نوین در فصلنامه رویکردهای پژوهشی
مدیریت و حسابداری
سال پنجم، شماره ۵۴، بهار ۱۴۰۰- جلد سوم، ص ۴۷-۳۸

Journal of New Research Approaches
in Management and Accounting
ISSN: 2588-4573
http://www.majournal.ir

# استفاده از هوش مصنوعی جهت پیش بینی و تشخیص بیماری فشار خون بالا

**فرشید صهبا** [۱]

**سیده فاطمه منتجبی ها** [۲]

**مهسا ناصر خاکی** [۳]

تاریخ دریافت: ۱۴۰۰/۰۱/۰۹ تاریخ چاپ: ۱۴۰۰/۰۱/۲۵

## چکیده

امروزه بیماری قلبی بسیار معمول و یکی از دلایل اصلی مرگ و میر به شمار می رود. به همین علت تشخیص درست و به موقع این بیماری بسیار مهم است. روش های تشخیصی و درمانی این بیماری عوارض جانبی زیاد و پرهزینه ای دارد؛ بنابراین محققان به دنبال روش های ارزان و با دقت بالا برای تشخیص این بیماری هستند. هوش مصنوعی امروزه در بسیار از سیستم ها مورد استفاده قرار می گیرد که مهمترین آن سیستم ها سلامت است. یکی از بیماری ها قلبی که هرساله جان بسیار از افراد بر اثر آن از دست می رود بیماری فشار خون است که در این تحقیق با استفاده از هوش مصنوعی به دنبال پیش بینی و تشخیص این بیماری هستیم.

## کلمات کلیدی

فشار خون، هوش مصنوعی، پیش بینی

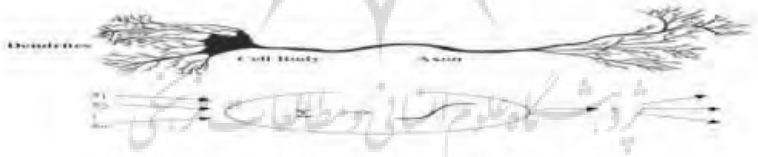
۱. استادیار، گروه مهندسی فناوری اطلاعات، دانشگاه غیاث الدین جمشید کاشانی، آبیک، ایران.
۲. دانشجو کارشناسی ارشد، گروه مهندسی فناوری اطلاعات، دانشگاه غیاث الدین جمشید کاشانی، آبیک، ایران.
۳. دانشجو کارشناسی ارشد، گروه مهندسی فناوری اطلاعات، دانشگاه غیاث الدین جمشید کاشانی، آبیک، ایران.

**مقدمه**

به تازگی، نیاز به شناخت موثر از اطلاعات غیر بدیهی در داده های متنی و مهم در تصمیم گیری، از یک گروه بسیار عظیم از داده ها به طور مداوم در حال افزایش است. ضرورت فرآیند کشف دانش، تبدیل داده ها به دانش به منظور کمک به تصمیم گیری است و به داده کاوی نسبت داده شده است.

اصطلاح بیماری های قلبی شامل بیماری های گوناگونی است که بر قلب تأثیر میگذارد. اصطلاح بیماری های قلبی عروقی، شامل طیف گسترده ای از شرایطی است که بر قلب و رگ های خونی و طریقه پمپ شدن و جریان خون در بدن اثر می گذارد (ازبکار و همکاران،۲۰۱۰). در این تحقیق منظور از بیماری قلبی بیماری فشار خون می باشد. الگوریتم k – means یکی از ابزار شناخته شده خوشه بندی است. داده ها را مطابق با مقادیر مشخصه شان به K دسته مجزا گروه بندی می کند. باید از قبل مقدار عدد صحیح و مثبت که نشان دهنده تعداد خوشه هاست، مشخص شود.

شبکه های عصبی از سیستم یادگیری پیچیده ی مغز که متشکل از مجموعه نرون ها می باشد الهام گرفته شده است. هدف از شبکه های عصبی مصنوعی ارائه روشهائی جهت استفاده از سخت افزارها (مدارات) و نرم افزارها (الگوریتم ها) برای ایجاد قابلیت های هوشمند به دستگاه ها، روبوت ها، برنامه ها و غیره می باشد که قادر به یادگیری حین فرآیند هستند. در شکل ۱ شباهت نرون مصنوعی با نرون های بیولوژیکی نمایش داده شده است. نرون مصنوعی کوچکترین واحد پردازشگر اطلاعات است ورودیهای نرون توسط یک نوع ارتباط به نام وزنبه نرون وارد میشوند



شکل۱- شباهت نرون مصنوعی با نرونهای بیولوژیکی[۲۹]

## فرآیند روش پیشنهادی

روند انجام پروژه برطبق متدولوژی داده کاوی CRISP - DM می باشد که بر روی مجموعه داده ها اعمال می شود. در ابتدا پالایش و تصحیح و یک دست کردن داده ها صورت می گیرد. پاکسازی و فیلتر کردن داده ها ممکن است برای جلوگیری از ایجاد قوانین و الگوهای فریبنده یا نامناسب انجام شود. در این پروژه، مجموعه داده بیماری های قلبی با حذف رکوردهای تکراری و تهیه مقادیر مفقود پاکسازی شده و علاوه بر آن نیز به شکل مناسب برای خوشه بندی تبدیل شده است (ساد و همکاران ۲۰۰۶)، سپس با استفاده از مدل feature selection برای شناسایی ویژگی های مهم در پیش بینی یک نتیجه

خاصئ استفاده می شود. در مرحله بعد با استفاده از الگوریتم های خوشه بندی با استفاده از الگوریتم K- Means دادهها خوشه بندی میشوند. سپس با استفاده از خوشه بندی با استفاده از مدل Kohonen نوعی از شبکه های عصبی را تولید می کند که برای خوشه بندی مجموعه داده ها در گروههای مجزا به کار می رود. پس از آن با استفاده از شبکه عصبی که در این پروژه روش پیشنهاد شده از یک پرسپترون چند لایه با الگوریتم پس انتشار برای آموزش بهره می گیرد. سپس با استفاده از مدل R tree & C که به ما این اجازه را می دهد که ملاحظات آینده را پیشگویی و دسته بندی کنیم و در پایان نود یک الگوریتم c5.0 را به کارگرفتیم تا بتواند یک درخت تصمیم و یا مجموعه ای از قوانین را بنا کند که با استفاده از ارزیابی نتایج آنها به یک نتیجه قابلاستفاده در مراکز پزشکی دست میابیم.

## داده های مورد نیاز

مجموعه داده مورد استفاده در این پروژه متعلق به مرجع (۳) میباشد. پارامترها و محدوده آنهادر زیر آورده شده است:

۱ – age: سن بیماران

۲ – sex: جنسیت (۱ = مرد، ۰ = زن)

۳– CP: نوع درد قفسه سینه

۴– trestbp: فشار خون در زمان استراحت

۵ – chol: کلسترول موجود در سرم خون

۶– FBS: قند خون ناشتا > ۱۲۰

۷ restecg: نتایج نوار قلب در وضعیت استراحت مقدار ۰: عادی – مقدار ۱: داشتن اختلال در موج ST یا T – مقدار ۲: نشان های بزرگ شدگی بطن چپ

۸ – thalach: بیشترین میزان ضربان قلب

۹– exang: آنژین یا درد قفسه سینه ناشی از ورزش

۱۰ – : oldpeak: پایین آمدن موج ST در ورزش نسبت به حالت استراحت

۱۱ – slope: شیب پیک بخش ST(مقدار ۱: شیب صعودی – مقدار ۲: بدون شیب – مقدار ۳: شیب نزولی)

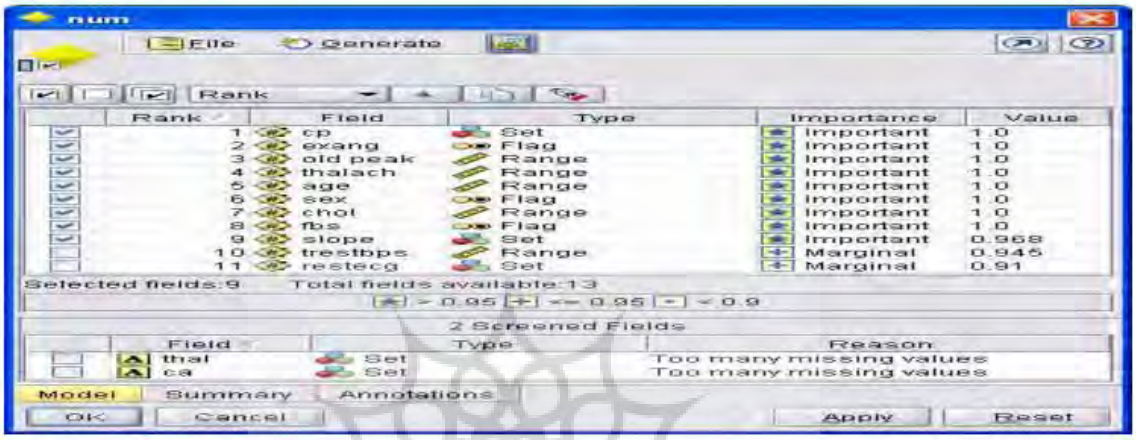
۱۲ –ca: تعداد رگهای اصلی (۰–۳) که با فلوروسکپی رنگ آمیزی شده اند

۱۳– thal: = طبیعی، ۶ = آسیب تثبیت شده، ۷ = آسیب قابل برگشت

۱۴ :um-مقدار. تقطر رگ کمتر از ۵۰ یعنی عدم تشخیص بیماری های قلبی وضعیت آنژیوگرافی ک بیماری صفت پیش بینی حمله قلبی– مقدار ۱: قطر رگ بیشتر از ۵۰٪ یعنی خطر حمله قلبی.

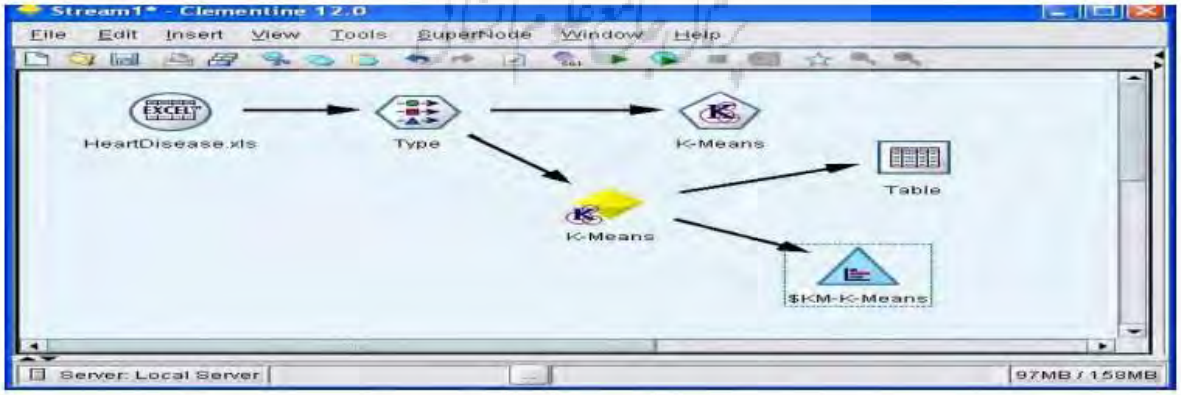
### مدل feature selection

قلبی، ویژگی های مدل انتخاب ویژگی برای شناسایی ویژگی های مهم در پیش بینی یک نتیجه خاص استفاده می شود. با اجرای این مدل بر داده های پیش پردازش شده بیماری thal و ca به دلیل وجو تعداد زی د ادی مقادیر گمشده و null از مجموعه داده حذف می شوند و در اجرای مدل ها شرکت نمی کنند. نتیجه اجرای مدل انتخاب ویژگی در شکل ۲ نشان




شکل۲-نتیجه اجرای مدلfeature selection

خوشه بندی با استفاده از الگوریتم K- Means الگوریتم k - means یکی از ابزار شناخته شده خوشه بندی است که k - means داده ها را مطابق با مقادیر مشخصه شان به دسته مجزا گروه بندی می کند. داده هایی که ویژگی هایی با مقادیر یکسان دارند در خوشه های یکسان دسته بندی می شوند (فرانک و همکاران،۲۰۰۳). در نرم افزار کلمنتاین مدل K - meansبرای خوشه بندی داده ها به صورت شکل ۳ به کار می رود.




شکل۳-مدلK-means درکلمنتاین

نتیجه اجرای مدل k - means نمایش داده شده است؛ و همانطور که مشخص است شماره ۳۹۸ بیمار یعنی ۷۲/۶۴ ٪ در خوشه ۱ و ۲۱۷ بیمار یعنی ۳۵/۲۸ ٪ در خوشه شماره ۲ قرار گرفته اند.

شکل ۴-نتیجه اجرای مدلK-means

## خوشه بندی با استفاده از مدل Ko honen

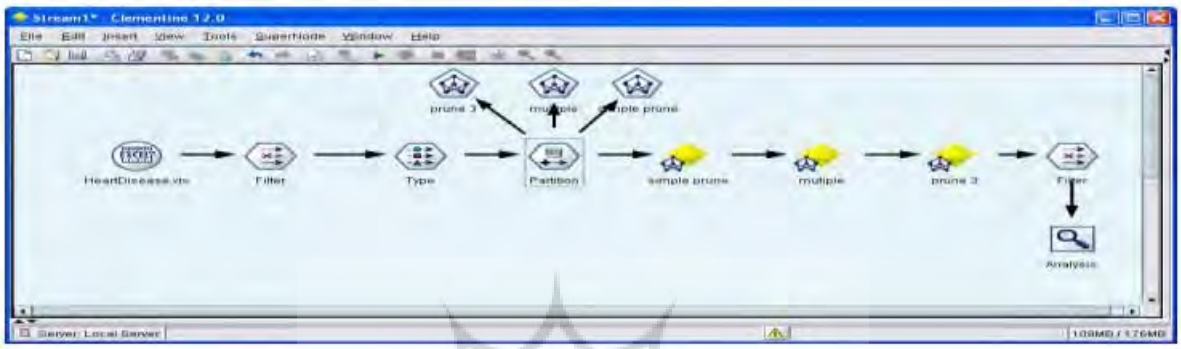
این گره نوعی از شبکه های عصبی را تولید می کند که برای خوشه بندی مجموعه داده ها در گروههای مجزا به کار می رود. وقتی شبکه به طور کامل ساخته شد رکوردهایی که شبیه هم هستند باید در نقشه خروجی نزدیک بهم ظاهر شوند. نتایج اجرای مدل Kohonen در زیر نشان داده شده است که مشخص می کند در هر مختصات X, Y چند رکورد شبیه هم وجوددارد. چند سطر ابتدای جدول خروجی مدل Kohonen شامل مختصات x, y برای هر سطر از پایگاه داده بیماری قلبی در جدول ۱ نشان داده شده است.

جدول ۱- بخشی ازخروجی مدلKohonen

| Age | Sex | Cp | Trestbps | chol | fbs | Res tecg | thalach | Exa ng | old peak | Slo pe | Num | Partition | KX$ Kohon en | KY$ Koho nen |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 40.0 | 1.0 | 2.0 | 140.0 | 289.0 | 0.0 | 0.0 | 172.0 | 0.0 | 0.0 | | 0.0 | 1_Training | 0 | 2 |
| 49.0 | 0.0 | 3.0 | 160.0 | 180.0 | 0.0 | 0.0 | 156.0 | 0.0 | 1.0 | 2.0 | 1.0 | 1_Training | 0 | 1 |
| 37.0 | 1.0 | 2.0 | 130.0 | 283.0 | 0.0 | 1.0 | 98.0 | 0.0 | 0.0 | | 0.0 | 1_Training | 3 | 0 |
| 48.0 | 0.0 | 4.0 | 138.0 | 214.0 | 0.0 | 0.0 | 108.0 | 1.0 | 1.5 | 2.0 | 1.0 | 2_Testing | 0 | 1 |
| 54.0 | 1.0 | 3.0 | 150.0 | | 0.0 | 0.0 | 122.0 | 0.0 | 0.0 | | 0.0 | 1_Training | 1 | 2 |
| 39.0 | 1.0 | 3.0 | 120.0 | 339.0 | 0.0 | 0.0 | 170.0 | 0.0 | 0.0 | | 0.0 | 1_Training | 1 | 2 |
| 45.0 | 0.0 | 2.0 | 130.0 | 237.0 | 0.0 | 0.0 | 170.0 | 0.0 | 0.0 | | 0.0 | 1_Training | 0 | 2 |
| 54.0 | 1.0 | 2.0 | 110.0 | 208.0 | 0.0 | 0.0 | 142.0 | 0.0 | 0.0 | | 0.0 | 1_Training | 0 | 2 |

## پیاده سازی مدل شبکه عصبی

شبکه های عصبی، یکی از قدرتمندترین توابع پیشبینی کنند محسوب میشوند. نرم افزار کلمنتاین با ترکیب کردن چندین خصوصیات سعی میکند تا شبکه های عصبی را از قرار گرفتن در دام ها و تله ها، نجات دهد. در مدل ارائه شده به بررسی شبکه عصبی prune، شبکه عصبی Multiple و شبکه عصبی Exhantive prune می پردازیم. اجرای مدل ها در کلمنتاین در شکل ۵ نمایش داده شده است

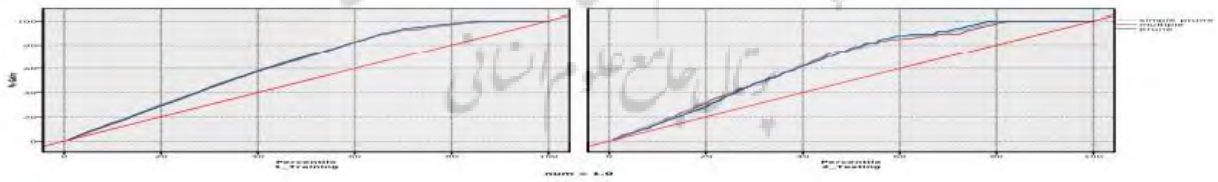



شکل ۵-اجرای سه مدل شبکه عصبی

مدل prune با ۱۲ نرون در لایه ورودی وا نرون در لایه خروجی و با تعداد ۱ لایه پنهان و ۲ نرون در آن اجرا شده است.

مدل Multple با ۱۸ نرون در لایه ورودی وا نرون در لایه خروجی و با تعداد ۱ لایه پنهان و ۷ نرون در آن اجرا شده است.

و مدل Exhantive prune با ۱۸ نرون در لایه ورودی و ۱ نرون در لایه خروجی و با تعداد ۲ لایه پنهان و ۳۰ نرون در لایه پنهان اول و ۲۰ نرون در لایه پنهان دوم اجرا شده است. نمودار Evaluation برای ارزیابی هرسه مدل شبکه عصبی اجرا شده در شکل ۶ رسم شده است.



شکل ۶-نمودار Evaluation

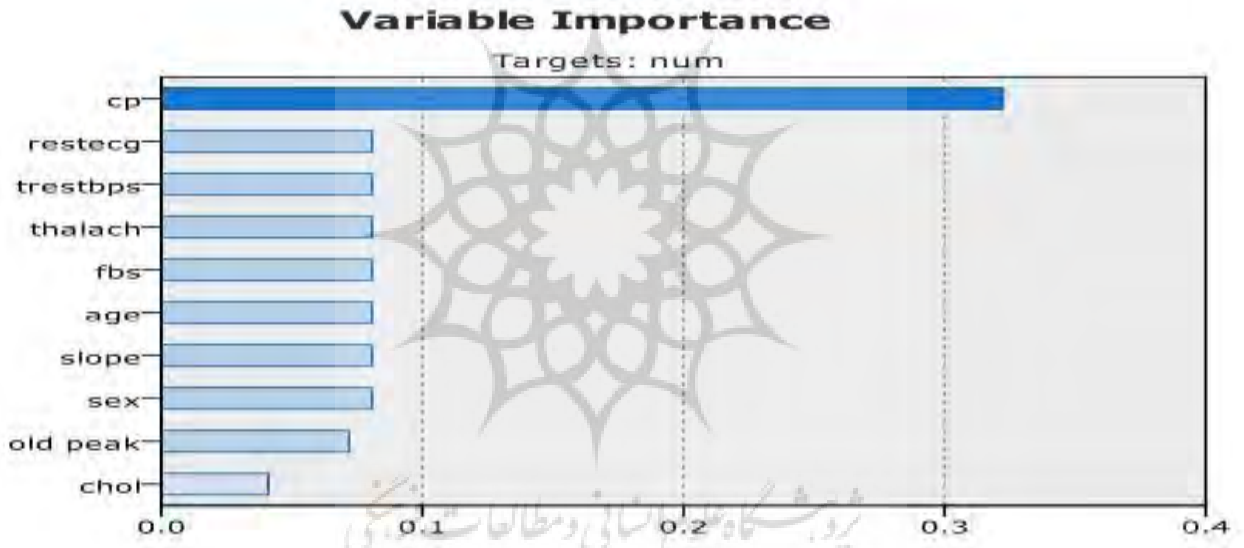
## پیاده سازی مدل R tree & C

R tree & C یک روش بر پایه قسمت بندی وجزء بندی بازگشتی است تا رکوردهای آموزشی را به داخل قطعه ها و اجزاء بخش بندی با کم کردن و حرص کردن بخشهای ناخالص و غیر مهم در هر مرحله، دسته بندی و تقسیم بندی کنیم. در این

قسمت سه مدل C & R tree با تعداد سطوح ۱، ۳ و ۵ لایه پیاده سازی شده است. این مدل ها در نرم افزار کلمنتاین به صورت شکل ۷ اجرا می شوند.

شکل ۷-مدل c&R tree در کلمنتاین

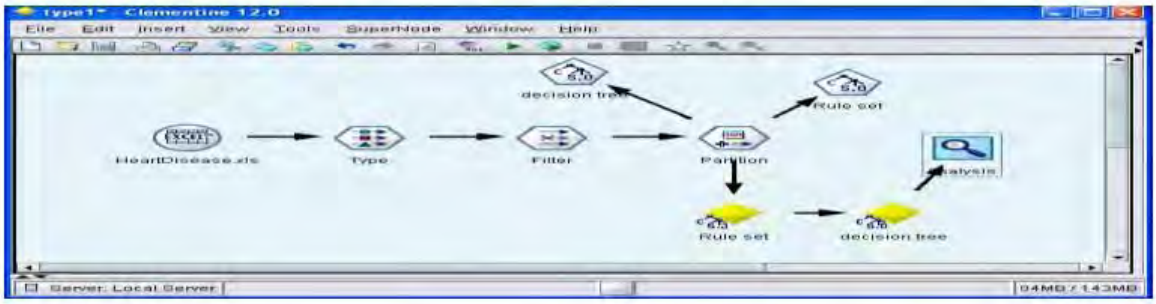
مدل tree C & R با افزایش تعداد سطوح تا ۱۱ سطح کارایی مدل بالاتر می رود اما از ۱۱ سطح به بعد افزایش تعداد سطوح تاثیری در کارایی مدل ندارد. اولویت ویژگی ها در مدل tree C & R با ۱۱ سطح در شکل ۸ نشان داده شده است.




شکل ۸-اولویت ویژگیها در مدل C&R tree با ۱۱ سطح

## پیاده سازی مدل C5. 0

مدل c5. 0 در برابر مواجه شدن با مسائلی که شامل اطلاعات از دسته رفته (ناقص) و دادههایی که شامل تعداد زیادی فیلد ورودی هستند کاملا قدرتمند است؛ و زمان زیادی برای آموزش نیاز ندارد تا بتواند پیشگویی و تخمین بزند، همچنین این مدل دارای قدرت بسیار بالایی در افزایش میزان دقت و صحت دسته بندی است. C5. ۰ یک روش خاص برای بالا بردن میزان نرخ صحت و دقت ارائه میدهد که به اختصار boosting نامیده می شود. این مدل ها در نرم افزار کلمنتاین به صورت شکل ۹ اجرا می شوند. این مدل ها به صورت شکل ۹ اجرا می شوند.




شکل۹-مدل c5.0 در کلمنتاین

## نتایج حاصل از پیاده سازی

در هر سه مدل شبکه عصبی از تعداد ۴۸۱ رکورد برای آموزش و ۱۳۴ رکورد برای تست استفاده شده است. پس با توجه به نتایج به دست آمده در مجموعه داده بیماران قلبی شبکه عصبی Exhamtive prune از مدل های دیگر شبکه عصبی بهتر عمل کرده است. در جدول ۲ نتایج ارزیابی مدل های پیشنهادی با یکدیگر مقایسه شده اند.

**جدول ۲. نتایج آموزش سبکه عصبی**

| نام مدل | شبکه عصبی | | C&R tree | | C5.0 | |
|---|---|---|---|---|---|---|
| قسمت | آموزشی | تست | آموزشی | تست | آموزشی | تست |
| درست | ٪۸۶/۹۲ | ٪۸۶/۸۴ | ٪۸۹/۱۷ | ٪۸۲/۴۷ | ٪۹۰/۲۲ | ٪۸۱/۷۵ |
| نادرست | ٪۱۳/۰۸ | ٪۱۳/۱۶ | ٪۱۰/۸۳ | ٪۱۷/۵۳ | ٪۹/۷۸ | ٪۱۸/۲۵ |

همانگونه که مشاهده می شود میزان دقت پیش بینی الگوریتم شبکه عصب ٪۸۴/۸۶، الگوریتم C & R tree، ٪۴۷/۸۲ الگوریتم C5. 0، ٪۸۱ /۷۵ می باشد و نتایج بررسی شده نشان دهنده این است مدل شبکه عصبی بر مجموعه داده بیماری های قلبی بهتر عمل می کند.

## نتیجه گیری

چنانچه مشخص است قلب یک اندام حیاتی برای بدن است و تشخیص بیماری قلبی یک اقدام مهم در حیطه پزشکی به حساب می آید. ما در این مقاله با در دست داشتن پایگاه داده ای از اطلاعات بیماران قلبی، در ابتدا به توضیح این اطلاعات

پرداختیم و سپس با بکارگیری تکنیک های داده کاوی به یک سیستم به عنوان سیستم پشتیبان تصمیم پزشکان دست یافتیم. این سیستم به پزشکان کمک میکند تا علاوه بر دانش و تخصص خود از تشخیص این سیستم نیز یاری بگیرند و در نهایت تشخیص دقیقتری در مورد بیماری بدهند.

**منابع**

[1] Frank Lemke and Johann-Adolf Mueller, "Medical data analysis using self-organizing data mining technologies"Systems Analysis Modelling Simulation, 2003, Vol. 43, no. 10, pp: 1399 - 1408.

[2] Ozbakır Lale, Bay kasoglu Adil and Kulluk Sinem. "A soft computing-based approach for integrated training and rule extraction from artificial neural networks:DIFACONN-miner", Applied Soft Coputing, 2010, Vol. 10,pp: 304-317.

[۳] Lee M. L., Lu H., Ling T.W. and Ko Y.K., "Cleansing Data for Mining and Warehousing," International Conference on Database and Expert Systems Applications (DEXA), Florence, Italy, August 1999.

[٤] F. H. Saad, B. de la Iglesia, and G. D. Bell, "A Comparison of Two Document Clustering Approaches for Clustering Medical Documents", International Conference on Data Mining (DM IN-06), 2006.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

# Use of artificial intelligence to predict and diagnose hypertension

**Farshid Sahba 1**
**Seyedeh Fatemeh Muntajbi 2**
**Mahsa Nasser Khaki 3**

Date of Receipt: 2021/03/29 Date of Issue: 2021/04/15

**Abstract**

Heart disease alone is a very valuable and one of the leading causes of death. That is why the correct and timely diagnosis of this disease is very important. Methods of diagnosis and treatment of this disease have many and costly side effects. If researchers are looking for cheap and high-precision methods to diagnose the disease. Inse artificial intelligence is used in many systems, the most important of which is health systems. One of the heart diseases that cause many people to lose their lives every year is hypertension, which in this study seeks to predict and diagnose this disease by using artificial intelligence.

**Keywords**

Blood pressure, artificial intelligence, prediction

1. Assistant Professor, Department of Information Technology Engineering, Ghiasuddin Jamshid Kashani University, Abyek, Iran.
2. Master student, Department of Information Technology Engineering, Ghiasuddin Jamshid Kashani University, Abyek, Iran.
3. Master student, Department of Information Technology Engineering, Ghiasuddin Jamshid Kashani University, Abyek, Iran.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی