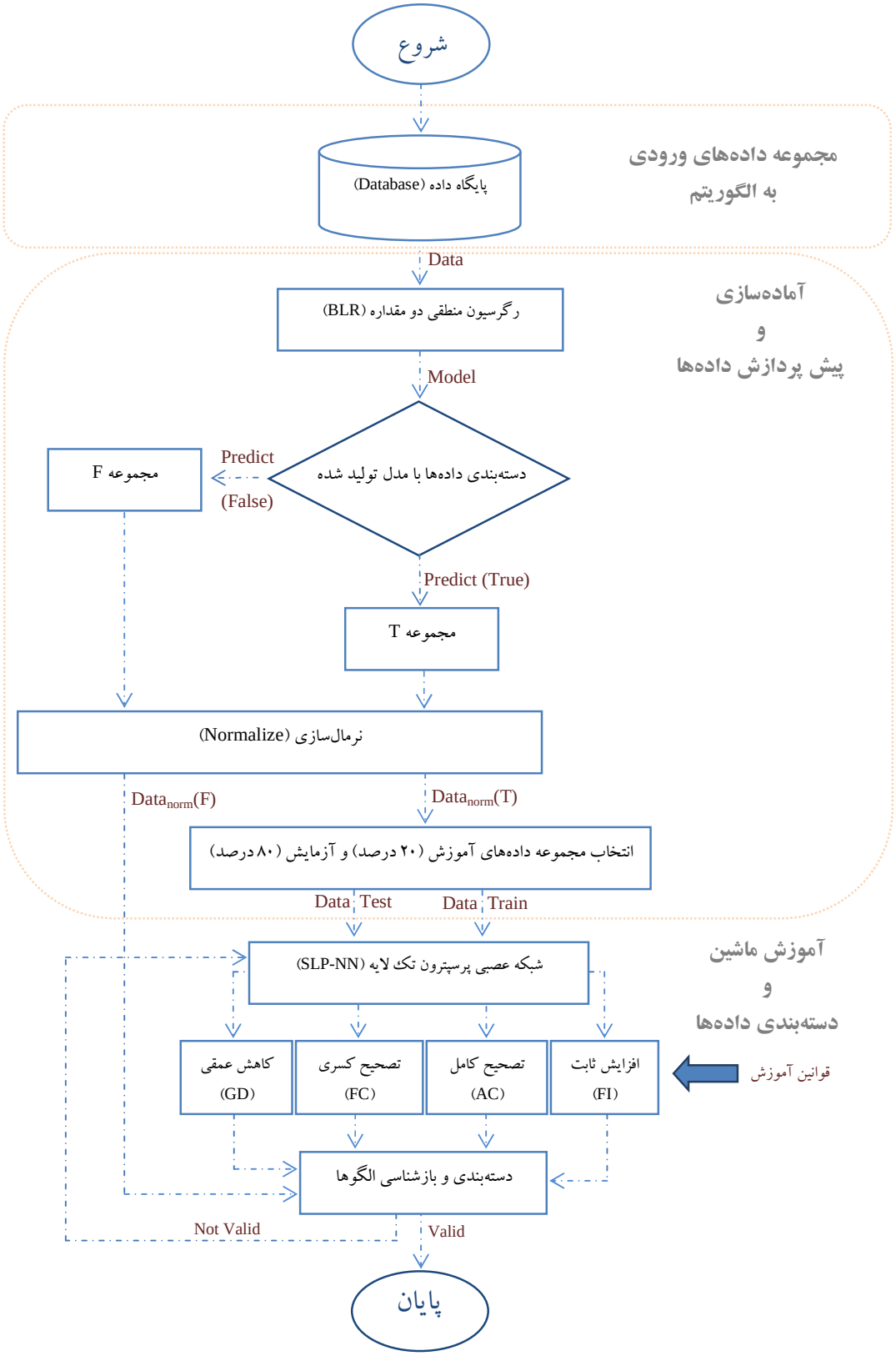
تحقیق حاضر یک مطالعه تشخیصی بود که بر اساس متغیرهای ورودی، به تصمیم‌گیری در مورد وضعیت وجود یا عدم وجود بیماری قلبی پرداخت. الگوریتم پیشنهادی ما از دو فاز اصلی تشکیل شده است: فاز اول، شامل بخش‌هایی برای آماده‌سازی داده‌ها قبل از انجام فرایند آموزش و دسته‌بندی است و فاز دوم، مربوط به بخش آموزش ماشین و دسته‌بندی داده‌ها برای بازشناسی و تفکیک ورودی‌های جدید از هم بر اساس شبکه عصبی می‌باشد. طرح کلی الگوریتم در تصویر شماره ۱ نشان داده شده است.

*۱. فاز پیش پردازش*

*۱-۱. نرمال‌سازی*

امروزه، در بیشتر مسایلی که با آن‌ها مواجه می‌شویم، مقادیر ورودی در مقیاس‌های یکسانی نیست و این مسأله، تأثیر منفی در روند همگرایی و افزایش زمان و تعداد دفعات تکرار اعمال آموزش خواهد گذاشت. پس بایستی داده‌ها را به نحو

[ Downloaded from jmums.mazums.ac.ir on 2026-11-02 ]




**تصویر شماره ۱: طرح الگوریتم پیشنهادی برای دسته‌بندی و بازشناسی الگوها**

[ Downloaded from jmums.mazums.ac.ir on 2026-11-02 ]

مؤثری نرمال کرد و تغییر داد تا این مشکلات برطرف شود. به این منظور، برای نرمالیزه کردن داده‌ها از روش نرمالیزاسیون آماری جهت همگن کردن داده‌ها استفاده شد که از رابطه زیر تبعیت می‌کند:

$$Data_{norm} = \frac{Data - Mean_{Data}}{Std_{Data}} \quad \text{(رابطه ۱)}$$

که در آن Data همان داده‌های خام ورودی و Mean و Std به ترتیب بیانگر میانگین داده‌ها و انحراف معیار داده‌ها در هر خصیصه می‌باشند.

*۱-۲. رگرسیون منطقی دو مقداره*

یکی از کاربردهای مهم رگرسیون، بررسی وابستگی‌های موجود در داده‌ها برای پیش‌بینی و به دست آوردن رفتار یک متغیر (Y)، بر حسب سایر متغیرها (X) می‌باشد (۱۹-۱۷). حال این رفتار می‌تواند خطی یا به شکل منحنی باشد. حالت خاصی از رگرسیون، رگرسیون منطقی دو مقداره است که یک تکنیک غیر خطی با یادگیری نظارتی، ساختاری آماری و بدون هوشمندی برای پیش‌بینی احتمالاتی یک اتفاق می‌باشد. اگر معادله آن به صورت زیر در نظر گرفته شود:

$$Y = \beta.X + \alpha \quad \text{(رابطه ۲)}$$

که در آن $\beta$ ضرایب تأثیر متغیرهای ورودی (X) و $\alpha$ مقدار ثابت (بایاس) می‌باشد، برای پیدا کردن احتمال وقوع هر کدام از کلاس‌ها، از رابطه زیر استفاده می‌گردد:

$$P = \frac{e^{Y}}{1+e^{Y}} \quad \text{(رابطه ۳)}$$

$$\begin{cases} \text{if} \quad ۰ \le P \le ۰/۵ \Rightarrow \text{Class 1} \\ \text{if} \quad ۰/۵ \le P \le ۱ \Rightarrow \text{Class 2} \end{cases}$$

به طوری که اگر $۰ \le P \le ۰/۵$ باشد، الگو متعلق به کلاس اول و اگر $۰/۵ < P \le ۱$ باشد، الگو متعلق به کلاس دوم می‌شود.

*۱-۳. نحوه انتخاب مجموعه داده‌های آموزش و آزمایش و روند یادگیری*

داده‌های آموزش بیشترین تأثیرگذاری را روی یادگیری ماشین دارند. روشی که ما برای تفکیک این داده‌ها به دو دسته آموزش و آزمایش پیشنهاد می‌کنیم، انتخاب ۲۰ به ۸۰ است که از تعریف قانون پارتو گرفته شده است (۲۱، ۲۰).

در این قانون، به ترتیب ۲۰ درصد و ۸۰ درصد جامعه آماری (داده‌ها) را اقلیت حیاتی و اکثریت کم اهمیت تشکیل می‌دهند. به عبارتی، اگر ۲۰ درصد از داده‌های مهم به درستی در آموزش شرکت یابند، امکان پیش‌بینی صحیح ۸۰ درصد از کل داده‌ها میسر خواهد شد. مجموعه دیگری که در روند آموزش استفاده شد، مجموعه اعتبارسنجی بود که در هر مرحله از آموزش، میزان بهبود دقت و میزان کاهش خطا توسط آن بررسی شد. در مدل پیشنهادی برای جلوگیری از زیاده برازش، تدابیری اتخاذ شد که اگر در طول یادگیری میزان خطای مجموعه اعتبارسنجی، ۷ مرتبه به صورت متوالی افزایش پیدا می‌کرد، یادگیری متوقف می‌شد و بیشترین مقدار دقت پیش از توقف، به عنوان پاسخ به خروجی فرستاده می‌شد.

*۲. فاز آموزش با شبکه عصبی پرسپترون تک لایه*

شبکه‌های عصبی به عنوان یک ابزار قدرتمند برای تجزیه و تحلیل پایگاه داده‌ها در مسایلی همچون دسته‌بندی، تشخیص، شناسایی و پشتیبانی از تصمیمات به کار گرفته می‌شوند (۲۴-۲۲). شبکه‌های عصبی توانایی یادگیری، مدل کردن، توسعه و بررسی دسته‌بندی‌های الگوها در یک پدیده بدون نیاز به هیچ دانش قبلی برای درک توزیع داده‌ها را دارا می‌باشند و به راحتی قابلیت انطباق با مسایلی مانند انواع تشخیص پزشکی و پیش‌بینی بیماری در افراد را دارند و علاوه بر سرعت بخشیدن در پاسخ‌گویی، از دقت بالایی نیز برخوردارند (۲۶-۲۳). شبکه‌های متفاوتی برای دسته‌بندی و مدل کردن مسایل به‌کار گرفته شده‌اند (۲۴).

شبکه عصبی مصنوعی پرسپترون تک لایه جزء ساده‌ترین شبکه‌های عصبی خطی محسوب می‌شود (۲۷). معادله این شبکه برای یک الگو در رابطه ۴ بیان شده است. فرض کنید یک مجموعه از الگوهای آموزشی $\{X_1,\ldots,X_j,\ldots,X_N\}$ داریم که هر بردار آن به یکی از دو کلاس $\omega_1$ و $\omega_2$ انتساب داده شده‌اند و N تعداد الگوهای ورودی می‌باشد. هدف پیدا کردن یک بردار وزن (در حالت تک نرونی)، $W = (w_1, \ldots, w_i, \ldots, w_R)$ و مقدار

[ Downloaded from jmums.mazums.ac.ir on 2026-11-02 ]

آستانه‌ای یا ترم بایاس $\theta$ است که R تعداد ورودی‌های هر الگو یا بعد آن می‌باشد $\{X_1, ..., X_i, ..., X_R\}$، همچنین Threshold یک تابع محرک است که در تنظیم مقدار خروجی (f) تأثیر دارد. در حالت کلی، اگر $f > 0$ باشد، الگو مربوط به کلاس $\omega_1$ و اگر مقدار $f > 0$ باشد، الگو در کلاس $\omega_2$ جای می‌گیرد.

$$f = \text{threshold}\left(\sum_{i=1}^{R} w_i . x_i, \theta\right) \qquad \text{(رابطه ۴)}$$

برای آموزش این شبکه، از چهار قانون یادگیری استفاده شد که شامل قوانین یادگیری افزایش ثابت (Fixed increment یا FI)، تصحیح کامل (AC یا Absolute correction)، تصحیح کسری (FC یا Fractional correction) و روش کاهش عمق (GD یا Gradient descent) می‌باشند (۲۸). قانونی که بهترین نتیجه را بدهد، به عنوان خروجی نهایی بازگردانده می‌شود. رابطه‌های این چهار قانون با تغییراتی که در اوزان و مقدار ترم بایاس به وجود می‌آورند، عمل یادگیری را انجام می‌دهند. اگر فرض کنیم:

$$y(j) = (x_1, x_2, ..., x_R, 1) \qquad \text{(رابطه ۵)}$$

$$W = (w_1, w_2, ..., w_R, \Theta) \qquad \text{(رابطه ۶)}$$

$$y(j) \rightarrow -y(j) \text{ for all } y(j) \in \omega_2 \qquad \text{(رابطه ۷)}$$

که y(j) نگاشتی از الگوی $X_j$ می‌باشد، مقدار عدد یک در نگاشت صورت گرفته، مربوط به ترم بایاس اضافه شده به بردار وزن است. آن گاه خواهیم داشت:

$$w(k+1) = \begin{cases} w(k) + cy(k) & if\ w(k).y(k) \le 0 \\ w(k) & if\ w(k).y(k) > 0 \end{cases} \qquad \text{(رابطه ۸)}$$

$$c = [|w(k).y(k)|/\|y(k)\|^2] \qquad \text{(رابطه ۹)}$$

$$c = \lambda \frac{|w(k).y(k)|}{\|y(k)\|^2} \qquad \text{(رابطه ۱۰)}$$

$$w(k+1) = \begin{cases} w(k) - c\nabla E & if\ w(k).y(k) \le 0 \\ w(k) & if\ w(k).y(k) > 0 \end{cases} \qquad \text{(رابطه ۱۱)}$$

رابطه ۸، معادله کلی برای تغییر یا عدم تغییر اوزان و ترم بایاس در سه قانون اول است. با توجه به این مطلب که در قانون افزایش ثابت، نرخ پارامتر یادگیری (c) مقداری بین [۰-۲] دارد. همچنین نرخ پارامتر یادگیری (c) در قانون تصحیح کامل و قانون تصحیح کسری، به ترتیب مقدارشان از رابطه ۹ و رابطه ۱۰ به دست می‌آید. در قانون تصحیح کسری، میزان تأثیر پارامتر یادگیری با متغیر $\lambda$ مشخص می‌شود که مقداری بین [۰-۲] است. رابطه ۱۱ نیز، معادله تغییر یا عدم تغییر اوزان و ترم بایاس در قانون کاهش عمقی است. با ذکر این موضوع که E مقدار میانگین مربع خطا است، $\nabla E$ مقدار گرادیان آن می‌باشد و نرخ پارامتر یادگیری (c) مقداری بین [۰-۲] است.

### ۳. داده‌های زیستی مورد بررسی

داده‌های مورد بررسی، از اطلاعات کلینیک کلیولند (Cleveland Clinic) موجود در سایت UCI جمع‌آوری شد و در حیطه تشخیص بیماری قلب بود (۲۹). این داده‌ها مربوط به ۲۷۰ نمونه بود که ۱۳ خصیصه از اطلاعات بالینی هر نمونه به همراه کلاس آن‌ها را شامل می‌شد (۲۹). این خصیصه‌ها شامل موارد زیر می‌باشند:

- سن
- جنسیت: مرد یا زن
- نوع درد قفسه سینه: شامل چهار وضعیت، آنژین صدری متداول (Angina)، آنژین صدری نابهنجار (Abnang)، درد غیر آنژینی (Notang) و بدون علامت (Asympt).
- فشار خون: در حالت استراحت
- میزان کلسترول خون (mg/dl)
- میزان قند خون ناشتا (mg/dl): بیشتر از ۱۲۰ میلی‌گرم به دسی‌لیتر است یا خیر.
- وضعیت نوار قلب در حالت استراحت: شامل چهار وضعیت، عادی (Normal)، غیر طبیعی (Abnormal)، اختلال در موج ST-T، هیپرتروفی بطن (Hypertrophy).
- حداکثر ضربان قلب
- میزان درد قلبی هنگام فعالیت: ناشی از آنژین
- پارامتر ST: میزان افت موج ST در حالت فعالیت نسبت به حالت استراحت
- شیب ST در اوج فعالیت: شامل سه وضعیت، شیب صعودی، مسطح و شیب نزولی

[ Downloaded from jmums.mazums.ac.ir on 2026-11-02 ]

- تعداد فلوروسکوپی: تعداد فلوروسکوپی های گرفته شده از رگ‌های اصلی به صورت رنگی
- شرایط قلب بیمار: شامل سه وضعیت، عادی، نقص ثابت و نقص برگشت‌پذیر

همچنین این پایگاه داده، افراد را در دو کلاس شامل محتمل به بیماری قلبی و غیر محتمل به بیماری قلبی قرار می‌دهد. هیچ کدام از داده‌هایی مورد بررسی، مقادیر گمشده و یا نامشخص نداشتند؛ یعنی تمامی اطلاعات نمونه‌ها موجود و مشخص بودند.

## یافته‌ها

روند پیاده‌سازی صورت گرفته توسط نرم‌افزار Matlab و یافته‌های به دست آمده به‌شرح زیر می‌باشد.

ابتدا با استفاده از روشی که در بخش ۱-۲ مطرح شد، ضرایب تأثیر و مقدار ثابت بهینه به دست آمد (جدول شماره ۱). فرمول مدل ریاضی تقریبی پدیده نیز در رابطه ۱۲ نوشته شده است. پس از آن، با مدل تولید شده، کل نمونه‌ها دسته‌بندی شدند. برای ارزیابی دقت و راستی آزمایی، نتایج پیش‌بینی مدل (نتایج واقعی) با مشاهدات حاصل از پدیده (نتایج مطلوب) مقایسه شدند (جدول شماره ۲).

$$Y = -۰/۰۱۷ X۱ + ۱/۵۴۹ X۲ \ldots + ۰/۷۱۳ X۱۳ - ۹/۵۵۰ \quad \text{(رابطه ۱۲)}$$

۲۳۰ نمونه از کل نمونه‌ها به درستی و ۴۰ نمونه باقی‌مانده به اشتباه در کلاس دیگر پیش‌بینی شدند (جدول شماره ۲). این دو دسته به ترتیب مجموعه T و مجموعه F نام‌گذاری شدند. این جداسازی داده‌ها به دو دسته F و T به این خاطر صورت پذیرفت که داده‌هایی که از حساسیت یا اختلال بیشتری نسبت به مابقی داده‌ها برخوردار هستند و در فرایند یادگیری امکان تولید خطا را افزایش می‌دهند، از چرخه آموزش کنار گذاشته شوند.

با رابطه ۱، داده‌های دو دسته نرمال شدند و سپس دسته T بر اساس مفاهیم بخش ۱-۳ به صورت تصادفی به دو دسته آموزش و آزمایش برای ورود به شبکه مطرح شده در بخش ۲، تفکیک گردیدند. در جدول شماره ۳ اطلاعات کلی مربوط به پارامترهای ورودی و نتایج خروجی به دست آمده بعد از اتمام آموزش آمده است. همچنین در نمودارهای شماره ۲، ۳، ۴ و ۵ روند بهبود دقت در هنگام آموزش با چهار روش یادگیری مطرح شده، ترسیم شده است.

**جدول شماره ۱: مقدار ضرایب تأثیر و ترم بایاس بهینه (β)**

| متغیر (ویژگی)ها | B |
|---|---|
| سن | -۰/۰۱۷ |
| جنسیت | ۱/۵۴۹ |
| نوع درد قفسه سینه | ۰/۷۰۱ |
| فشار خون | ۰/۰۲۵ |
| میزان کلسترول (mg/dl) | ۰/۰۰۷ |
| میزان قند خون ناشتا (mg/dl) | -۰/۷۶۶ |
| وضعیت نوار قلب در حالت استراحت | ۰/۳۰۴ |
| حداکثر ضربان قلب | -۰/۰۲۲ |
| میزان درد قلبی هنگام فعالیت | ۰/۸۱۶ |
| پارامتر ST | ۰/۳۴۴ |
| شیب ST در اوج فعالیت | ۰/۴۵۱ |
| تعداد فلوروسکوپی | ۱/۱۶۱ |
| شرایط قلب بیمار | ۰/۷۱۳ |
| مقدار ثابت (ترم بایاس) | -۹/۵۵۰ |

**جدول شماره ۲: مقایسه نتایج واقعی مدل با نتایج مطلوب موجود**

| مشاهدات | | پیش‌بینی* | | |
|---|---|---|---|---|
| | | کلاس | | درست (درصد) |
| | | ۱ | ۲ | |
| کلاس | ۱ | ۱۳۳ | ۱۷ | ۸۸/۶۶ |
| | ۲ | ۲۳ | ۹۷ | ۸۰/۸۳ |
| درصد کل | | | | ۸۵/۱۸ |

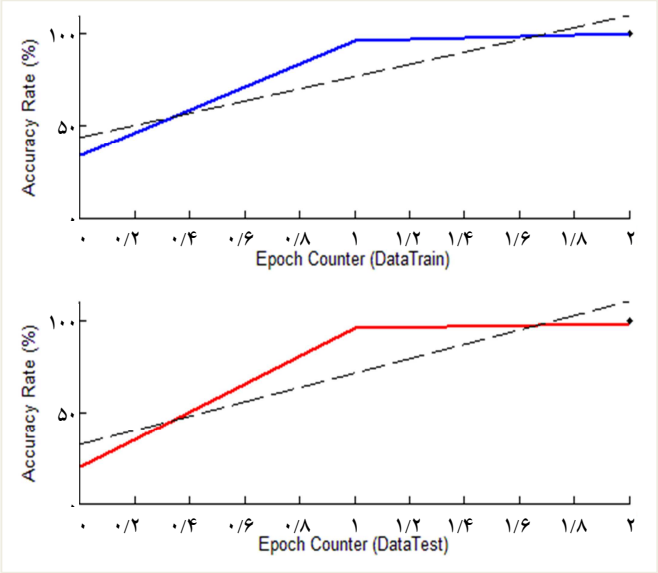
* برش ارزشی برابر ۰/۵ می‌باشد.

مشاهده می‌شود که مقدار خروجی یادگیری با افزایش ثابت نتیجه بهتری را تولید می‌کند و از شناسایی خوبی برخوردار است. در مدل پیشنهادی، چون مقادیر اولیه اوزان و ترم بایاس به صورت تصادفی انتخاب می‌شوند، امکان این که مقادیر خروجی و میزان دقت نتایج در هر اجرا کمی تغییر کنند، وجود دارد. با این حال، میانگین به دست آمده از نتایج چندین بار اجرای برنامه روی هر یک از ترکیب‌های مختلف داده‌ها (شامل ۱۰ ترکیب مختلف)، در مجموعه آموزش و آزمایش نشان داده شد.

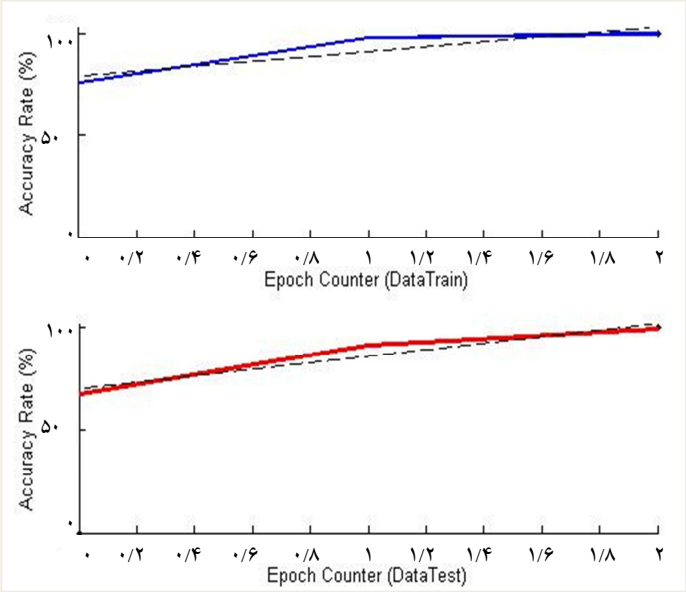
[ Downloaded from jmums.mazums.ac.ir on 2026-11-02 ]

**جدول شماره ۳: نتایج به دست آمده بعد از اتمام آموزش**

| روش یادگیری (Rule) | نرخ پارامتر ورودی | تعداد تکرار (Epoch) | دقت بازشناسی نمونه (درصد) (Recognition accuracy) | | | تعداد شناسایی | |
|---|---|---|---|---|---|---|---|
| | | | آموزش | آزمایش | مجموعه F | درست | غلط |
| Fixed Inc. | ۱ | ۳ | ۱۰۰ | ۹۸/۳۳ | ۳۲/۵ | ۲۴۰ | ۳۰ |
| AbsoluteCor. | – | ۳ | ۱۰۰ | ۹۸/۸۸ | ۱۲/۵ | ۲۳۳ | ۳۷ |
| Fractional Cor. | ۲ | ۳ | ۱۰۰ | ۹۷/۲۲ | ۲۷/۵ | ۲۳۶ | ۳۴ |
| Gradient Des. | ۱ | ۴ | ۱۰۰ | ۹۵/۰۰ | ۲۷/۰۰ | ۲۳۲ | ۳۸ |

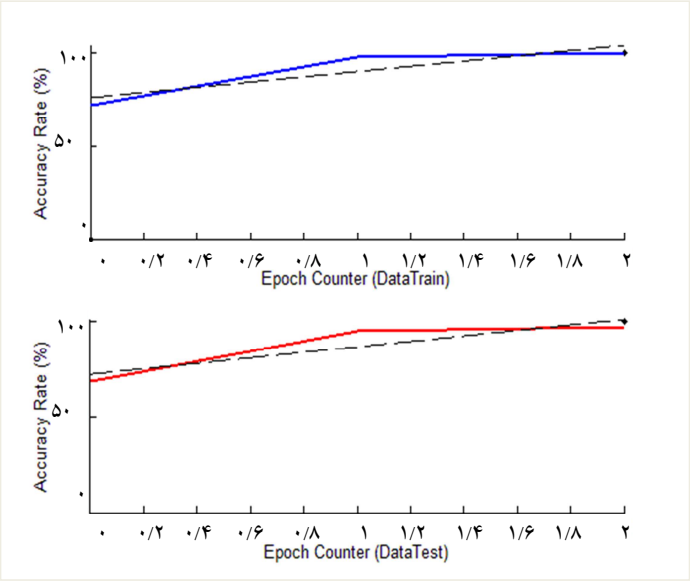



**نمودار شماره ۲: روند افزایش دقت در قانون افزایش ثابت**

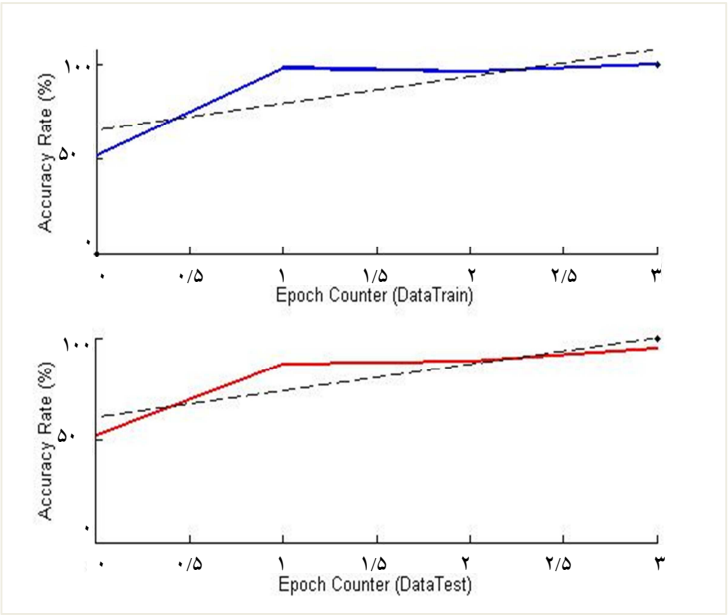



**نمودار شماره ۳: روند افزایش دقت در قانون تصحیح کامل**

[ Downloaded from jmums.mazums.ac.ir on 2026-11-02 ]




**نمودار شماره ٤: روند افزایش دقت در قانون تصحیح کسری**




**نمودار شماره ٥: روند افزایش دقت در قانون کاهش عمقی**

## بحث

در بیشتر کارهایی که در زمینه پزشکی صورت گرفته است، از الگوریتم‌های غیر خطی و پیچیده استفاده شده است و این مسأله باعث افزایش زمان پاسخگویی سیستم می‌شود. همچنین یادگیری از روی تعداد زیادی الگو به عنوان مجموعه آموزش (بیش از ۷۰ درصد) در این الگوریتم‌ها باعث شده است که علاوه بر صرف زمان بیشتر، هوشمندی سیستم نیز زیر سؤال برود. مدل پیشنهادی، یک الگوریتم خطی و ساده با دقت و سرعت تحلیل بالاتر است که تنها با ۲۰ درصد از داده‌ها قابل آموزش و اجرا است. بین این روش با روش‌های مطرح دیگر در جدول شماره ۴ مقایسه‌ای صورت گرفته است که نتایج آن ادعای مدل پیشنهاد شده را ثابت کرده است.

[ Downloaded from jmums.mazums.ac.ir on 2026-11-02 ]

جدول شماره ٤: مقایسه نتایج الگوریتم‌های مختلف ارایه شده

| توضیحات | بیشترین دقت (درصد) | میانگین دقت (درصد) | میانگین خطا (درصد) | در سال | الگوریتم یادگیری عنوان کوتاه | عنوان کامل |
|---|---|---|---|---|---|---|
| تعداد قانون تولید شده برابر ۳۱ | ۸۷/۹۱ | ۸۳/۷۴ | ۱۶/۲۶ | ۲۰۰۶ | HNFB-1 (۳۰) | Inverted hierarchical neuro-fuzzy binary space partitioning |
| - | ۸۷/۴ | ۸۶/۸ | ۱۳/۲ | ۲۰۰۸ | ANN-FNN (۳۱) | artificial neural network and fuzzy neural network |
| با معیار همبستگی و اندازه جمعیت برابر ۱۶ | ۸۵/۸۱ | ۸۵/۱۵ | ۱۴/۸۵ | ۲۰۰۹ | GA-SVM (۳۲) | genetic algorithms and support vector machines |
| - | ۹۱/۸۶ | ۸۷/۶۳ | ۱۲/۳۷ | ۲۰۱۰ | PCA-DE (۳۳) | P rincipal component analysis and differential evolution classifier |
| در الگوریتم شبکه عصبی پس انتشار | ۹۰/۳ | ۸۷/۸۸ | ۱۲/۱۲ | ۲۰۱۰ | GDM/AG (۳۴) | gradient descent with momentum and adaptive gain |
| شامل ۳۰ الگوریتم یادگیری ماشین | ۸۰/۴۹ | ۸۰/۴۹ | ۱۹/۵۱ | ۲۰۱۱ | RF ensemble (۳۵) | Rotation forest ensemble classifier |
| - | ۹۴/۸۲ | ۸۵/۴۸ | ۱۴/۵۲ | ۲۰۱۲ | APPM-ANN (۳۶) | Artificial photosynthesis and phototropism mechanism and artificial neural network |
| با تغییر روی تعداد اعضای کلاس ها | ۹۱/۰۵ | ۸۵/۵۲ | ۱۴/۴۸ | ۲۰۱۳ | MFS-CFS (۳۷) | motivated feature selection and computer feature selection |
| روش پیشنهادی | ۹۰/۰۰ | ۸۸/۸۹ | ۱۱/۱۱ | - | BLR-SLPNN | binary logistic regression and single layer perceptron neural network |

در واقع، هدف اصلی تولید این مدل این بود که با توسعه و استفاده همگانی از این مدل، هر فردی بتواند تنها با داشتن اطلاعات بالینی و آزمایشگاهی خود، بدون نیاز به پزشک، یک تشخیص اولیه به نسبت دقیق در مورد بروز بیماری قلبی در خود پیش از بحرانی شدن شرایط داشته باشد و سپس برای تشخیص دقیق‌تر به پزشک متخصص مراجعه کند. حتی یک پزشک برای حصول اطمینان از تشخیص خود (بیماری قلبی) نیز می‌تواند از این مدل بهره ببرد.

## References


1. Australian Bureau of Statistics. Causes of Death, Australia [Online]. [cited 2009]; Available from: URL: http://abs.gov.au/ausstats/abs@.nsf/Products/696C1CF9601E4D8DCA25788400127BF0?
2. Benjamin King RJ, Robins MW. Cancer Biology. New York, NY: Pearson/Prentice Hall; 2006.
3. Shabestan News Agency. Report of death rates about 2012 year, have a record of heart disease [Online]. [cited 2012 Aug 11]; Available from: URL: http://www.shabestan.ir/NSite/FullStory/News/?Serv=0&Id=161273&Mode
4. Chilnick LD. Heart Disease: an essential guide for the newly diagnosed. Boston, MA: Da Capo Press; 2008.
5. Health Central. Heart Disease. Definition [Online]. [cited 2010]; Available from: URL: http://www.healthcentral.com/encyclopedia/408/458.html#Definition of Heart Disease
6. King L. Taking on Heart Disease. New York, NY: Rodale; 2004.
7. Silverstein A, Silverstein Nunn L. Heart Disease. Washington, DC: Twenty-First Century Books; 2006.
8. Huang Z, Li J, Su H, Watts GS, Chen H. Large-scale regulatory network analysis from microarray data: modified Bayesian network learning and association rule mining. Decision Support Systems 2007; 43(4): 1207-25.
9. Ishibuchi H, Kuwajima I, Nojima Y. Prescreening of candidate rules using association rule mining and pareto-optimality in genetic rule selection. Lecture Notes in Computer Science 2007; 4693: 509-16.
10. Karabatak M, Cevdet Ince M. An expert system for detection of breast cancer based on association


[ Downloaded from jmums.mazums.ac.ir on 2026-11-02 ]

rules and neural network. Expert Systems with Applications 2009; 36(2): 3465-9.
11. Mi Shin A, Hee Lee I, Ho Lee G, Joon Park H, Seop Park H, Yoon K, et al. Diagnostic analysis of patients with essential hypertension using association rule mining. Healthc Inform Res 2010; 16(2): 77-81.
12. Wang Z, Hoy WE. Is the Framingham coronary heart disease absolute risk function applicable to Aboriginal people? Med J Aust 2005; 182(2): 66-9.
13. Eberhart RC, Dobbins RW, Webber WRS. CaseNet: a neural network tool for EEG waveform classification. Minneapolis, MN: IEEE Xplore; 1989. p. 60-8.
14. Cios KJ, Chen K, Langenderfer RA. Use of neural networks in detecting cardiac diseases from echocardiographic images. IEEE Eng Med Biol Mag 1990; 9(3): 58-60.
15. Palaniappan S, Awang R. Intelligent heart disease prediction system using data mining techniques. Washington, DC: IEEE Computer Society; 2008. p. 108-15.
16. Avci E, Turkoglu I. An intelligent diagnosis system based on principle component analysis and ANFIS for the heart valve diseases. Expert Systems with Applications 2009; 36(2): 2873-8.
17. Inoue T, Taguri M, Hoshi M. Prediction of nitrogen oxide concentration by a regression model. Atmospheric Environment 1986; 20(12): 2325-37.
18. Inoue T, Hoshi M, Taguri M. Stability of two prediction schemes for hourly nitrogen oxide concentrations by a regression model. Atmospheric Environment 1987; 21(4): 920-42.
19. Andrés Sánchez J. Calculating insurance claim reserves with fuzzy regression. Fuzzy Sets and Systems 2006; 157(23): 3091-108.
20. Juran JM. Universals in Management Planning and Controlling. TPOK/Juran Institute 1954; 1: 1-13.
21. Pareto V, Bongiorno A, Livingston A, Rogers JH. Trattato di sociologia generale. London, UK: Jonathan Cape Ltd; 1935.
22. Horn W. Artificial Intelligence in Medicine. New York, NY: Springer; 1999. p. 201-4.
23. Zhang GP. Neural networks for classification: a survey. IEEE Systems, Man, and Cybernetics Society 2000; 30(4): 451-62.
24. Bishop CH. Neural Networks for Pattern Recognition. New York, NY: Oxford University Press, Inc; 1995.
25. Baxt WG. Application of artificial neural networks to clinical medicine. Lancet 1995; 346(8983): 1135-8.
26. Haykin S. Neural Networks: A comprehensive foundation. 2nd ed. New Jersey, NJ: Prentice Hall; 1999.
27. Webb AR. Statistical pattern recognition. 2nd ed. Hoboken, NJ: Wiley; 2002.
28. Nilsson NJ. The mathematical foundations of learning machines. San Francisco, CA: Morgan Kaufmann Publishers; 1965.
29. Center for Machine Learning and Intelligent Systems. Cleveland heart disease data details [Online].[cited 2010 Feb]; Available from: URL: http://archive.ics.uci.edu/ml/machine-learning-databases/heart-disease/heartdisease. names
30. Goncalves LB, Vellasco MMBR, Pacheco MAC, de Souza FJ. Inverted hierarchical neuro-fuzzy BSP system: a novel neuro-fuzzy model for pattern classification and rule extraction in databases. IEEE Systems, Man, and Cybernetics Society 2006; 36(2): 236-48.
31. Kahramanli H, Allahverdi N. Design of a hybrid system for the diabetes and heart diseases. Expert Systems with Applications 2008; 35(1-2): 82-9.
32. Tan KC, Teoh EJ, Yu Q, Goh KC. A hybrid evolutionary algorithm for attribute selection in data mining. Expert Systems with Applications 2009; 36(4): 8616-30.
33. Luukka P, Lampinen J. A Classification method based on principal component analysis and differential evolution algorithm applied for prediction diagnosis from clinical EMR heart data sets. Computational Intelligence in Optimization 2010; 7: 236-83.
34. Nawi NM, Ghazali R, Mohd Salleh MN. The development of improved back-propagation neural networks algorithm for predicting patients with heart disease. Information Computing and Applications 2010; 6377: 317-24.
35. Ozcift A, Gulten A. Classifier ensemble construction with rotation forest to improve medical diagnosis performance of machine learning algorithms. Comput Methods Programs Biomed 2011; 104(3): 443-51.
36. Cui Z, Yang C, Sanyal S. Training artificial neural networks using APPM. International Journal of Wireless and Mobile Computing 2012; 5: 168-74.
37. Nahar J, Imam T, Tickle KS, Phoebe Chen Y. Computational intelligence for heart disease diagnosis: A medical knowledge driven approach. Expert Systems with Applications 2013; 40(1): 96-104.

[ Downloaded from jmums.mazums.ac.ir on 2026-11-02 ]