مقاله پژوهشی

مجله انفورماتیک سلامت و زیست پزشکی
مرکز تحقیقات انفورماتیک پزشکی
دوره هشتم شماره سوم. ص ۳۱۴-۳۰۴، ۱۴۰۰

ارائه مدل بهینه تهیه دارو با استفاده از داده‌کاوی

آزیتا کوهستانی[۱]، امیراشکان نصیری‌پور[۲]*، مهدی ریاحی فر[۲]

• دریافت مقاله: ۱۴۰۰/۷/۶ • پذیرش مقاله: ۱۴۰۰/۹/۱۷

مقدمه: کمبود منابع مالی و نقدینگی معضل اصلی بیمارستان‌ها می‌باشد. داروخانه‌ها از بخش‌های تأثیرگذار بر گردش مالی بیمارستان‌ها هستند که به دلیل عدم پیش‌بینی مصرف و تهیه دارو، در انتهای سال با اضافه موجودی، حجم زیاد داروهای تاریخ مصرف گذشته و گاهی کمبود دارو مواجه می‌شوند؛ لذا پیش‌بینی مصرف دارو با استفاده از داده‌های گذشته‌نگر موجود، منجر به بهبود مدیریت منابع در بیمارستان‌ها می‌شود. به دلیل توانمندی بالای داده‌کاوی در مدل‌سازی مسائل پزشکی، از الگوریتم‌های منتخب برای تعیین مدل بهینه تهیه دارو استفاده گردید.

روش: در این مطالعه مقطعی، برای بررسی انواع الگوریتم‌های داده‌کاوی، فرم اطلاعاتی بر اساس اهداف طراحی، سپس در قالب گزارش در سیستم اطلاعات بیمارستانی تعریف گردید و با کمک نرم‌افزار Crystal Report داده‌ها استخراج گردید. برای ارائه مدل، دقت الگوریتم‌های پیش‌بینی داده‌کاوی KNN, SVM, NN, Random Forest, LR, Adaboost بر اساس معیارهای MSE, RMSE, MAE, R2 در نرم‌افزار Weka بررسی شدند.

نتایج: روش Adaboost با معیارهای MAE, R2 و RMSE (۰/۷۸، ۲۴۷، ۸۲۷) و روش جنگل تصادفی با معیارهای (۰/۶، ۱۱۷۰، ۱۸۶۸) نسبت به بقیه مدل‌ها بالاترین دقت را داشته و میزان خطا را به نسبت بیشتری کاهش می‌دهند. سایر روش‌ها با معیارهای فوق در پیش‌بینی مسئله پژوهش عملکرد ضعیف‌تری را دارند.

نتیجه‌گیری: نتایج پژوهش حاکی از دقت بالای روش Adaboost و جنگل تصادفی در مقایسه با سایر روش‌های مورد بررسی است. درصد کمی از بیمارستان‌ها برای مدیریت تهیه دارو، برنامه‌ریزی می‌کنند؛ لذا به مدیران بیمارستان‌ها و داروخانه‌ها پیشنهاد می‌شود که در مدیریت واحدهای متبوع‌شان از داده‌کاوی استفاده کنند.

کلیدواژه‌ها: دارو، بیمارستان، داده‌کاوی، الگوریتم‌های پیش‌بینی

•ارجاع: کوهستانی آزیتا، نصیری‌پور امیراشکان، ریاحی فر مهدی. ارائه مدل بهینه تهیه دارو با استفاده از داده‌کاوی مجله انفورماتیک سلامت و زیست پزشکی ۱۴۰۰؛ ۸(۳): ۱۴-۳۰۴.

۱. کارشناسی ارشد مدیریت خدمات بهداشتی و درمانی، گروه مدیریت خدمات بهداشتی و درمانی، واحد الکترونیکی، دانشگاه آزاد اسلامی، تهران، ایران

۲. دکترای مدیریت خدمات بهداشتی و درمانی، دانشیار، گروه مدیریت خدمات بهداشتی و درمانی، واحد الکترونیکی، دانشگاه آزاد اسلامی، تهران، ایران

*** نویسنده مسئول:** امیراشکان نصیری‌پور

آدرس: تهران، خیابان پاسداران، نیستان نهم، نبش خیابان رام، پلاک ۵، طبقه سوم، دانشکده مدیریت، گروه مدیریت خدمات بهداشتی و درمانی

• شماره تماس: ۰۹۱۲۴۱۳۹۵۴۱ **• Email:** nasiripour@srbiau.ac.ir

[Downloaded from jhbmi.com on 2026-05-24]

مقدمه

تأمین سلامتی به عنوان یکی از چالش‌های مهم زندگی بشر و درمان بیماری‌ها یکی از راه‌های دستیابی به سلامتی است. در این راستا، دارو یکی از ابزارهای مهم در کسب سلامتی و حفظ و ارتقاء آن می‌باشد. تهیه، تأمین و توزیع به‌موقع و صحیح دارو از اهمیت زیادی برای حفظ سلامتی برخوردار است، به طوری‌که بر اساس تصمیمات کنفرانس آلماتا (Alma-Ata)، در سال ۱۹۸۷، تأمین داروهای اساسی از سوی سازمان بهداشت جهانی به عنوان جزء هشتم مراقبت‌های بهداشتی اولیه معرفی گردید. توجه به سیستم تأمین، نگهداری و توزیع منظم دارو امر مهمی است، چرا که وجود یک سیستم کارآمد، دسترسی افراد به خدمات دارویی را ممکن ساخته و بدین ترتیب از مراجعه بخش عظیمی از مردم به مراکز خصوص جلوگیری می‌شود. از دیدگاه سازمان جهانی بهداشت، مدیریت ضعیف دارو در بخش دولتی، به خصوص در کشورهای درحال‌توسعه موضوع تعیین کننده‌ای است که بهینه‌سازی وضعیت آن می‌تواند به حفظ سرمایه این کشورها کمک کند [۱]. همچنین برنامه‌ریزی کارا برای منابع و هزینه‌ها در بیمارستان‌ها نقشی حیاتی در مدیریت بیمارستان و بخش‌های آن ایفا می‌کند [۲].

داروخانه یکی از بخش‌های مهم و درآمدزا در بیمارستان‌ها می‌باشد که کنترل و نظارت بر فرایند مصرف دارو را بر عهده داشته و دسترسی به‌موقع به دارو، امنیت در مصرف دارو و استفاده مؤثر و مقرون به صرفه از دارو را برای بیماران و کارکنان بیمارستان تضمین می‌نماید [۳] که می‌توان با بهبود خرید دارو و افزایش بهره‌وری آن، به نحوی به وضعیت اقتصادی بیمارستان‌ها کمک کرد. این حقیقت که مصرف دارو بر اساس شیوع بیماری‌ها در فصول مختلف تغییر می‌کند از جمله فاکتورهایی هستند که باید در خرید دارو در نظر گرفته شوند. [۴].

در حال حاضر کمبود منابع مالی و نقدینگی معضل اصلی بیمارستان‌ها می‌باشد و یکی از بخش‌های تأثیرگذار در وضعیت گردش مالی بیمارستان‌ها، داروخانه بیمارستان‌ها هستند. که به دلیل عدم پیش‌بینی تهیه دارو در انتهای فصل معمولاً با اضافه موجودی، حجم زیادی از داروهای تاریخ مصرف گذشته که غیرقابل استفاده هستند و در برخی از موارد کمبود برخی از داروها مواجه می‌شوند. از طرف دیگر در حال حاضر سیستم‌های اطلاعات بیمارستان، حجم زیادی از داده‌های مربوط به بیمارستان از جمله مصرف دارو را نگهداری می‌کنند، که این داده‌ها مورد پردازش قرار نمی‌گیرند [۵].

پژوهش‌های مختلفی در زمینه استفاده از تکنیک‌های داده‌کاوی در علوم پزشکی انجام گرفته است. در مطالعه‌ای تحت عنوان الگوی پیش‌بینی تقاضای سالانه داروی مصرفی ایران [۶]، کارایی دو روش پیش‌بینی شبکه عصبی مصنوعی و برازش منحنی در مقایسه با روش متداول نرخ رشد مرکب بررسی شد. نتایج نشان داد، عملکرد بهتر روش شبکه عصبی و برازش منحنی در پیش‌بینی میزان فروش ریالی دارو را نسبت به روش مرسوم نشان داد و در مواقعی که سوابق داده‌های تجربی گذشته، کم است، روش برازش منحنی، کارایی بهتری داشته؛ اما با داده‌های ورودی بیشتر، روش شبکه عصبی کارایی بهتری نسبت به دو روش دیگر خواهد داشت.

کمالی [۷] پژوهشی تحت عنوان پیش‌بینی تقاضای دارو با استفاده از شبکه‌های عصبی مصنوعی (Artificial Neural Network)ANN انجام داد. نتایج حاکی از آن است که روش شبکه عصبی مصنوعی در مقایسه با روش کلاسیک میانگین متحرک خودهمبسته یکپارچه (Autoregressive Integrated Moving)ARIMA دارای عملکرد بهتری بوده و قادر است میزان تقاضای دارو را دقیق‌تر پیش‌بینی نماید. صالح‌آبادی [۸]، تکنیک شبکه‌های عصبی را در یکی از شرکت‌های پخش دارو به منظور دسته‌بندی مشتریان صنعت پخش دارو به کار برد. طبق نتایج، مشتریان به سه بخش با ویژگی‌های خاص تفکیک شدند. Sadarina و همکاران [۹]، تکنیک‌های داده‌کاوی را برای بازاریابی محصولات دارویی پیاده‌سازی نمودند که نتایج نشان داد الگوریتم‌های درخت تصمیم و روش طبقه‌بندی، در گرفتن تصمیمات مهم در بازاریابی دارویی کمک کرده است.

با توجه به مشکلات مالی که گریبان‌گیر بیمارستان‌های کشور است. هم‌چنین اهمیت و نقشی که داروخانه‌ها در گردش مالی و درآمد بیمارستان‌های کشور دارند. هدف این پژوهش پیش‌بینی مصرف دارو با استفاده از تکنیک‌های داده‌کاوی، در داروخانه یک بیمارستان بزرگ به منظور افزایش بهره‌وری مالی آن است. همچنین با استفاده از داده‌های گذشته‌نگر و تکنیک‌های داده‌کاوی بتوان مدلی برای پیش‌بینی مصرف دارو ارائه داد. در واقع ارائه یک مدل مبتنی بر تکنیک‌های داده‌کاوی که بتوان از آن برای پیش‌بینی تهیه دارو استفاده کرد. به همین دلیل باید روش‌های مختلف پیش‌بینی را اعمال کرد تا بتوان مقایسه‌ای

[Downloaded from jhbmi.com on 2026-05-24]

بین روش‌ها و انتخاب بهترین روش داشته باشیم؛ لذا سؤال اصلی پژوهش این است که مدل بهینه تهیه دارو با تکنیک داده‌کاوی در بیمارستان هدف کدام است؟ چرا که تهیه بهینه دارو از یک سو افزایش بهره‌وری مالی بیمارستان و از سوی دیگر پاسخگویی بیشتر در برابر نیازهای درمانی بیماران را به دنبال خواهد داشت.

روش

پژوهش حاضر با توجه به هدف از نوع پژوهش‌های کاربردی، توصیفی و به صورت مقطعی و کمی است. مراحل اصلی روش پژوهش در سه بخش تشریح می‌گردد.

مرحله ۱: به منظور بررسی وضعیت سیستم اطلاعات دارویی جهت استخراج داده‌ها ابتدا فرم ثبت اطلاعات شامل عناصر اطلاعاتی (سال، کد دارو، نام دارو، میزان مصرف ماهانه و فصلی دارو، جمع مصرفی دارو) توسط پژوهشگر طراحی گردید. سپس در بررسی میدانی سیستم اطلاعات بیمارستانی، نیاز به طراحی عناصر اطلاعاتی در قالب گزارش احساس شد، لذا گزارش فوق طراحی گردید و سپس با استفاده از نرم‌افزار Crystal Reports داده‌های دارویی ذخیره شده در پایگاه داده SQL SERVER سیستم اطلاعات بیمارستان به تفکیک سال‌های ۱۳۹۸ و ۱۳۹۹ در قالب فایل اکسل استخراج گردید.

مرحله ۲: در این پژوهش، با توجه به ماهیت داده‌ها و هدف آن، در مرحله پیش‌پردازش و آماده‌سازی داده‌ها، عملیات پاک‌سازی داده‌ها، انتخاب ویژگی، نمونه‌برداری و تبادل داده‌ها ضروری به‌نظر رسید. این کار در نرم‌افزار Weka، با استفاده از ابزار پیش‌پردازش انجام گردید.

برای دستیابی به نتیجه دقیق و منسجم، داده‌ها به صورت فصلی جدا گردید. در روش فصول متوالی، تقاضای دارو در قالب پنجره‌ای با سایز چهار، از فصول متوالی (بهار، تابستان، پاییز و زمستان) در آن نگهداری شده است. در پایگاه داده، داروهایی موجود هستند که ممکن است میزان تقاضای آن‌ها در بعضی از سال‌ها بسیار پایین یا صفر باشد و یا این که در تمام سال تقاضای تقریباً یکسانی داشته باشند. این داده‌ها نمی‌توانند در روند اجرای مسئله پیش‌بینی تأثیر بسزایی داشته باشند به این دلیل که واریانس این داده‌ها پایین است و میزان تغییرات آن‌ها در ماه‌ها، فصول و سال‌های مختلف بسیار پایین است. بین ترتیب داده‌هایی که تقاضای آن‌ها بیشترین واریانس را داشتند به عبارتی در فصول مختلف بیشترین تکرار را داشتند به عنوان مجموعه داده پایه وارد نرم‌افزار شدند.

مرحله ۳: در این مرحله، الگوریتم‌های انتخاب شده شامل: جنگل تصادفی، ماشین بردار پشتیبان، شبکه عصبی، رگرسیون خطی، k نزدیک‌ترین همسایه و Adaboost به ترتیب در محیط نرم‌افزار Weka بر روی مجموعه داده‌ها اعمال گردید و خطای پیش‌بینی با استفاده از چهار معیار ارزیابی MSE (Mean Squared Error)، MAE (Mean Absolute Error)، RMSE (Root Mean Squared Error)، R^2 محاسبه گردید و مدل‌ها با استفاده از آن‌ها مورد مقایسه و بررسی قرار گرفتند و در نهایت، دانش کشف شده به شیوه قابل درک تبدیل و ارائه شد. در عمل روشی مطلوب است که به طور هم‌زمان ضریب تشخیص آن بالاتر ($R^2>۰/۶$) و خطای آن (با توجه به ماهیت داده‌ها) کمتر باشد؛ به‌عبارت‌دیگر کم بودن مقدار RMSE و بالا بودن ضریب R^2 بیانگر دقت قابل قبول مدل و برتری آن نسبت به مدل‌های دیگر است. MAE و MSE هر کدام به شیوه‌ای اختلاف بین داده‌های مشاهداتی و محاسباتی را نشان می‌دهند که هر چه این مقادیر کمتر باشند یعنی کار نتیجه بخش‌تر بوده است.

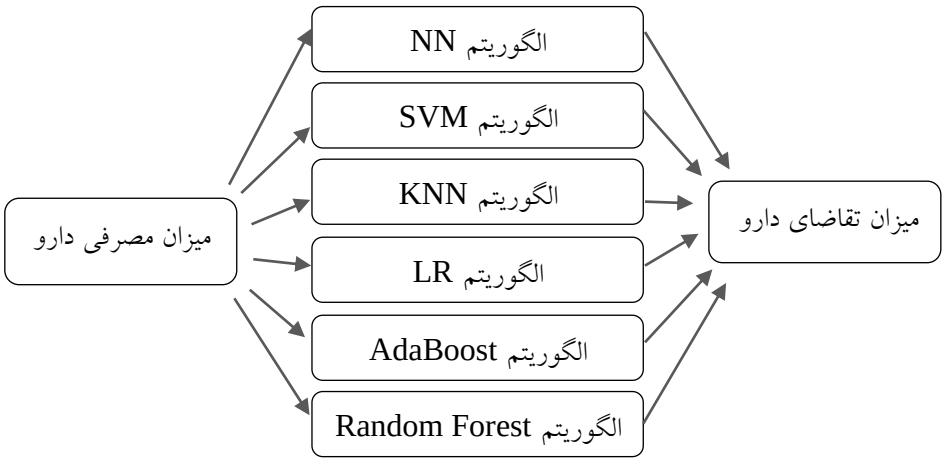
داده‌کاوی از الگوریتم‌های متعدد یادگیری به منظور تولید مدل‌های پیش‌بینی استفاده می‌کند. نتایج مطالعه رضایی نشان داد که هر سه الگوریتم Decision Tree, SVM, ANN قادر به پیش‌بینی مدت اقامت بیماران قلبی با درجات مختلف دقت هستند؛ اما یافته‌ها نشان داد که الگوریتم SVM بهترین مدل است [۱۰]. مقدسی و همکاران نیز بیان کردند، سازمان‌هایی که از داده‌کاوی به منظور بهبود کیفیت خدمات سلامت استفاده می‌کنند سریع‌تر به قله‌های موفقیت دست خواهند یافت و می‌توانند در جهت پیشگیری، تشخیص، درمان و نیز ارائه خدمات با کیفیت به مشتریان سلامت گام‌های جدیدی را بردارند [۱۱].

پژوهش Yeh و همکاران که به منظور پیش‌بینی بیماری‌های عروق مغزی با استفاده از تکنیک‌های داده‌کاوی انجام شد نیز نشان داد که الگوریتم درخت تصمیم به عنوان بهترین مدل پیش‌بینی برای بیماری‌های عروق مغزی می‌باشد [۱۲]. نتایج پژوهش مهدی‌پور و همکاران حاکی از قدرت بالای مدل شبکه عصبی برای پیش‌بینی سکته مغزی با دقت ۹۳/۷ درصد بود [۱۳]. مظاهری و همکاران با هدف ارائه مدلی برای پیش‌بینی روش درمانی مناسب بیماری قلبی به منظور کاهش هزینه‌های

[Downloaded from jhbmi.com on 2026-05-24]

درمان و ارائه خدمات با کیفیت بهتر توسط پزشکان، به عملکرد بهتر الگوریتم‌های درخت تصمیم نسبت به شبکه عصبی در این زمینه دست یافتند [۱۴].

با توجه به نتایج دست یافته در پژوهش‌های پیشین و وضعیت موجود، در این پژوهش نیز بر اساس مقایسه روش‌های داده‌کاوی، مدل مفهومی پژوهش به شکل ذیل ارائه گردید.



شکل ۱: مدل مفهومی پژوهش

نتایج

نتایج اجرای الگوریتم شبکه عصبی

همان‌طور که نتایج جدول ۱ نشان می‌دهد، در روش NN بر روی مجموعه داده‌ها، نوسانات تقاضا بسیار بالا است و این باعث می‌شود پیش‌بینی بسیار مشکل شود. همان‌طور که مشاهده می‌شود بالاترین مقدار R^2 مربوط به فصل بهار (۰/۷۶۴) است که نشان از پیش‌بینی نسبتاً دقیق این روش در فصل بهار است و مقادیر R^2 برای فصول تابستان و پاییز به ترتیب ۰/۰۵و ۰/۲۹۴ به دست آمد که حاکی از این است که این روش در فصول فوق، دقت پیش‌بینی بالایی را ندارد و در کل این روش به جزء در فصل بهار در سایر فصول دقت پیش‌بینی بالایی را نشان نمی‌دهد.

جدول ۱: نتایج اجرای الگوریتم NN در روش فصول متوالی

فصول / معیار	MSE	RMSE	MAE	R^2
بهار	۲۲۲۵۲۶۳۰/۸۳۲	۴۷۱۷/۲۶۹	۳۱۰۸/۵۶۴	۰/۷۶۴
تابستان	۴۶۳۲۳۴۸/۷۲۱	۲۱۵۲/۲۸۹	۴۸۹/۱۰۷	۰/۰۵۰
پاییز	۱۶۵۸۴۰۴۳/۲۷۷	۴۰۷۲/۳۵۱	۱۹۵۴/۵۹۸	۰/۲۹۴
زمستان	۱۵۹۴۰۷۰۲/۳۵۲	۳۹۹۲/۵۸۱	۲۴۴۰/۱۱۸	۰/۵۹۰

نتایج اجرای الگوریتم آدابوست

بر اساس نتایج حاصل در جدول ۲، مقادیر R^2 در الگوریتم Adaboost، در سه فصل بهار، زمستان و پاییز نشان از توان بالای این روش برای پیش‌بینی می‌باشد و تنها در فصل تابستان دارای دقت پایین می‌باشد که می‌توان آن را به دلیل تغییر مصرف دارو به دنبال تغییر شیوع بیماری‌ها در فصول مختلف دانست. از سوی دیگر خطاهای پیش‌بینی نیز اعداد کوچکی را در مقایسه با سایر الگوریتم‌ها نشان می‌دهد؛ بنابراین روش Adaboost خطا را به میزان خیلی زیادی کاهش داده و دقت پیش‌بینی خوبی را نشان می‌دهد.

[Downloaded from jhbmi.com on 2026-05-24]

جدول ۲: نتایج اجرای الگوریتم Adaboost در روش فصول متوالی

فصول	MSE	RMSE	MAE	R^2
بهار	۳۴۷۸۵۰۰۰	۱۸۶/۵۰۷	۵۸/۹۳۳	۰/۹۹۷
تابستان	۵۰۷۵۱۷۵/۸۶۷	۲۲۵۲/۸۱۵	۶۵۳/۸۰۰	۰/۱۵۰
پاییز	۴۹۸۷۹۱/۸۶۷	۷۰۶/۲۵۲	۲۳۰/۴۰۰	۰/۹۶۱
زمستان	۲۷۳۹۵/۶۳۳	۱۶۵/۵۱۶	۴۵/۲۳۳	۰/۹۹۷

نتایج اجرای الگوریتم k نزدیک‌ترین همسایه

یافته‌های پژوهش در جدول ۳ نشان می‌دهد، در روش KNN، نیز مقادیر R^2 در هر چهار فصل کمتر از ۰/۶ را نشان می‌دهد و این مطلب بیانگر دقت پایین این روش در امر پیش‌بینی مسئله پژوهش است. از سوی دیگر به همان نسبت نیز خطاهای پیش‌بینی مقادیر بزرگی را نشان می‌دهد؛ بنابراین این روش از دقت لازم بری پیش‌بینی برخوردار نیست و خطای پیش‌بینی آن خیلی بالا می‌باشد.

جدول ۳: نتایج اجرای الگوریتم KNN در روش فصول متوالی

فصول / معیار	MSE	RMSE	MAE	R^2
بهار	۱۰۰۵۳۳۹۵/۲۴۰	۳۱۷۰/۷۰۹	۲۳۳۴/۲۰۰	۰/۲۰۳
تابستان	۳۶۴۲۵۹۹/۳۶۰	۱۹۰۸/۵۵۹	۷۴۹/۵۴۷	۰/۱۷۵
پاییز	۱۰۸۰۳۴۹۸/۰۳۷	۳۲۸۶/۸۶۸	۲۲۰۷/۲۰۰	۰/۱۵۷
زمستان	۷۷۰۰۷۸۱/۶۹۲	۲۷۷۵/۰۲۸	۲۰۳۵/۵۸۰	۰/۲۳۲

نتایج اجرای الگوریتم رگرسیون خطی

بر اساس یافته‌ها در جدول ۴، مقادیر R^2 در الگوریتم LR، نیز در تمامی فصول کمتر از ۰/۶ می‌باشد و مقادیر معیارهای خطای پیش‌بینی نیز بسیار بالا می‌باشند؛ بنابراین این روش نیز نتوانسته با توجه به مجموعه داده‌های پژوهش دقت بالایی را در امر پیش‌بینی نشان دهد.

جدول ۴: نتایج اجرای الگوریتم LR در روش فصول متوالی

فصول / معیار	MSE	RMSE	MAE	R^2
بهار	۹۹۲۰۱۴۴/۰۸۳	۳۱۴۹/۶۲۶	۲۳۴۵/۴۶۲	۰/۲۱۴
تابستان	۳۱۹۷۹۶۹/۹۴۵	۱۷۸۸/۲۸۷	۹۲۸/۵۲۷	۰/۲۷۵
پاییز	۱۱۰۰۲۴۸۳/۴۱۱	۳۳۱۶/۹۹۹	۲۰۹۶/۷۴۸	۰/۱۴۲
زمستان	۸۴۰۶۲۴۸/۸۹۴	۲۸۹۹/۳۵۳	۲۱۹۹/۹۲۸	۰/۱۶۲

نتایج اجرای الگوریتم ماشین بردار پشتیبان

همان‌طور که نتایج جدول ۵ نشان می‌دهد، در روش SVM بر روی مجموعه داده‌ها، مقادیر R^2 در هر چهار فصل بسیار پایین و کمتر از ۰/۶ می‌باشد و مقادیر معیارهای خطا نیز اعداد بالایی را نشان می‌دهد؛ لذا با توجه به این که مدل مناسب باید دارای ضریب تعیین بیشتر از ۰/۶ و خطای پیش‌بینی کم باشد، لذا این روش نیز بر روی مجموعه داده‌های مورد بررسی پژوهش، عملکرد خوبی را نشان نمی‌دهد.

[Downloaded from jhbmi.com on 2026-05-24]

جدول ۵: نتایج اجرای الگوریتم SVM در روش فصول متوالی

فصول معیار	MSE	RMSE	MAE	R^2
بهار	۱۳۰۷۵۹۸۲/۶۰۴	۳۶۱۶/۰۷۳	۲۲۹۹/۸۸۱	۰/۰۳۷
تابستان	۴۶۴۷۹۸۳/۴۳۰	۲۱۵۵/۹۱۸	۴۸۵/۱۵۲	۰/۰۵۳
پاییز	۱۶۶۳۴۴۳۳/۰۸۸	۴۰۷۸/۵۳۳	۱۹۵۴/۲۸۹	۰/۲۹۸
زمستان	۱۲۲۷۸۷۹۳/۹۹۵	۳۵۰۴/۱۱۱	۲۳۴۷/۸۴۲	۰/۲۲۵

نتایج اجرای الگوریتم جنگل تصادفی

بر اساس نتایج نشان داده شده در جدول ۶، مقادیر R^2 در الگوریتم RF، در سه فصل زمستان، بهار، پاییز به ترتیب ۰/۷۲۸، ۰/۷۲۵، ۰/۶۰۳ و بالاتر از ۰/۶ است. در این روش نیز تنها در فصل تابستان شاهد ضریب تعیین پایین هستیم، که این مطلب را که مصرف دارو تحت تأثیر شیوع بیماری‌ها با توجه به فصول سال، تغییر می‌کند را قوت می‌بخشد. هم‌چنین با توجه به پایین بودن معیارهای خطا، بنابراین این روش توانسته است با دقت بالا پیش‌بینی را انجام دهد و خطای پیش‌بینی را نیز کاهش دهد.

جدول ۶: نتایج اجرای الگوریتم RF در روش فصول متوالی

فصول / معیار	MSE	RMSE	MAE	R^2
بهار	۳۴۶۴۱۳۷/۴۱۶	۱۸۶۱/۲۱۹	۱۲۹۸/۹۰۴	۰/۷۲۵
تابستان	۲۹۰۴۸۷۷/۳۷۷	۱۷۰۴/۳۷۰	۶۱۵/۴۵۰	۰/۳۴۲
پاییز	۵۰۹۱۹۵۸/۵۴۳	۲۲۵۶/۵۳۷	۱۵۰۷/۶۰۴	۰/۶۰۳
زمستان	۲۷۲۸۸۴۹/۸۸۹	۱۶۵۱/۹۲۳	۱۲۵۸/۸۶۹	۰/۷۲۸

نتایج ارزیابی الگوریتم‌ها

نتایج نشان داده شده در جدول ۷، حاکی از برتری روش Adaboost با بیشترین مقدار R^2، (۰/۸) و به ترتیب با کمترین میزان RMSE و MAE، (۸۲۸)، (۲۴۷) نسبت به بقیه روش‌ها است. در رتبه دوم الگوریتم Random Forest، به ترتیب با مقادیر R^2، MAE، RMSE، ۰/۶، ۱۱۷۰، ۱۸۶۸ نتیجه بهتری را برای پیش‌بینی نشان می‌دهد. بسته به نوع داده‌ها، هر کدام از این الگوریتم‌ها می‌تواند نتیجه متفاوتی را ارائه نماید.

جدول ۷: نتایج ارزیابی الگوریتم‌ها

الگوریتم / معیار	MSE	RMSE	MAE	R^2
NN	۱۴۸۵۲۴۳۱/۲۹۶	۳۷۳۳/۶۲۳	۱۹۹۸/۰۹۷	۰/۴۲۵
Adaboost	۱۴۰۹۰۳۷/۰۹۲	۸۲۷/۷۲۵	۲۴۷/۰۹	۰/۷۸
KNN	۸۰۵۰۰۶۸/۵۸۲	۲۷۸۵/۲۹۱	۱۸۳۱/۶۳۲	۰/۱۹۲
LR	۸۱۳۱۷۱۱/۵۸۳	۲۷۸۸/۵۶۶	۱۸۹۲/۶۶۶	۰/۱۹۸
SVM	۱۱۶۵۹۲۹۸/۲۷۹	۳۳۳۸/۶۵۹	۱۷۷۱/۷۹۱	۰/۱۵۳
Random Forest	۳۵۴۷۴۵۵/۸۰۶	۱۸۶۸/۵۱۲	۱۱۷۰/۲۰۷	۰/۶۰۰

[Downloaded from jhbmi.com on 2026-05-24]

مدل نهایی پژوهش

بدین ترتیب با توجه به یافته‌های پژوهش و با توجه به این‌که در عمل روشی مطلوب است که به طور همزمان ضریب تشخیص آن بالاتر (R^2>۰/۶) و خطای آن (با توجه به ماهیت داده‌ها) کمتر باشد؛ بنابراین مدل نهایی پژوهش ارائه می‌گردد.

شکل ۲: مدل نهایی پژوهش

بحث و نتیجه‌گیری

بر اساس نتایج حاصل از این مطالعه از بین مدل‌های موردبررسی، الگوریتم Adaboost با در نظر گرفتن معیارها، در پیش‌بینی تهیه دارو توان بهتری داشته است که به برتری این الگوریتم در برخی از مطالعات نیز اشاره شده است. از جمله در مطالعه‌ای که توسط زارع و همکاران [۱۵]، به منظور پیش‌بینی عوامل خطر مغزی در کودکان صورت گرفت، عملکرد روش‌های داده‌کاوی آدابوست، بگینگ و الگوریتم C4.5 مورد مقایسه قرار گرفتند. در بین این روش‌ها مدل آدابوست، عملکرد موفق‌تری در تشخیص و مدل‌سازی داشت؛ که هم‌راستا با نتایج مطالعه حاضر می‌باشد.

دومین مدلی که عملکرد مناسب آن برای پیش‌بینی تهیه دارو تأیید گردید الگوریتم جنگل تصادفی است که نتایج مشابه را می‌توان در سایر مطالعات نیز مشاهده کرد. از جمله ایوب‌زاده و همکاران [۱۶]، مدل جنگل تصادفی را بهترین مدل برای طبقه‌بندی مدت اقامت بیماران دانستند. نتایج مطالعه ناطقی‌نیا و همکاران [۲]، نیز هم‌راستا است و دقت بسیار زیاد این الگوریتم را در پیش‌بینی مدت بستری‌های غیرضروری و اقامت طولانی‌مدت در بخش مراقبت‌های ویژه معرفی کرد. در همین راستا سعدی و همکاران [۱۷]، به نتایج مشابهی دست یافت که حاکی از دقت قابل‌قبول جنگل تصادفی در پیش‌بینی احتمال ابتلاء به بیماری قلبی و عروقی بود.

Ranjan [۱۸]، بیان کرد که مدل عصبی در پیش‌بینی فروش دارو و کشف دارو بیشترین کاربرد را دارد. نتایج مطالعه کمالی [۷]، عملکرد بهتر مدل شبکه عصبی در پیش‌بینی میزان تقاضای دارو (آنتی‌بیوتیک) را نشان داد. در مطالعه داودی و رئیسی [۶]، مدل شبکه عصبی دارای عملکرد بهتری در پیش‌بینی میزان فروش ریالی دارو نسبت به سایر روش‌ها داشت. Qingkui و Junhu [۱۹]، نیز نقش شبکه عصبی در پیش‌بینی تقاضا برای استفاده از خدمات و امکانات بیمارستان را مثبت ارزیابی نمودند. یوسفی [۲۰]، مدل شبکه عصبی را یکی از مدل‌های مناسب در پیش‌بینی عوامل مؤثر در بستری مجدد بیماران و تأثیر مثبت در کاهش نرخ بستری مجدد عنوان نمود. این در حالی است که نتایج پژوهش حاضر کاملاً متفاوت از نتایج مطالعات فوق می‌باشد و الگوریتم شبکه عصبی در پژوهش حاضر عملکرد مناسبی برای پیش‌بینی تهیه دارو نداشت که یافته‌های این پژوهش هم‌راستا با نتایج مطالعه قاضی‌سعیدی و همکاران [۲۱] که مدل شبکه عصبی را مدل ضعیفی در شناسایی و پیش‌بینی احتمال ابتلای فرد به بیماری آسم دانست. این امر را می‌توان ناشی از ماهیت متفاوت مجموعه داده‌ها و نیز تعداد کم داده‌های ورودی دانست. همان‌طور که داودی و رئیسی [۶]، نیز در مطالعه خود بیان کردند اگر چه روش شبکه عصبی در پیش‌بینی ریالی فروش دارو عملکرد بهتری از روش مرسوم دارد؛ اما با داده‌های ورودی بیشتر می‌توان به کارایی بهتری در روش شبکه عصبی دست یافت.

در الگوریتم KNN نیز نتایج مناسبی کسب نشد و این الگوریتم جهت پیش‌بینی تهیه دارو دقت بسیار پایینی در تمامی فصول در مقایسه با سایر الگوریتم‌های مورد بررسی داشت. این در حالی است که عسگری و همکاران [۲۲]، عملکرد مدل

[Downloaded from jhbmi.com on 2026-05-24]

KNN را در پیش‌بینی مرگ‌ومیر در بخش مراقبت‌های ویژه تأیید نمودند. نتایج مطالعه ناطقی‌نیا و همکاران [۲]، نیز تأیید کننده عملکرد مناسب مدل KNN در رتبه سوم مدل‌های مورد بررسی در پیش‌بینی مدت اقامت بیمار در بخش‌های ویژه بود. لعل دشتی و همکاران [۲۳]، نیز دقت الگوریتم KNN را در مقایسه با دو الگوریتم دیگر برای پیش‌بینی سندرم متابولیک بسیار بالا دانستند. تفاوت نتایج به‌دست آمده را می‌توان به دلیل ماهیت متفاوت داده‌ها و نیز محدود بودن تعداد آن‌ها دانست.

مدل رگرسیون خطی، یکی دیگر از الگوریتم‌های مورد بررسی در پژوهش حاضر بود که عملکرد آن دقت پایینی را در پیش‌بینی تهیه دارو نشان داد. نتایج پژوهش حاضر متفاوت با نتایج مطالعه عسگری و همکاران [۲۲]، بود که مدل LR عملکرد بهتری در پیش‌بینی مرگ‌ومیر در بخش مراقبت‌های ویژه نسب به سایر الگوریتم‌های مطالعه داشت. هم‌چنین ایوب زاده و همکاران [۲۴]، مدل LR را دارای عملکرد مناسب برای پیش‌بینی شیوع کووید ۱۹ در ایران دانستند. استفاده از رگرسیون خطی به دلیل سادگی در فهم و نتایج، می‌تواند روش مناسبی برای پیش‌بینی باشد، اما باید توجه داشت نتایج چنین مطالعاتی بسته به داده‌های مورد استفاده می‌تواند متفاوت باشد. مطلبی که در مورد پژوهش حاضر نیز صدق می‌کند و به نظر می‌رسد نتایج به دست آمده در این مدل را نیز می‌توان به دلیل محدودیت داده‌ها دانست و افزایش تعداد داده‌ها در سال‌های متوالی، احتمال کسب نتایج بهتر را تقویت می‌کند.

نتایج اجرای الگوریتم ماشین بردار پشتیبان، عملکرد مناسبی را در تمامی فصول و در مجموع برای پیش‌بینی تهیه دارو ندارد و این در حالی است که سایر مطالعات عملکرد مناسب این الگوریتم را در مطالعات خود بیان می‌کنند از جمله Zheng و همکاران [۲۵] الگوریتم بردار پشتیبان را با دقت ۷۸/۴ درصد نسبت به مدل جنگل تصادفی و شبکه عصبی دارای دقت بالاتری در شناسایی صحیح بیماران بستری شده و کاهش نرخ کلی بستری مجدد بیمارستان است. نتایج مطالعه عسگری و همکاران [۲۲]، نیز تأیید کننده عملکرد مناسب ماشین بردار پشتیبان در پیش‌بینی مرگ و میر در بخش مراقبت‌های ویژه است. در این راستا نتایج درمحمدی و همکاران [۲۶]، نیز حاکی از بالاترین کارایی الگوریتم SVM در تعیین علل ناباروری دارد. عملکرد مناسب ماشین بردار پشتیبان برای پیش‌بینی مدت اقامت بیماران قلبی با درجات مختلف در پژوهش رضایی و همکاران [۲۷]، نیز تکرار گردید.

نتایج این پژوهش عملکرد بهتر روش Adaboost را نسبت به سایر الگوریتم‌ها نشان داد که می‌تواند به دلیل سهولت استفاده و اصلاح پارامترهای کمتر باشد. طبق مطالعات پیشین در زمینه یادگیری ماشین، Adaboost دقت بالایی در مورد داده‌های تصویر یا متن دارد و از آنجایی که ماهیت داده‌های این مطالعه مشابه داده‌های متنی می‌باشد؛ لذا عملکرد بالای الگوریتم فوق در امر پیش‌بینی در این مطالعه قابل توجیه خواهد بود.

هم‌چنین روش جنگل تصادفی روش برتر، برای کار با داده‌های مفقود شده ارائه می‌دهد. مقادیر از دست رفته با متغیری که بیشتر در یک گروه خاص ظاهر می‌شود، جایگزین می‌شود. از آنجایی که این روش از تشکیل جنگل‌های تصادفی یا به عبارت دیگر یادگیری گروهی در امر پیش‌بینی استفاده می‌کند لذا امکان overfit شدن (بیش برازش) در آن به شدت کاهش می‌یابد که مناسب بودن عملکرد پیش‌بینی این الگوریتم در این مطالعه قابل توجیه می‌باشد.

نمی‌توان به طور قطعی راجع به برتری یک مدل اظهارنظر کرد. کارایی هر مدل می‌تواند بسته به مجموعه داده متفاوت باشد. هر کدام از این مدل‌ها ممکن است در یک مجموعه داده پیش‌بینی دقیقی داشته باشد؛ اما در مجموعه داده دیگر، کارایی ضعیفی را از خود نشان دهد.

در نهایت، نتایج این پژوهش نشان داد، بهره‌گیری از روش‌های پیش‌بینی و تکنیک‌های مشابه، با ارائه تصویری دقیق‌تر از میزان تهیه و مصرف دارو در سطح بیمارستان، امکان برنامه‌ریزی صحیح و تخصیص مناسب منابع را فراهم می‌نماید. به این ترتیب مدیران اجرایی بیمارستان می‌توانند با استفاده از الگوی پیش‌بینی تهیه دارو، برآورد دقیق‌تری از منابع خود داشته و در نهایت در کنار بهبود بهره‌وری و مدیریت بهینه تهیه دارو در سازمان خود، خدمات قابل‌قبول‌تری نیز به بیماران ارائه دهند.

در محیط رقابتی امروز، سازمان‌هایی که به‌واسطه استفاده از فن‌آوری‌های نوین هم‌چون داده‌کاوی بتوانند داده‌ها را در راستای بهبود کیفیت خدمات سلامت به کار برند، سریع‌تر به قله موفقیت خواهند رسید؛ لذا پیشنهاد می‌گردد تا سازمان‌های سلامت از این عرصه بازنمانند. در این راستا استفاده از یافته‌های پژوهش به صورت کاربردی توصیه می‌شود:

۱- با توجه به ارتقاء سیستم‌های اطلاعات بیمارستانی و ذخیره داده‌های مالی، بالینی در این سیستم‌ها، لزوم طراحی دستگاه‌های نرم‌افزاری جهت استفاده از این داده‌های خام بسیار ضروری به نظر می‌رسد؛ لذا پیشنهاد می‌شود، طراحی

[Downloaded from jhbmi.com on 2026-05-24]

سیستم‌های نرم‌افزاری به منظور استفاده از این داده‌های خام بر مبنای مدل‌های ارائه شده در این پژوهش در دستور کار مدیران ارشد بیمارستان‌ها و مسئولین دارویی داروخانه‌ها قرار گیرد.

۲- احصاء فرآیندها در بیمارستان‌ها از جمله راه‌های بهبود منابع مالی و جلوگیری از اتلاف منابع می‌باشد؛ لذا استفاده از نتایج این پژوهش جهت بهبود و کوتاه نمودن فرایندهای دارویی توصیه می‌گردد.

۳- اصلی‌ترین رکن در سازمان‌ها، نظارت می‌باشد؛ لذا استفاده از سیستم‌های نرم‌افزاری و بهبود فرآیندها اگر چه از الزامات می‌باشد، ولی کافی نمی‌باشد و بدون نظارت بر برنامه‌ها به نتایج مطلوبی دست پیدا نخواهیم کرد. در این راستا نیز استفاده از نتایج حاصل از این پژوهش برای شناسایی نقاط ضعف و مناطق قابل بهبود در زنجیره تأمین دارو توسط سازمان‌های دارویی با هدف توسعه برنامه‌های نظارتی پیشنهاد می‌گردد.

۴- به نظر می‌رسد، عدم آشنایی مدیران ارشد بیمارستان‌ها و مسئولین دارویی با فن‌آوری‌های جدید از جمله معضلات بیمارستان‌ها جهت پیاده‌سازی بسیاری از فرآیندهای علمی می‌باشد. از سوی دیگر عدم ریسک‌پذیری مدیران ارشد و تمایل آن‌ها به پیاده‌سازی روش‌هایی که توسط سایر سازمان‌ها تجربه شده و در مدت زمان کوتاهی اجرایی می‌شود نیز یکی دیگر از موانع پیش‌رو برای استفاده از نتایج مطالعات مشابه می‌باشد؛ لذا لزوم برگزاری دوره‌های توانمندسازی مدیران ارشد بیمارستان‌ها و آشنایی هر چه بیشتر آنان با مزایای فن‌آوری‌های نوین پیشنهاد می‌گردد.

تعارض منافع

این مقاله مستخرج از پایان‌نامه کارشناسی ارشد گروه مدیریت خدمات بهداشتی و درمانی، واحد الکترونیک دانشگاه آزاد اسلامی با عنوان «ارائه مدل بهینه تهیه دارو با استفاده از داده‌کاوی» است.

در انجام مطالعه حاضر، نویسندگان هیچ‌گونه تضاد منافعی نداشته‌اند. پژوهش حاضر از سوی هیچ سازمانی مورد حمایت مالی قرار نگرفته است.

References

1. World Health Organization (WHO). Handbook of supply management at first-level health care facilities; 2006 [cited 2020 Jan 2]. Available from: https://apps.who.int/iris/handle/10665/69731 https://apps.who.int/iris/bitstream/handle/10665/69731/WHO_HIV_2006.03adapt_eng.pdf

2. Nateghinia S, Hajiesmaeili MR, Shafagh SO, Goharani R, Zangi M, Ahmadi H. Prediction of the Length of Stay of Patients in the Neuro-Critical Care Unit Using Data Mining Techniques. Journal of Iranian Society Anaesthesiology and Intensive Care 2018; 40(1): 22-33. [In Persian]

3. Kelly WN. Pharmacy What It Is and How It Works. USA: CRC Press; 2017.

4. Asadi FA, Moghaddasi H, Hosseini A, Maserrat E. A survey on pharmacy information system at hospitals affiliated to Shahid Beheshti University of Medical Sciences 2009. Journal of Health Administration 2010;13(41):31-40. [In Persian]

5. Hong T, Dong M, Zhao J, Fu X, Chen Y. The application of information technology in the hospital pharmacy management based on HIS. International Symposium on Information Technologies in Medicine and Education; Hokkaido, Japan: IEEE; 2012. p. 604-67. doi: 10.1109/ITiME.2012.6291379

6. Davodi N, Raissi S. Forecasting Model for Annual Drug Demand in Iran. Journal of Health Administration 2019;22(3):89-104. [In Persian]

7. Kamali J. Prediction of drug demand using artificial neural networks (Case study: 500-hijrat distribution company antibiotic capsules) [dissertation]. Mashhad: Payam-e Noor University of Mashhad; 2017. [In Persian]

8. Salehabadi S. Categorizing customers of the drug distribution industry using data scavenger. International Conference on Management, Economics and Industrial Engineering; 2015 Jun 16; Tehran: Modiran Ideh Pardaz Paytakht Viera Institute; 2015. [In Persian]

9. Sadarina P, Kothari M, Gondaliya J. Implementing data mining techniques for marketing of pharmaceutical products. International Journal of Computer Applications & Information Technology 2013;2(1): 1-4.

10. Rezaei Inanlou R, Soleimanian Gharehchopogh F, Mojahedi M. A model for temperament (Mizaj) recognition based on fuzzy decision tree in traditional Persian Medicine. Journal of Islamic and Iranian Traditional Medicine 2018; 9(3):227-37. [In Persian]

11. Moghaddasi H, Hoseini A, Asadi F, Jahanbakhsh M. Application of Data Mining In Health. Health Information Management 2012; 9(2): 297-304. [In Persian]

12. Yeh DY, Cheng CH, Chen YW. A predictive model for cerebrovascular disease using data mining. Expert Systems with Applications 2011;38(7):8970-7.

13. Mehdipour Y, Ebrahimi S, Karimi A, Alipour J, Khammarnia M, Siasar F. Presentation a model for

[Downloaded from jhbmi.com on 2026-05-24]

prediction of cerebrovascular accident using data mining algorithm. Sadra Medical Journal 2016;4(4):255-66. [In Persian]
14. Mazaheri S, Ashouri M, Bechari Z. Predicting the treatment of heart disease using data mining algorithms. Journal of Payavard Salamat 2017; 11(3): 287-96. [In Persian]
15. Delavar SZ, Bakhshi E, Soleimani F, Biglarian A. Classification regression methods in identification of the interactive effects of the risk factors of cerebral palsy in children. Iranian Journal of Epidemiology 2014;10(2): 33-8. [In Persian]
16. Ayyoubzadeh SM, Ghazisaeedi M, Kalhori SR, Hassaniazad M, Baniasadi T, Maghooli K, et al. A study of factors related to patients' length of stay using data mining techniques in a general hospital in southern Iran. Health Information Science and Systems 2020;8(1):1. doi:10.1007/s13755-020-0099-8
17. Saadi P, Zeinalnezhad M, Movahedi Sobhani F. Modeling and Predicting the Risk of Coronary Artery Disease Using Data Mining Algorithms. Journal of Health and Biomedical Informatics 2021;8(2):193-207. [In Persian]
18. Ranjan J. Data mining in pharma sector: benefits. International Journal of Health Care Quality Assurance 2010; 22(1): 82-92. https://doi.org/10.1108/09526860910927970
19. Qingkui C, Junhu R. Study on the Demand Forecasting of Hospital Stocks Based on Data Mining and BP Neural Networks. International Conference on Electronic Commerce and Business Intelligence; 2009 Jun 6-7; China: Beijing; 2009. p. 284-9.
20. Yousefi M, ShahBahrami A, FaridiMasooleh M, Bagheri A. Modeling the prediction of readmission in cardiac hospital using data data mining. 6th International Conference on Industrial and Systems Engineering; 2020 Sep 9; Mashhad: Ferdowsi University Mashhad; 2020. [In Persian]
21. Ghazisaeedi M, Sheikhtaheri A, Behniafard N, Aghaei Meybodi F, Kargar Bideh M. A Proposed Model to Identify Factors Affecting Asthma using Data Mining. Scientific Journal of Ilam University of Medical Sciences 2019;27(1):203-12. [In Persian]
22. Asgari P, Atashi A, Meraji M, Miri M. The Comparison of Selected Data-mining techniques in ICU Mortality Risk Prediction in Imam Hossein hospital. Journal of Modern Medical Information Sciences 2019;5(2):59-67. [In Persian] doi: 10.29252/jmis.5.2.59
23. Lal Dahti J, Mohammadi M, Padidaran Moghadam F. A Method for the Diagnosis of Metabolic Syndrome based on KNN Data Mining Algorithm: A case study in Shohada-ye Kargar Hospital in Yazd, Iran. Journal of Health and Biomedical Informatics 2018;4(4):291-304. [In Persian]
24. Ayyoubzadeh SM, Ayyoubzadeh SM, Zahedi H, Ahmadi M, Kalhori SR. Predicting COVID-19 incidence through analysis of google trends data in iran: data mining and deep learning pilot study. JMIR Public Health Surveill 2020;6(2):e18828. doi:10.2196/18828
25. Zheng B, Zhang J, Yoon SW, Lam SS, Khasawneh M, Poranki S. Predictive modeling of hospital readmissions using metaheuristics and data mining. Expert Systems with Applications 2015;42(20):7110-20.
26. Dormohammadi S, Alizadeh S, Asghari M, Shami M. Proposing a prediction model for diagnosing causes of infertility by data mining algorithms. Journal of Health Administration 2014;17(57): 46-57. [In Persian]
27. Rezaee Hachesu PR, Ahmadi M, Alizadeh S, Sadoughi F. Use of data mining techniques to determine and predict length of stay of cardiac patients. Healthcare Informatics Research 2013;19(2):121-9. doi: https://doi.org/10.4258/hir.2013.19.2.121

[Downloaded from jhbmi.com on 2026-05-24]

Original Article

Journal of Health and Biomedical Informatics
Medical Informatics Research Center
2021; 8(3): 304-314

An Optimal Model for Medicine Preparation Using Data Mining

Koohestani Azita[1], Nasiripour Amir Ashkan[2]*, Riahifar Mahdi[2]

• **Received: 28 Sep 2021** • **Accepted: 8 Dec 2021**

Introduction: Lack of financial resources and liquidity are the main problems of hospitals. Pharmacies are one of the sectors that affect the turnover of hospitals and due to lack of forecast for the use and supply of medicines, at the end of the year, encounter over-inventory, large volumes of expired medicines, and sometimes shortage of medicines. Therefore, medicine prediction using available retrospective data leads to improved resource management in hospitals. Due to the high capability of data mining in modeling medical problems, selected algorithms were used to determine the optimal model of medicine preparation.

Method: In this cross-sectional study, to investigate different types of data mining algorithms, an information form was developed based on the design objectives and then defined in the form of reports in the hospital information system. The data were extracted using Crystal Report software. To develop the model, the accuracy of the data mining prediction algorithms including KNN, SVM, NN, Random Forest, LR, and Adaboost was examined based on MSE, RMSE, MAE, and R^2 criteria in Weka software.

Results: Concerning R2, MAE, and RMSE criteria, Adaboost method (0.78, 247, 827) and random forest method (0.6, 1170, 1868) had the highest accuracy compared to other models and reduced the error rate more. Other methods with the above criteria had poorer performance in predicting the research problem.

Conclusion: The results of this study indicated that the Adaboost and random forest methods are more accurate than other methods. A small percentage of hospitals plan to manage the preparation of medicines; thus, it is suggested that managers of hospitals and pharmacies use data mining in the management of their respective units.

Keywords: Medicine, Hospital, Data Mining, Prediction Algorithms

• **Citation:** Koohestani A, Nasiripour AA, Riahifar M. An Optimal Model for Medicine Preparation Using Data Mining. Journal of Health and Biomedical Informatics 2021; 8(3): 304-14. [In Persian]

1. M.Sc. in Health Services Management, Department of Health Services Management, Electronic Branch, Islamic Azad University, Tehran, Iran
2. Ph.D. in Health Services Management, Associate Professor, Department of Health Services Management, Electronic Branch, Islamic Azad University, Tehran, Iran

***Corresponding Author:** Amir Ashkan Nasiripour
Address: Department of Health Services Management, Electronic Branch, No. 5, Third Floor, Ram St., 9th Neystan, Pasdaran Ave., Tehran, Iran
• **Tel:** 09124139541 • **Email:** nasiripour@srbiau.ac.ir

[Downloaded from jhbmi.com on 2026-05-24]