مجله انفورماتیک سلامت و زیست پزشکی
مرکز تحقیقات انفورماتیک پزشکی
دوره هفتم شماره سوم. ص ۳۱۷-۳۰۴، ۱۳۹۹

مقاله پژوهشی

پیش‌بینی خطر ابتلاء به بیماری پوکی استخوان با استفاده از درخت تصمیم و شبکه عصبی

امین عابدینی[۱]، عفت جبارپور[۲]*، عباسعلی کشتکار[۳]

• دریافت مقاله: ۹۸/۸/۱۱ • پذیرش مقاله: ۹۸/۱۱/۱۴

مقدمه: پوکی استخوان یکی از دلایل مهم ناتوانی و مرگ در افراد مسن است. هدف این تحقیق تعیین عوامل تأثیرگذار در بروز پوکی استخوان و ارائه مدل پیش‌بینی کننده‌ای جهت سرعت بخشیدن به تشخیص وکاهش هزینه‌ها می‌باشد.

روش: در این مطالعه بنیادی توصیفی مدل جدیدی جهت شناسایی ویژگی‌های تأثیرگذار بر پوکی استخوان ارائه شده است. اطلاعات مربوط به ۴۰۸۳ نفر زن جهت کشف دانش با ابزار داده کاوی Clementine12 مورد بررسی قرار گرفته است. با استفاده از الگوریتم‌های داده‌کاوی شامل درخت تصمیم و شبکه عصبی مصنوعی قوانینی استخراج شده که می‌توان به عنوان الگویی برای پیش‌بینی وضعیت بیماران از آن‌ها استفاده کرد و در نهایت دقت مدل‌های ساخته شده با یکدیگر مقایسه شده‌اند.

نتایج: این تحقیق مدل‌های متعدد را بر روی تعداد ویژگی‌های متفاوت بررسی و نتایج حاصل را به منظور یافتن بهترین مدل پیش‌بینی کننده از نظر دقت و صحت با هم مقایسه می‌کند. دقت طبقه‌بندی مدل شبکه عصبی MLP با ۹۲/۱۴ درصد از دیگر الگوریتم‌های به کار رفته در این مطالعه بیشتر است. با توجه به شناسایی عوامل تأثیرگذار بر پوکی استخوان می‌توان برای یک نمونه جدید احتمال ابتلاء به این بیماری را پیش‌بینی کرد.

نتیجه‌گیری: سازمان‌های متولی مراقبت‌های بهداشتی و سلامت همواره حجم زیادی از اطلاعات را جمع‌آوری می‌کنند در حالی که این اطلاعات و داده‌ها به درستی مورد استفاده قرار نمی‌گیرند. این مطالعه نشان می‌دهد با کشف الگوها و روابط پنهان در این داده‌ها می‌توان از آن‌ها در جهت بهبود کیفیت ارائه خدمات تشخیصی و درمانی بهره برد.

کلیدواژه‌ها: پوکی استخوان، داده‌کاوی، درخت تصمیم، شبکه عصبی مصنوعی، کلمنتاین

•ارجاع: عابدینی امین، جبارپور عفت، کشتکار عباسعلی. پیش‌بینی خطر ابتلاء به بیماری پوکی استخوان با استفاده از درخت تصمیم و شبکه عصبی. مجله انفورماتیک سلامت و زیست پزشکی ۱۳۹۹؛ ۷(۳): ۱۷-۳۰۴.

۱. کارشناسی ارشد هوش مصنوعی، کامپیوتر و فناوری اطلاعات، دانشگاه آزاد اسلامی قزوین، قزوین، ایران
۲. دانشجوی دکتری مدیریت صنعتی، دانشکده مدیریت، دانشگاه تهران، تهران، ایران
۳. دانشیار، گروه توسعه آموزش علوم سلامت، دانشکده بهداشت، دانشگاه علوم پزشکی تهران، تهران، ایران
* **نویسنده مسئول:** عفت جبارپور
آدرس: تهران، وزارت علوم، تحقیقات و فناوری
•شماره تماس: ۸۲۲۳۴۱۸۷ -۰۲۱
• **Email:**ejabarpoor@yahoo.com

[Downloaded from jhbmi.com on 2026-01-30]

مقدمه

پوکی استخوان بیماری است که با کاهش تراکم استخوان و از دست رفتن کیفیت ریزساختار استخوان، شناخته می‌شود که خود منجر به افزایش خاصیت شکنندگی استخوان و افزایش خطر شکستگی می‌گردد [۱]. سازمان جهانی بهداشت پوکی استخوان را به صورت کاهش تراکم، استخوان به میزان ۲/۵ انحراف معیار کمتر از متوسط حداکثر تراکم استخوانی در افراد جوان جامعه تعریف کرده است [۲].

عوامل ژنتیکی و نژادی از عوامل مهم تأثیرگذار بر تراکم استخوانی محسوب می‌شوند [۳،۴]، همچنین عوامل فیزیولوژیک، محیطی و سبک زندگی می‌توانند به طور معنی‌داری در دست‌یابی به حداکثر تراکم استخوانی و حفظ آن در طول زندگی نقش داشته باشند [۶، ۵]. برخی از این عوامل شامل تأثیر هورمون‌های جنسی در سنین نوجوانی [۷]، تغذیه کافی و وزن کلی بدن [۸] و همچنین میزان فعالیت فیزیکی می‌باشند [۹].

تعیین عوامل خطرساز پوکی استخوان در جوامع و کشورهای مختلف و در سطوح مختلف اجتماعی- اقتصادی می‌تواند در طرح‌ریزی برنامه‌های پیشگیری از پوکی استخوان کمک کننده باشد. پوکی استخوان یکی از دلایل مهم ناتوانی و مرگ در افراد مسن است. میزان مرگ‌ومیر ناشی از شکستگی لگن (سر استخوان ران) در سال اول بعد از شکستگی، در افراد مسن، به حدود ۲۰ می‌رسد و نیمی از این افراد در باقی‌مانده عمر خویش دچار درجاتی از ناتوانی خواهند بود [۱۰].

امروزه روش‌های پیشگیری و درمانی مختلفی در دسترس بوده یا در حال ارائه هستند. پیشگیری و درمان پوکی استخوان شامل روش‌های دارویی و غیردارویی است. در هر دو مرحله پیشگیری و درمان، هدف انجام هر چه زودتر اقداماتی است که منجر به حفظ تراکم استخوان و مانع از دست رفتن یک دستی و صحت ساختار استخوانی شود تا به این ترتیب از بروز شکستگی‌های پاتولوژیک جلوگیری شود [۱۰]. پیش‌بینی می‌شود که بیش از ۷۵٪ از شکستگی‌های ناشی از پوکی استخوان در ۵۰ سال آینده در کشورهای در حال توسعه رخ دهد [۱۱]. کشور ما ایران نیز همچون سایر کشورهای در حال توسعه در ۵۰ سال آینده دارای جمعیت قابل‌توجهی از افراد کهن سال خواهد بود و به نظر می‌رسد طرح‌ریزی برنامه‌های پیشگیری از وقوع پوکی استخوان می‌بایست جزء اولویت‌های بهداشتی کشورمان قرار گیرد.

با توجه به پیشرفت روزافزون تکنولوژی اطلاعات باید به دنبال الگوهای جالب جهت تشخیص پزشکی، پیشگیری و درمان باشیم. یکی از روش‌های مناسب برای پیش‌بینی و تشخیص در حوزه‌های پزشکی رویکرد داده‌کاوی می‌باشد. داده‌کاوی روشی داده مدار و بر مبنای یادگیری و کشف الگوی پنهان در میان داده‌های حقیقی می‌باشد که از این الگو جهت پیش‌بینی برای موارد مشابه استفاده می‌کند [۱۲].

داده‌کاوی یا کشف دانش از میان پایگاه‌های اطلاعاتی علمی است که برای تصمیم‌گیری‌های هوشمندانه بسیار کاربرد دارد. انواع مختلفی از تکنیک‌های داده‌کاوی در استخراج اطلاعات که می‌تواند برای پیش‌بینی مورد استفاده قرار گیرد وجود دارد که از آن جمله می‌توان به شبکه‌های عصبی، درخت تصمیم و رگرسیون لجستیک اشاره کرد [۱۴، ۱۳].

Moudani و همکاران بر روی سیستم هوشمند پیش‌بینی پوکی استخوان کار کردند در این بررسی از روش جنگل تصادفی برای شناسایی موارد پوکی استخوان استفاده شده است [۱۵].

Zhou و همکاران با استفاده از داده‌کاوی مدلی برای عوامل خطر ابتلاء به پوکی استخوان بر اساس طب سنتی چین و طب مدرن غربی ارائه کردند. بدین منظور از دو روش ماشین بردار پشتیبان و جنگل تصادفی استفاده کردند [۱۶].

شریف‌خانی و همکاران بر روی ارائه مدلی برای پیش‌بینی احتمال ابتلاء به بیماری پوکی استخوان با استفاده از الگوریتم‌های درخت تصمیم کار کردند [۱۷].

Sathawane و همکاران بر روی داده‌کاوی در سوابق بالینی به منظور پیشگویی خطر ابتلاء به پوکی استخوان کار کردند [۱۸].

متأسفانه تعداد تحقیقات انجام شده با موضوع استفاده از داده‌کاوی در خصوص بیماران پوکی استخوان در مقایسه با سایر بیماری‌های خاص مانند سرطان بسیار محدود بوده و بیشتر تحقیقات این حوزه با مطالعات بالینی و یا با استفاده از روش‌های آماری صورت پذیرفته است. به خصوص در داخل کشور تحقیقاتی که با موضوع استفاده از داده‌کاوی در خصوص بیماران پوکی استخوان انجام شده باشد بسیار محدود می‌باشند؛ لذا هدف این تحقیق این است تا با استفاده از الگوریتم‌های داده‌کاوی ویژگی‌های مؤثر در ابتلا به پوکی استخوان را پیش‌بینی شود.

[Downloaded from jhbmi.com on 2026-01-30]

روش

از آنجایی که اطلاعات و داده‌های این پژوهش بزرگ می‌باشد، روش‌های درخت تصمیم و شبکه عصبی انتخاب شد [۲۰، ۱۹]. برای داشتن یک داده‌کاوی مؤثر علاوه بر نیاز به داده‌های مرتبط، باید از یک فرآیند و روش داده‌کاوی مناسب استفاده شود. روشی که کلیه مراحل داده‌کاوی اعم از جمع‌آوری داده، آماده‌سازی داده، مدل‌سازی و ارزیابی را در برگیرد.

جمع‌آوری و توصیف داده: سوابق کلینیکی مورد استفاده در این تحقیق مربوط به افراد مراجعه کننده به مرکز تحقیقات پوکی استخوان در پژوهشگاه علوم غدد و متابولیسم دانشگاه علوم پزشکی تهران می‌باشد. این داده‌ها شامل اطلاعات مربوط به ۴۰۸۳ نفر خانم که طی سال‌های ۱۳۸۴ تا ۱۳۸۸ شمسی به این مرکز مراجعه نموده و تشکیل پرونده داده‌اند، می‌باشد که در فایل اکسل جمع‌آوری شده است. تعداد ۴۲۵ ویژگی در این مجموعه داده ثبت گردیده است که شامل اطلاعات فردی، سبک زندگی و اطلاعات بیماری‌های افراد می‌باشد. تعدادی از این ویژگی‌ها عبارت‌اند از: سن، وزن، قد، شاخص توده بدن، سن قاعدگی، تعداد بارداری، مدت شیردهی، سقط جنین، سن یائسگی، اعمال جراحی، بیماری‌های زمینه‌ای، مصرف داروهای مختلف. تمامی افراد مراجعه کننده از نظر این که فرد به بیماری پوکی استخوان مبتلا می‌باشد یا خیر برچسب‌گذاری شده‌اند. در این مطالعه از تعداد ۴۰۸۳ خانم مراجعه کننده ۳۶۵ نفر به بیماری پوکی استخوان مبتلا بوده اند.

آماده‌سازی داده: هدف از این مرحله افزایش کیفیت داده‌های مورد استفاده می‌باشد به نحوی که بتوان مجموعه داده مناسبی برای فازهای بعدی و مدل‌سازی داده‌ها تأمین شود. مراحل آماده‌سازی داده به شکل زیر انجام گردید.

پاک‌سازی داده‌ها: این مرحله شامل هموار کردن نویزها، شناخت و حذف داده‌های پرت (دورافتاده)، برطرف کردن ناسازگاری‌ها و پر کردن داده‌های از دست رفته می‌باشد.

یکسان‌سازی شکل‌ها: این موضوع نیز یکی از مسائل مهم به هنگام تجمیع داده‌ها است که به دلیل اهمیت آن و پنهان بودن زوایای آن از دید داده‌کاوان، به شکل مجزا مورد اشاره قرار می‌گیرد. به طور مثال در داده‌های ما گاهی مقدار وزن به گرم یا به کیلوگرم آمده است که باید تمامی داده‌ها به صورت یکسان و با یک شکل ویرایش گردند.

نرمال‌سازی داده‌ها: داده‌هایی شامل سن، وزن، قد، شاخص توده بدن، مدت شیردهی، تعداد بارداری و ... همه این داده‌ها گاهی مقادیر بزرگی دارند که می‌توانند تأثیر بیشتری را بر آنالیزها بگذارند. کلیه داده‌هایی که بدین‌صورت می‌باشند در این مرحله نرمالیزه می‌گردد.

ترکیب برخی داده‌ها: به منظور تسهیل در مدل‌سازی و استخراج نتایج قابل استناد از لحاظ پزشکی، داده‌ها از مقادیر پیوسته به مقادیر گسسته و دارای مفهوم پزشکی تبدیل می‌گردد.

جهت کاهش ابعاد از بین ۴۰۰ ویژگی اولیه در داده‌ها، با توجه به نظرخواهی از فرد متخصص در پژوهشگاه علوم غدد و متابولیسم و با کمک از مقالات علمی [۲۵-۲۱] در زمینه پوکی استخوان، ۴۵ ویژگی جهت مدل‌سازی انتخاب گردید.

مدل‌سازی داده

کلاس‌بندی با استفاده از درخت تصمیم (Chi-Squared Automatic Interaction Detection) CHAID

این روش که Chi-squared نیز نامیده می‌شود بر خلاف درخت‌های C&R Tree و QUEST می‌توانند درختی غیربایری تولید کند، بدین معنا که بعضی از تقسیمات بیش از دو شاخه دارند. فیلدهای هدف و پیش‌بینی شده می‌توانند هم به صورت دامنه‌ای وهم به صورت طبقه‌بندی شده باشند. البته زمان محاسبه این الگوریتم طولانی‌تر است [۲۸-۲۶].

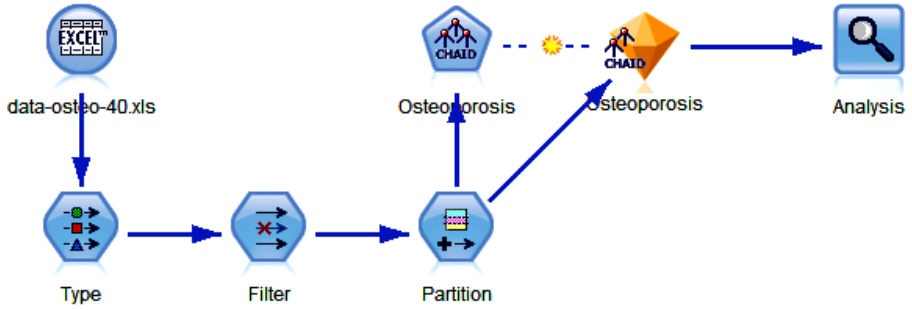
در ابتدا CHAID اهمیت هر یک از متغیرهای پیش‌بینی کننده را برای پیش‌بینی متغیر هدف با توجه به آزمون استقلال کای دو تعیین می‌کند. اگر بیش از یک مورد از این روابط از لحاظ آماری مهم باشند، CHAID پیش‌بینی کننده‌ای را انتخاب می‌کند که از سایرین مهم‌تر باشد (کوچک‌ترین مقدار p). اگر پیش‌بینی کننده‌ای بیش از دو طبقه داشته باشد، این طبقات مقایسه شده و طبقاتی که تأثیر آن‌ها در نتیجه یکسان باشد، با هم ادغام می‌شوند. این فرآیند ادغام طبقات، هنگامی به اتمام می‌رسد که تمام طبقات باقی مانده در یک سطح معینی در آزمون باهم متفاوت باشند. در مورد پیش‌بینی‌کنندگان مجموعه‌ای، تمام طبقات قابل ترکیب هستند؛ ولی در مورد نوع ترتیبی تنها طبقات مجاور قابل ادغام می‌باشند.

مدل‌سازی CHAID

در این جا با توجه به اثرات کاهش ابعاد در داده‌کاوی، مدل‌سازی با ۴۵ ویژگی منتخب در مرحله کاهش ابعاد شروع گردید. ویژگی بیماری پوکی استخوان به عنوان ویژگی هدف مشخص شد و ویژگی‌های دیگر به عنوان ویژگی‌های ورودی انتخاب شدند. درخت تصمیم ساخته شده با روش CHAID و

[Downloaded from jhbmi.com on 2026-01-30]

مجموعه قواعد (Rule set) برای این درخت را به دست آورده و در آخر به کمک گره آنالیز، دقت طبقه بندی مدل مشخص گردید (شکل ۱).



شکل ۱: مدل CHAID

مدل‌سازی Exhaustive CHAID

با استفاده از روش Exhaustive CHAID مدل اجرا شد و دقت طبقه‌بندی مدل و ویژگی‌های مهم آن محاسبه گردید. در این حالت نیز مانند مدل قبل و به همان ترتیب طراحی گردید. در کلیه مدل‌های ویژگی پوکی استخوان به عنوان ویژگی هدف مشخص و ویژگی‌های دیگر به عنوان ویژگی‌های ورودی انتخاب شدند. در ابتدا ۴۵ ویژگی برای این مدل در نظر گرفته شد. سپس درخت تصمیم ساخته شده با روش CHAID Exhaustive، مجموعه قواعد (Rule set) و همچنین دقت طبقه‌بندی برای داده های آموزشی و آزمایشی به دست آورده شد.

در نخستین مرحله داده‌های آموزشی را در جریان تعریف شده به عنوان ورودی به گره CHAID وارد کرده و پارامترهای مدل به صورت زیر تنظیم شد.

Method: Chaid
Tree growing algorithm:CHAID & exhaustive CHAID
Maximum tree depth: 7
Chi-square for categorical targets: Pearson

کلاس‌بندی با استفاده از شبکه عصبی

شبکه‌های عصبی جزء پرکاربردترین روش‌های پیش‌بینی هستند. شبکه‌های عصبی سیستم‌های دینامیکی هستند که با پردازش روی داده‌های تجربی، دانش نهفته آن را استخراج می‌کنند. پردازش اطلاعات در شبکه‌های عصبی روشی همانند مغز انسان دارد. این شبکه از تعداد زیادی از عناصر پردازشی به هم پیوسته (نورون‌ها) تشکیل شده که به صورت موازی کار می‌کنند [۲۹]. شبکه‌های عصبی هوشمند می‌توانند با تحلیل داده‌ها و ارتباط بین ورودی و خروجی‌های واقعی، جنبه‌های پنهان و روابط پنهان در داده‌ها را استخراج نماید. این دانش از وزن‌هایی که به شبکه عصبی اختصاص می‌دهد قابل استنتاج است [۳۰]. شبکه‌های عصبی مصنوعی از معماری سه لایه شامل: لایه ورودی، پنهان و خروجی تشکیل شده است. هر لایه شامل تعدادی نورون یا گره می‌باشد. مدل پرسپترون چند لایه کاربرد موفقیت‌آمیزی در حل برخی مسائل از جمله شناسایی الگو و تخمین تابع دارد.

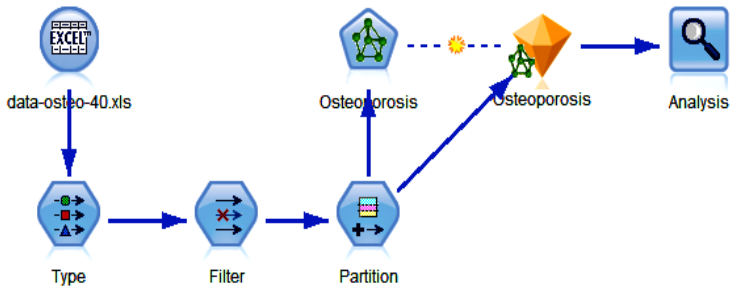
رایج‌ترین مدل شبکه‌های عصبی که مورد استفاده قرار می‌گیرد، مدل پرسپترون چندلایه می‌باشد. در این مدل یک تابع که معمولاً غیرخطی می‌باشد بر روی جمع‌وزنی ورودی‌ها اعمال شده و خروجی مدل ساخته می‌شود. متغیرهای ورودی مسئله به لایه اولیه یا لایه ورودی وارد می‌شوند و متغیرهای پاسخ در لایه نهایی یا خروجی قرار می‌گیرند و در این لایه‌ها برای هر متغیر یک گره در نظر گرفته می‌شود. لایه‌های بین این دو به لایه‌های پنهان موسوم هستند. روش قطعی برای تعیین تعداد گره‌های لازم در این لایه‌ها وجود ندارد و اغلب از روش سعی و خطا تعداد مناسب این گره‌ها تعیین می‌شود. [۳۲،۳۳].

به صورت کلی، شبکه را با استفاده از قواعد و داده‌ها، آموزش

[Downloaded from jhbmi.com on 2026-01-30]

داده و بـا اسـتفاده از قابلیـت یـادگیری شـبکه، الگوریتم‌هـای متنوعی پیشنهاد می‌گردد که همگـی سـعی در نزدیـک کـردن خروجی تولید شده توسط شـبکه بـه خروجـی ایـده‌آل و مـورد انتظار را دارند [۳۴].

برای استفاده از مدل شبکه عصبی ابتدا باید آمـوزش صـورت بگیرد. داده‌ها به دو دسته آموزشی و آزمایشی تقسیم گردید کـه بـه ترتیـب نسـبت ۳۰،۷۰ دارنـد. بعـد از دسـته‌بندی داده‌هـا و آموزش مدل بـه وسـیله نرم‌افـزار Clementine یـک شـبکه عصبی دو لایه با ۹ نود میانی تشکیل شد (شکل ۲).



شکل ۲: مدل شبکه عصبی

مدل‌سازی شبکه عصبی با روش MLP

شـــبکه عصـــبی پرســـپترون چندلایـــه (Multi-Layer Perceptron)MLP یـکـی از رایج‌تـرین انـواع شـبکه‌های عصبی است. در این حالت، تعـداد نورون‌هـای لایـه ورودی بـا تعداد عناصر بردار ورودی و تعداد نورون‌های لایـه خروجـی بـا تعداد عناصر بردار خروجی برابر در نظر گرفته می‌شوند [۳۵].

در قسمت تنظیمات گره شبکه عصبی در نرم‌افـزار کلمنتـاین در ابتدا روش MLP انتخاب گردیـد. در تنظیمـات ورودی ایـن مدل همانند مدل‌سازی درخـت تصـمیم ورودی ۴۵ ویژگـی در نظر گرفته شد و نتایج مشخص گردید.

مدل‌سازی شبکه عصبی با روش RBF

یکی دیگر از انواع شبکه عصبی مصنوعی، شبکه بـا تـابع پایـه شعاعی (Radial basis function) RBF است. در این نـوع از شبکه سیگنال‌های ورودی به طـور مسـتقیم وارد سـلول‌های لایه مخفی می‌شوند و توابـع فعالیـت در ایـن شـبکه‌ها محلـی می‌باشند. در این نوع از شبکه، تعداد سلول‌های لایـه خروجـی برابر با تعداد خروجی‌ها است [۳۶].

در ایـن قسـمت تنظیمـات گـره شـبکه عصـبی در نرم‌افـزار Clementine روش RBF انتخاب و بر اسـاس ورودی مـدل بررسی شد.

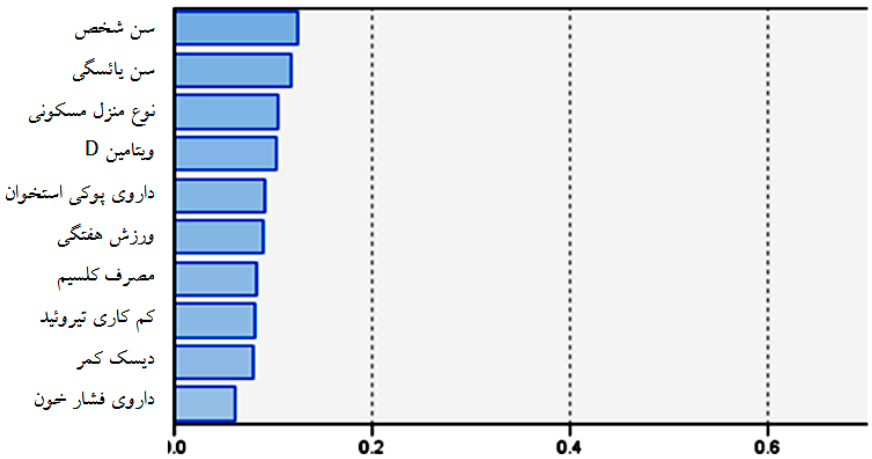
نتایج

ارزیابی مـدل تهیـه شـده توسـط درخـت تصـمیم CHAID

به منظور ارزیابی مدل ارائه شده توسط ایـن روش، در نرم‌افـزار Clementine داده‌های آموزشـی و داده‌هـای آزمایشـی را بـه طور مجزا به گره Analysis وارد نمـوده و دقـت مـدل تهیـه شده مورد ارزیابی قرار گرفت.

نتایج حاصل از ارزیابی مدل درخت تصمیم بـا روش CHAID که برای مجموعه داده آموزشی، آزمایشـی و همچنـین قـوانین استخراج شده از آن به شرح زیر است:

[Downloaded from jhbmi.com on 2026-01-30]



شکل ۳: ده ویژگی مهم CHAID

در این مدل فاکتورهای «سن شخص، سن یائسگی، نوع منزل مسکونی، ویتامین D، داروی پوکی استخوان، ورزش هفتگی، مصرف کلسیم، کم کاری تیروئید، دیسک کمر و داروی فشار خون» به عنوان فاکتورهایی که بیشترین تأثیر را برروی پوکی استخوان دارند، شناسایی کرد (شکل ۳).

برخی از قوانین استخراج شده از این مدل در این مرحله با سطح اطمینان بیش از ۹۰ درصد به شرح زیر می‌باشند:

قانون ۱: احتمال این که فردی ویتامین D، داروی پوکی استخوان و کلسیم مصرف نکند و یائسه هم نشده باشد و سالم باشد، ۹۷/۸ درصد است.

قانون ۲: احتمال این که فردی ویتامین D و داروی پوکی استخوان مصرف نکند، یائسه هم نشده باشد، کلسیم مصرف کند و سالم باشد، ۹۴/۴ درصد است.

قانون ۳: احتمال این که فردی داروی پوکی استخوان، داروی فشار خون و ویتامین D مصرف نکند، یائسه شده باشد، کمتر از ۴ بارداری داشته باشد، در خانه حیاط دار زندگی کند و دیسک کمر نداشته باشد و سالم باشد، ۹۲/۳ درصد است.

قانون ۴: احتمال این که فردی داروی پوکی استخوان، داروی فشارخون و ویتامین D مصرف نکند، یائسه شده باشد، کمتر از ۴ بارداری داشته، در خانه حیاط دار زندگی کند، همچنین دیسک کمر داشته باشد و سالم باشد، ۹۶/۶ درصد است.

قانون ۵: احتمال این که فردی داروی پوکی استخوان و ویتامین D مصرف نکند، یائسه شده باشد، کمتر از ۴ بارداری داشته باشد، در آپارتمان زندگی کند، کم کاری تیروئید نداشته باشد و در هفته حداکثر سه روز ورزش کند و سالم باشد، ۹۲/۸ درصد است.

قانون ۶: احتمال این که فردی داروی پوکی استخوان و ویتامین D مصرف نکند، یائسه شده باشد و بیشتر از ۴ بارداری داشته باشد و سالم باشد، ۹۶/۸ درصد است.

قانون ۷: احتمال این که فردی داروی پوکی استخوان مصرف نکند، ویتامین D مصرف کند و حداکثر ۳ روز در هفته ورزش کندو سالم باشد، ۹۷/۴ درصد است.

دقت طبقه‌بندی مدل CHAID برای این مدل در جدول ۱ نشان داده شده است.

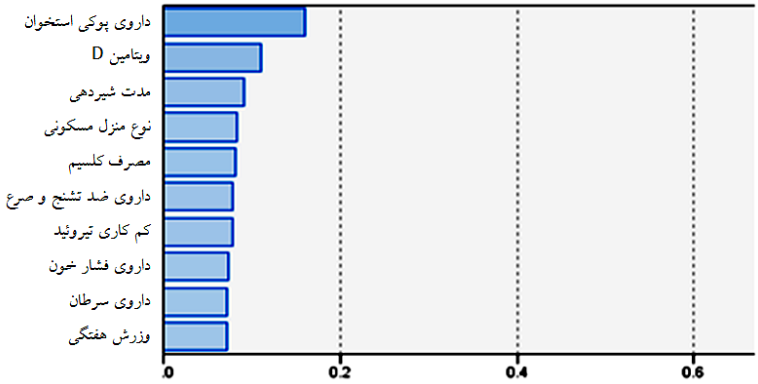
جدول ۱: دقت طبقه‌بندی در مدل CHAID

	تعداد ویژگی	آزمایشی	آموزشی
CHAID	۴۵	٪۹۰/۴۶	٪۹۱/۲۹

[Downloaded from jhbmi.com on 2026-01-30]

ارزیابی مدل با روش Exhaustive -CHAID

نتایج حاصل از ارزیابی مدل درخت تصمیم با روش EX-CHAID برای مجموعه داده آموزشی، آزمایشی و همچنین قوانین استخراج شده از آن به شرح زیر می‌باشد.



شکل ٤: ده ویژگی مهم در مدل Exhaustive CHAID

مدل ایجاد شده توسط Exhaustive CHAID در مرحله اول فاکتورهای «داروی پوکی استخوان، ویتامین D، مدت شیردهی، نوع منزل مسکونی، مصرف کلسیم، داروی ضد تشنج و صرع، کم کاری تیروئید، داروی فشارخون، داروی سرطان، وزرش هفتگی» را به عنوان فاکتورهایی که بیشترین تأثیر را بر روی پوکی استخوان دارند، شناسایی نموده است. (شکل ۴)

برخی از قوانین استخراج شده از این مدل در این مرحله با سطح اطمینان بیش از ۹۰ درصد به شرح زیر می‌باشند:

قانون ۱: احتمال این که فردی ویتامین D، داروی پوکی استخوان و کلسیم مصرف نکند، یائسه نشده باشد و سالم باشد، ۹۷/۸ درصد است.

قانون ۲: احتمال این که فردی ویتامین D و داروی پوکی استخوان مصرف نکند، کلسیم مصرف کند، یائسه نشده باشد و سالم باشد، ۹۴/۴ درصد است.

قانون ۳: احتمال این که فردی ویتامین D، داروی پوکی استخوان، داروی چربی خون و داروی انعقاد خون مصرف نکند، در خانه حیاطدار زندگی کند، یائسه شده باشد و سالم باشد، ۹۵/۷ درصد است.

قانون ۴: احتمال این که فردی ویتامین D، داروی پوکی استخوان مصرف نکند، در آپارتمان زندگی کند، کم کاری تیروئید نداشته باشد، شیردهی داشته باشد، یائسه شده باشد و سالم باشد، ۰/۹۱ درصد است.

قانون ۵: احتمال این که فردی ویتامین D، داروی پوکی استخوان مصرف نکند، در آپارتمان زندگی کند، بیش از ۶ سال شیردهی داشته باشد، یائسه شده باشد و سالم باشد، ۹۵/۳ درصد است.

قانون ۶: احتمال این که فردی داروی پوکی استخوان و داروی سرکوب گر سیستم ایمنی مصرف کند، داروی سرطان مصرف نکند و سالم باشد، ۹۵/۶ درصد است.

دقت طبقه‌بندی مدل Exhaustive -CHAID در جدول ۲ نشان داده شده است.

جدول ۲: دقت طبقه‌بندی در مدل Exhaust Chaid

	تعداد ویژگی	آزمایشی	آموزشی
Exhaust CHAID	۴۵	٪۹۱/۰۵	٪۹۱/۶۷

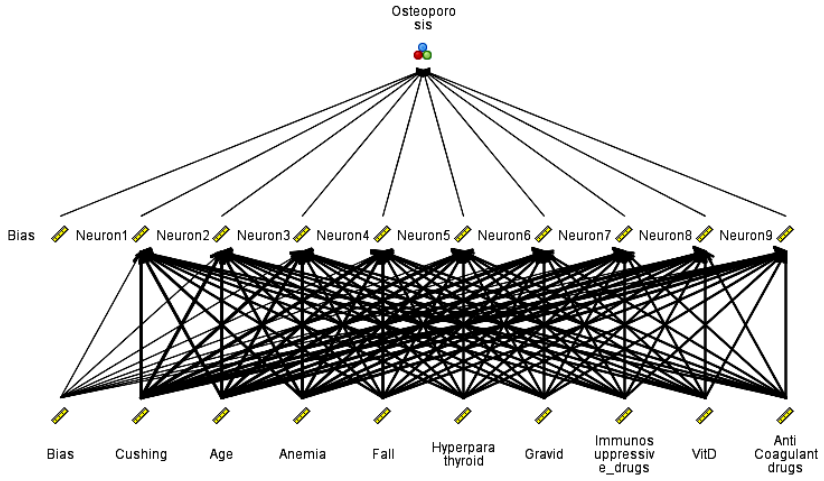
[Downloaded from jhbmi.com on 2026-01-30]

ارزیابی مدل تهیه شده توسط شبکه عصبی

در مدل تهیه شده توسط شبکه عصبی علاوه بر ترسیم گرافِ روابط مابین متغیرها، فاکتورهای دارای بیشترین اهمیت نیز شناسایی شدند. به منظور ارزیابی دقت این کلاسه‌بندی تولید شده، داده‌های آزمایشی و آموزشی به طور مجزا به گره Analysis وارد نموده و دقت مدل تهیه شده مورد ارزیابی قرار گرفت.

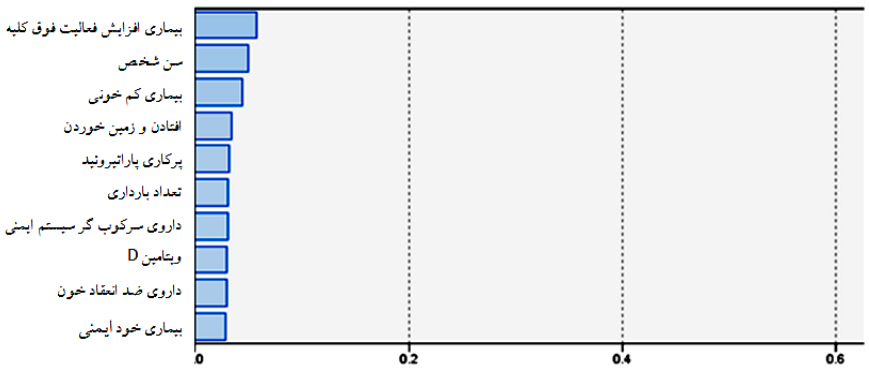
ارزیابی مدل شبکه عصبی روش MLP

در قسمت تنظیمات گره شبکه عصبی در نرم‌افزار کلمنتاین در ابتدا روش MLP را انتخاب و با ۴۵ ویژگی ورودی نتایج زیر به دست آمد (شکل ۵).



شکل ۵: شبکه عصبی طراحی شده با کلمنتاین

مدل ایجاد شده توسط شبکه عصبی (روش MLP) فاکتورهای «بیماری افزایش فعالیت فوق کلیه، سن شخص، بیماری کم‌خونی، افتادن و زمین خوردن، پرکاری پاراتیروئید، تعداد بارداری، داروی سرکوب گر سیستم ایمنی، ویتامین D، داروی ضد انعقاد خون، بیماری خود ایمنی» را به عنوان فاکتورهایی که بیشترین تأثیر را برروی پوکی استخوان دارند شناسایی نموده است (شکل ۶).



شکل ٦: ده ویژگی مهم شبکه عصبی روش MLP

[Downloaded from jhbmi.com on 2026-01-30]

جدول ۳: دقت طبقه‌بندی در مدل شبکه عصبی روش MLP

	تعداد ویژگی	آزمایشی	آموزشی
ANN (MLP)	۴۵	٪۹۲/۱۴	٪۹۲/۸۷

ارزیابی مدل شبکه عصبی روش RBF

در قسمت تنظیمات گره شبکه عصبی در نرم‌افزار کلمنتاین روش RBF انتخاب و سپس با ۴۵ ویژگی ورودی نتایج به دست می‌آید.

مدل ایجاد شده توسط شبکه عصبی (روش RBF) فاکتورهای «داروی چربی خون، داروی پوکی استخوان، جراحی گوارش، پرکاری تیروئید، بیماری افزایش فعالیت فوق کلیه، پرکاری پاراتیروئید، افتادن و زمین خوردن، بیماری کم‌خونی، ویتامین D، چربی خون بالا» را به عنوان فاکتورهایی که بیشترین تأثیر را بر روی پوکی استخوان دارند شناسایی نموده است (شکل ۷).

شکل ۷: ده ویژگی مهم شبکه عصبی روش RBF

در روش RBF دقت طبقه‌بندی برای داده‌های آموزشی ۹۱/۲۸ درصد و برای داده‌های آزمایشی ۹۰/۴۱ درصد می‌باشد که دقت مناسبی به نظر می‌رسد (جدول ۴).

جدول ٤: دقت طبقه‌بندی در مدل شبکه عصبی روش RBF

	تعداد ویژگی	آزمایشی	آموزشی
ANN (RBF)	۴۵	٪۹۰/۴۱	٪۹۱/۲۸

مدل ایجاد شده توسط شبکه عصبی روش RBF فاکتورهای «داروی چربی خون، داروی پوکی استخوان، جراحی گوارش، پرکاری تیروئید، بیماری افزایش فعالیت فوق کلیه، پرکاری پاراتیروئید، افتادن و زمین خوردن، بیماری کم‌خونی، ویتامین D، چربی خون بالا» ر ا به عنوان فاکتورهایی که بیشترین تأثیر را بر روی پوکی استخوان دارند، شناسایی نموده است.

بحث و نتیجه‌گیری

نتایج به دست آمده به صورت خلاصه به این صورت آورده می‌شود (جدول ۶).

[Downloaded from jhbmi.com on 2026-01-30]

جدول ٦: دقت طبقه‌بندی مدل‌های مختلف

	CHAID	Ex CHAID	ANN(MLP)	ANN(RBF)
آموزشی	٪۹۱/۲۹	٪۹۱/۶۷	٪۹۲/۸۷	٪۹۱/۲۸
آزمایشی	٪۹۰/۴۶	٪۹۱/۰۵	٪۹۲/۱۴	٪۹۱/۴۱

- مدل درخت تصمیم CHAID: دقت طبقه‌بندی ۹۰/۴۶ درصد می‌باشد. این مدل ویژگی‌های «سن شخص، سن یائسگی، نوع منزل مسکونی، ویتامین D، داروی پوکی استخوان، ورزش هفتگی، مصرف کلسیم، کم کاری تیروئید، دیسک کمر و داروی فشارخون» را به عنوان فاکتورهای ارزشمند و مهم در پوکی استخوان معرفی کرده است.
- مدل درخت تصمیم Exhaust-CHAID: به طور کلی دقت طبقه‌بندی توسط این مدل ۹۱/۰۵ درصد می‌باشد که دقت مناسبی به نظر می‌رسد. این مدل ویژگی‌های زیر را به عنوان ویژگی‌های مهم در بروز پوکی استخوان مشخص کرده است: «داروی پوکی استخوان، ویتامین D، مدت شیردهی، نوع منزل مسکونی، مصرف کلسیم، داروی ضد تشنج و صرع، کم کاری تیروئید، داروی فشار خون، داروی سرطان، ورزش هفتگی».
- مدل شبکه عصبی در روش MLP: مشخص شد که بیشترین دقت طبقه‌بندی مربوط به این مدل بوده و مقدار آن ۹۲/۱۴ درصد می‌باشد، در این مدل فاکتورهای «بیماری افزایش فعالیت فوق‌کلیه، سن شخص، بیماری کم‌خونی، افتادن و زمین خوردن، پرکاری پاراتیروئید، تعداد بارداری، داروی سرکوب گر سیستم ایمنی، ویتامین D، داروی ضد انعقاد خون، بیماری خود ایمنی» بیشترین تأثیر را بر روی پوکی استخوان داشته‌اند. همچنین این مدل بیشترین دقت را در بین مدل‌های ارائه شده در این تحقیق دارا می‌باشد.
- مدل شبکه عصبی در روش RBF: دقت طبقه‌بندی ۹۰/۴۱ درصد دارد و فاکتورهای «داروی چربی خون، داروی پوکی استخوان، جراحی گوارش، پرکاری تیروئید، بیماری افزایش فعالیت فوق کلیه، پرکاری پاراتیروئید، افتادن و زمین خوردن، بیماری کم‌خونی، ویتامین D، چربی خون بالا» را به عنوان ویژگی‌هایی که بیشترین تأثیر را بر روی پوکی استخوان دارند شناسایی نموده است.
- با بررسی فاکتورهایی که در هر مدل به عنوان فاکتورهای مهم شناخته شده‌اند مشاهده شد که سن شخص، مصرف ویتامین D، مصرف کلسیم، کم کاری تیروئید، مصرف داروی سرکوب‌گر سیستم ایمنی، مصرف داروی فشارخون، میزان ورزش هفتگی، پرکاری تیروئید و پاراتیروئید، احتمال و استعداد سقوط افراد، بیماری کم‌خونی، مدت شیردهی و چربی خون بالا از مهم‌ترین عوامل ابتلا به پوکی استخوان شناخته شده‌اند.
- علاوه بر این، ویژگی‌های سابقه پوکی استخوان در خانواده، بیماری دیابت، تعداد بارداری، مصرف داروی تشنج و صرع، مصرف داروهای سرطان، دیسک کمر و مصرف داروهای ضد انعقاد خون از فاکتورهای تأثیرگذار در ابتلای اشخاص به پوکی استخوان شناسایی شده‌اند. نکته قابل توجه این است که مطالعات بالینی و پزشکی در این زمینه نشان می‌دهند که عوامل خطر ابتلاء به پوکی استخوان عبارت‌اند از: سن شخص (بیش از ۶۵ سال)، یائسگی زودرس (قبل از ۴۵ سالگی)، مصرف کلسیم (یا از دست دادن کلسیم) ، مصرف چای و قهوه، عدم تحرک و ورزش، وزن پایین (شاخص توده بدن)، اختلالات تیروئید و پارا تیروئید و .... [۱۰].

دقت مدل پیش‌بینی کننده پوکی استخوان معرفی شده Wang و همکاران در مدل ترکیبی درخت تصمیم و شبکه عصبی مصنوعی با در نظر گرفتن ۳۳ ویژگی و بر روی ۲۹۳۴ زن، برابر ۷۰/۵ درصد است [۳۷].

Gao و همکاران در تحقیقی با بررسی یک ویژگی استخراج شده از طیف فرکتال تصویر میکروسی تی به پیش‌بینی پوکی استخوان در افراد پرداخته‌اند. آن‌ها مطالعه خود را با در نظر گرفتن ۱ ویژگی و بر روی ۱۴ فرد سالم و ۱۴ فرد مبتلا به پوکی استخوان انجام دادند. از درخت تصمیم C4.5 برای مدل پیش‌بینی کننده استفاده کرده‌اند. دقت به دست آمده ۹۲/۹ درصد است [۳۸].

Moudani و همکاران به پیش‌بینی ابتلاء به پوکی استخوان با استفاده از روش جنگل تصادفی درخت تصمیم پرداخته‌اند. این مطالعه با در نظر گرفتن ۱۵ ویژگی و بر روی ۲۸۴۵ نفر انجام گردید. مدل به کار رفته چهار کلاس بدون خطر، کم خطر، خطر متوسط، خطر زیاد و خطر شدید را برای احتمال ابتلاء به پوکی استخوان پیش‌بینی می‌کند. در این تحقیق ۱۵ ویژگی را مورد تحلیل قرار داده‌اند و دقت ۹۹/۹۲ درصد را به دست آورده‌اند [۱۵].

[Downloaded from jhbmi.com on 2026-01-30]

Zhou و همکاران با استفاده از داده‌کاوی مدلی برای عوامل خطر ابتلاء به پوکی استخوان بر اساس طب سنتی چین و طب مدرن غربی ارائه کردند. بدین منظور از دو روش ماشین بردار پشتیبان و جنگل تصادفی استفاده کردند. نتایج بررسی‌های آن‌ها نشان داد علائم و فاکتورهای ذکر شده در طب سنتی نقش مهم‌تری در ارزیابی خطر ابتلاء به پوکی استخوان دارند. از میان متغیرهای ذکر شده، سال‌های پس از یائسگی از عوامل خطر مهم در ابتلاء به پوکی استخوان محسوب می‌شود [۳۹].

در مطالعه شریف‌خانی و همکاران نشان دادند مدل شبکه عصبی مصنوعی با روش Multiple بیشترین دقت را برای پیش‌بینی افراد مبتلا به پوکی استخوان دارد. این مطالعه با ۶۰ ویژگی در نظر گرفته شده از اطلاعات پرسشنامه‌ای ۶۷۰ نفر انجام گردید و دقتی برابر با ۹۵/۷۰ را داشته است [۱۷].

با مطالعه شرایط و نتایج تحقیقات قبلی در حوزه داده‌کاوی و پوکی استخوان، مشخص شد این مطالعه از لحاظ حجم نمونه و تعداد ویژگی مورد بررسی، بالاترین تعداد را داشته است. همچنین استفاده از روش‌های درخت تصمیم و شبکه عصبی انتخاب ویژگی، یکی دیگر نقاط قوت این مطالعه محسوب می‌شود. همچنین در این تحقیق مدل‌های متعدد بر روی تعداد ویژگی‌های متفاوت بررسی و نتایج حاصل را به منظور یافتن بهترین مدل پیش‌بینی کننده از نظر دقت و صحت با هم مقایسه شد. با توجه به موارد ذکر شده، نتایج حاصل از داده‌کاوی بر روی سوابق کلینیکی مورد مطالعه نشان می‌دهد ویژگی‌های پیش‌بینی شده توسط الگوریتم‌های داده‌کاوی در مطالعه حاضر، هم راستا و هم جهت با نتایج بالینی به دست آمده از مطالعات پزشکی و یافته‌های پژوهش‌های گذشته نیز می‌باشد.

در پایان توجه به این نکته حائز اهمیت است که، سازمان‌های متولی مراقبت‌های بهداشتی و سلامت همواره حجم زیادی از اطلاعات را جمع‌آوری می‌کنند در حالی که این اطلاعات و داده‌ها به درستی مورد استفاده قرار نمی‌گیرند. این مطالعه نشان می‌دهد با کشف الگوها و روابط پنهان در این داده‌ها می‌توان از آن‌ها به درستی در جهت بهبود کیفیت ارائه خدمات تشخیصی و درمانی بهره برد.

تعارض منافع

در انجام این پژوهش هیچ‌گونه تعارض منافعی برای نویسندگان وجود نداشته است.

References

1. Eastell R. Treatment of postmenopausal osteoporosis. N Engl J Med 1998;338(11):736-46. doi: 10.1056/NEJM199803123381107https://www.jri.ir/documents/fullpaper/fa/170.pdf

2. World Health Organization (WHO). Assessment of fracture risk and its application to screening for postmenopausal osteoporosis: report of a WHO study group [meeting held in Rome from 22 to 25 June 1992]. Available from: https://apps.who.int/iris/bitstream/handle/10665/39142/WHO_TRS_843.pdf?sequence=1 World Health Organization. Assessment of fracture risk and its application to screening for postmenopausal osteoporosis: report of a WHO study group [meeting held in Rome from 22 to 25 June 1992]. World Health Organization; 1994.

3. Seeman E, Hopper JL, Bach LA, Cooper ME, Parkinson E, McKay J, et al. Reduced bone mass in daughters of women with osteoporosis. N Engl J Med 1989;320(9):554-8. doi: 10.1056/NEJM198903023200903

4. Slemenda CW, Christian JC, Williams CJ, Norton JA, Johnston CC. Genetic determinants of bone mass in adult women: a reevaluation of the twin model and the potential importance of gene interaction on heritability estimates. J Bone Miner Res 1991;6(6):561-7. doi: 10.1002/jbmr.5650060606

5. Bonjour JP, Theintz G, Law F, Slosman D, Rizzoli R. Peak bone mass. Osteoporos Int 1994;4:7-13. doi: 10.1007/BF01623429

6. Saadi HF, Reed RL, Carter AO, Qazaq HS, Al Suhaili AR. Bone density estimates and risk factors for osteoporosis in young women. East Mediterr Health J 2001;7(4-5):730-7.

7. Ruiz JC, Mandel C, Garabedian M. Influence of spontaneous calcium intake and physical exercise on the vertebral and femoral bone mineral density of children and adolescents. J Bone Miner Res 1995;10(5):675-82. doi: 10.1002/jbmr.5650100502

8. Rizzoli R, Bonjour JP, Ferrari SL. Osteoporosis, genetics and hormones. J Mol Endocrinol 2001;26(2):79-94. doi: 10.1677/jme.0.0260079

9. Bass S, Pearce G, Bradney M, Hendrich E, Delmas PD, Harding A, Seeman E. Exercise before puberty may confer residual benefits in bone density in adulthood: studies in active prepubertal and retired female gymnasts. J Bone Miner Res 1998;13(3):500-7. doi: 10.1359/jbmr.1998.13.3.500.

10. Ardeshir Larijani MB, Mohajeri Tehrani MR., Soltani Z, Pazhouhi M. Mosteoporosis, Prevention, Diagnosis and Treatment. Journal of Reproduction and Infertility 2005; 6(1): 5-24. [In Persian]

[Downloaded from jhbmi.com on 2026-01-30]

11. Sambrook PN, Dequiker J, Rasp HH. Metabolic bone disease, Report of a WHO study group, Assessment of fracture risk and its application to screening for postmenopausal Osteoporosis. Geneva: WHO Technical Report Series; 1994.
12. Lakshmi KV, Padmavathamma M. Modeling an expert system for diagnosis of gestational diabetes mellitus based on risk factors. Journal of Computer Engineering 2013;8(3):29-32.
13. Stafford GC, Kelley PE, Syka JE, Reynolds WE, Todd JF. Recent improvements in and analytical applications of advanced ion trap technology. International Journal of Mass Spectrometry and Ion Processes 1984;60(1):85-98. https://doi.org/10.1016/0168-1176(84)80077-4
14. Vijiyarani S, Sudha S. Disease prediction in data mining technique–a survey. International Journal of Computer Applications & Information Technology 2013;2(1):17-21.
15. Moudani W, Shahin A, Chakik F, Rajab D. Intelligent predictive osteoporosis system. International Journal of Computer Applications 2011;32(5):28-37.
16. Zhou XH, Li SL, Tian F, Cai BJ, Xie YM, Pei Y, et al. Building a disease risk model of osteoporosis based on traditional Chinese medicine symptoms and western medicine risk factors. Stat Med 2012;31(7):643-52. doi: 10.1002/sim.4382
17. Sharifkhani M, Alizadeh S, Abbasi M, Ameri H. Providing a Model for Predicting the Risk of Osteoporosis Using Decision Tree Algorithms. J Mazandaran Univ Med Sci 2014; 24(116):110-18. [In Persain]
18. Sathawane MK, Tuteja MR. Data Mining in Clinical Records to foretell the risk of Osteoporosis; 2015.
19. Nahar N, Ara F. Liver disease prediction by using different decision tree techniques. International Journal of Data Mining & Knowledge Management Process 2018;8(2): 1-9. doi: 10.5121/ijdkp.2018.8201
20. El_Jerjawi NS, Abu-Naser SS. Diabetes prediction using artificial neural network. International Journal of Advanced Science and Technology 2018; 21: 55-64. doi:10.14257/ijast.2018.121.05
21. Fitzpatrick LA. Secondary causes of osteoporosis. Mayo Clin Proc 2002;77(5):453-68. doi: 10.4065/77.5.453
22. Bogoch ER, Elliot-Gibson V, Wang RY, Josse RG. Secondary causes of osteoporosis in fracture patients. J Orthop Trauma 2012;26(9):e145-52. doi: 10.1097/BOT.0b013e3182323f2c
23. Raisz LG. Screening for osteoporosis. N Engl J Med 2005;353(2):164-71. doi: 10.1056/NEJMcp042092
24. Mäkitie RE, Costantini A, Kämpe A, Alm JJ, Mäkitie O. New insights into monogenic causes of osteoporosis. Front Endocrinol (Lausanne) 2019; 10: 70. doi: 10.3389/fendo.2019.00070
25. Zanker J, Duque G. Osteoporosis in older persons: old and new players. J Am Geriatr Soc 2019;67(4):831-40. doi: 10.1111/jgs.15716
26. Akin M, Eyduran E, Reed BM. Use of RSM and CHAID data mining algorithm for predicting mineral nutrition of hazelnut. Plant Cell, Tissue and Organ Culture (PCTOC) 2017;128(2):303-16. doi: 10.1007/s11240-016-1110-6
27. Milanović M, Stamenković M. CHAID decision tree: Methodological frame and application. Economic Themes 2016;54(4):563-86. doi: 10.1515/ethemes-2016-0029
28. Akin M, Eyduran E, Reed BM. Use of RSM and CHAID data mining algorithm for predicting mineral nutrition of hazelnut. Plant Cell, Tissue and Organ Culture (PCTOC). 2017;128(2):303-16. doi: 10.1007/s11240-016-1110-6
29. Zarkogianni K, Vazeou A, Mougiakakou SG, Prountzou A, Nikita KS. An insulin infusion advisory system based on autotuning nonlinear model-predictive control. IEEE Trans Biomed Eng 2011;58(9):2467-77. doi: 10.1109/TBME.2011.2157823
30. Su MC. Use of neural networks as medical diagnosis expert systems. Computers in Biology and Medicine 1994;24(6):419-29. https://doi.org/10.1016/0010-4825(94)90040-X
31. Jaafar SF, Ali DM. Diabetes mellitus forecast using artificial neural network (ANN). Asian Conference on Sensors and the International Conference on New Techniques in Pharmaceutical and Biomedical Research; 2005 Sep 5-7; Kuala Lumpur, Malaysia, Malaysia: IEEE; 2005. p. 135-9. doi: 10.1109/ASENSE.2005.1564523
32. Zurada JM. Introduction to Artificial Neural Systems. 1st ed.West Group; 1992.
33. Amiri Chaijan R, Khosh Taghaza M, Montazer GH, Minaee S, Alizadeh M. Estimation of head rice yield using artificial neural networks for fluidized bed drying of rough rice. Isfahan University of Technology-Journal of Crop Production and Processing 2009;13(48):285-98. [In Persian]
34. Wesolowski M, Suchacz B. Artificial neural networks: theoretical background and pharmaceutical applications: a review. J AOAC Int 2012;95(3):652-68. doi: 10.5740/jaoacint.sge_wesolowski_ann
35. Chatterjee SP, Pandya AS. Artificial Neural Networks in Drug Transport Modeling and Simulation–II. Artificial Neural Network for Drug Design, Delivery and Disposition 2016; 243-61. https://doi.org/10.1016/B978-0-12-801559-9.00012-0
36. Haykin S, Lippmann R. Neural networks, a comprehensive foundation. International Journal of Neural Systems 1994;5(4):363-4.
37. Wang W, Richards G, Rea S. Hybrid data mining ensemble for predicting osteoporosis risk. Engineering in Medicine and Biology 27th Annual Conference; 2006 Jan 17-18; Shanghai, China: IEEE; 2006. p. 886-9. doi: 10.1109/IEMBS.2005.1616557
38. Gao Z, Hong W, Xu Y, Zhang T, Song Z, Liu J. Osteoporosis diagnosis based on the multifractal spectrum features of micro-ct images and c4. 5 decision tree. First International Conference on Pervasive Computing, Signal Processing and

[Downloaded from jhbmi.com on 2026-01-30]

Applications; 2010 Sep 17-19; Harbin, China: IEEE; 2010. p. 1043-47. doi: 10.1109/PCSPA.2010.257

39. Englebardt SP, Nelson R. Health Care Informatics: An Interdisciplinary Approach. 1st ed. St. Louis: Mosby; 2002.

[Downloaded from jhbmi.com on 2026-01-30]

Original Article

Journal of Health and Biomedical Informatics
Medical Informatics Research Center
2020; 7(3): 304-317

Predicting the Risk of Osteoporosis Using Decision Tree and Neural Network

Abedini Amin[1], Jabarpour Efat[2*], Keshtkar Abbasali[3]

• **Received: 02 Nov 2019** • **Accepted: 03 Feb 2020**

Introduction: Osteoporosis is one of the major causes of disability and death in elderly people. The objective of this study was to determine the factors affecting the incidence of osteoporosis and provide a predictive model to accelerate diagnosis and reduce costs.
Method: In this fundamental descriptive study, a new model was proposed to identify the factors affecting osteoporosis. Data related to 4083 women were investigated with Clementine12, the data mining tool, to discover knowledge. Using data mining algorithms, including decision tree and artificial neural network, some rules were extracted that can be used as a model to predict the condition of patients and finally, the accuracy of the proposed models were compared.
Results: This study examined several models on a number of different characteristics and compared the results in terms of accuracy to find the best predictive model. The classification accuracy of the MLP neural network model was 92.14% which was higher than that of the other algorithms used in this study. According to the identification of factors affecting osteoporosis, the risk of developing this disease can be predicted for a new sample.
Conclusion: Healthcare organizations are always gathering a lot of information while this data is not used properly. This study showed that the hidden patterns and relationships in this data can be discovered and used to improve the quality of diagnostic and treatment services.

Keywords: Osteoporosis, Data Mining, Decision Tree, Artificial Neural Network, Clementine

• **Citation:** Predicting the Risk of Osteoporosis Using Decision Tree and Neural Network. Journal of Health and Biomedical Informatics 2020; 7(3): 304-17. [In Persian]

1. M.Sc. in Artificial Intelligence, Computer and IT Engineering, Islamic Azad University of Qazvin, Qazvin, Iran
2. Ph.D. Candidate in Industrial Management, Faculty of Management, University of Tehran, Tehran, Iran
3. Associate Professor, Health Sciences Education Development Dept., Faculty of Public Health, Tehran University of Medical Sciences, Tehran, Iran
***Corresponding Author**: Efat Jabarpour
Address: Ministry of Science, Research, and Technology
• **Tel:** 021-82234187 • **Email:** ejabarpour@yahoo.com

[Downloaded from jhbmi.com on 2026-01-30]