مقاله پژوهشی

مجله انفورماتیک سلامت و زیست پزشکی
مرکز تحقیقات انفورماتیک پزشکی
دوره سوم، شماره اول. ۹-۱، ۱۳۹۵

کاربرد داده کاوی در پیش بینی بقای پیوند کلیه و شناسایی متغیرهای تأثیرگذار در بقای کلیه پیوندی

محترم میرزایی[۱]، سیدمحمد فیروزآبادی[۲]*

• دریافت مقاله: ۹۵/۳/۱۱ • پذیرش مقاله: ۹۵/۳/۲۷

مقدمه: نارسایی مزمن کلیه از بیماری‌های شایع در ایران و جهان است و پیوند کلیه روش درمانی مناسب و مؤثرترین راهبرد در بین بیماران مبتلا به نارسایی مزمن کلیه است. این مطالعه جهت پیش‌بینی بقای کلیه پیوندی و شناسایی عوامل مؤثر بر آن و ارائه یک مدل به منظور صحت پیش‌بینی بالاتر انجام شد.

روش: مطالعه حاضر به روش گذشته‌نگر بر روی داده‌های ۴۲۳ مورد پرونده بیماران پیوند کلیه در سال‌های ۹۰-۱۳۸۵ در مرکز آموزشی-درمانی افضلی‌پور شهر کرمان انجام گرفت. از طبقه‌بندهای شبکه عصبی، درخت تصمیم و ماشین بردار پشتیبان جهت پیش بینی بقای پیوند کلیه و از همجوشی اطلاعات برای ترکیب نتایج طبقه‌بندهای ذکر شده، به منظور طراحی مدلی با صحت بالاتر استفاده شد. همچنین، برای شناسایی متغیرهای تأثیرگذار در بقای کلیه پیوندی از الگوریتم ژنتیک و جهت تحلیل داده‌ها و اجرای الگوریتم‌ها از نرم افزارهای Clementine 12 و Weka استفاده شد.

نتایج: صحت حاصل از سه روش شبکه عصبی، درخت تصمیم و ماشین بردار پشتیبان به ترتیب برابر با ۹۴ ، ۹۲و ۹۲ درصد و صحت به دست آمده از روش همجوشی اطلاعات برابر ۹۵/۷۴ درصد بوده است. همچنین، ویژگی‌های شاخص توده بدنی و جنسیت گیرنده، سن دهنده، همسانی گروه خونی دهنده و گیرنده و سابقه پیوند کلیه توسط الگوریتم ژنتیک به عنوان متغیرهای تأثیرگذار در بقای پیوند کلیه شناسایی شد که صحت پیش‌بینی مدل پیشنهادی با این تعداد ویژگی ۹۱/۶۷ درصد بوده است.

نتیجه‌گیری: با استفاده از همجوشی اطلاعات، می‌توان صحت نتایج طبقه‌بندها را افزایش داد. همچنین، الگوریتم ژنتیک روش مناسبی جهت شناسایی ویژگی‌های بهینه است.

کلید واژه‌ها: همجوشی اطلاعات، شبکه عصبی، درخت تصمیم، ماشین بردار پشتیبان، الگوریتم ژنتیک

•ارجاع: میرزایی محترم، فیروزآبادی سیدمحمد. کاربرد داده کاوی در پیش بینی بقای پیوند کلیه و شناسایی متغیرهای تأثیرگذار در بقای کلیه پیوندی. مجله انفورماتیک سلامت و زیست پزشکی ۱۳۹۵؛ ۳(۱): ۹-۱.

۱. کارشناس ارشد انفورماتیک پزشکی، گروه انفورماتیک پزشکی، دانشکده علوم پزشکی، دانشگاه تربیت مدرس، تهران، ایران.

۲. دکترای مهندسی پزشکی، استاد، گروه انفورماتیک پزشکی، دانشکده علوم پزشکی، دانشگاه تربیت مدرس، تهران، ایران.

*** نویسنده مسؤول:** تهران، تقاطع بزرگراه جلال آل احمد و چمران، دانشگاه تربیت مدرس، دانشکده علوم پزشکی، گروه انفورماتیک پزشکی.

•شماره تماس: ۰۹۱۲۱۳۰۷۰۸۸ • **Email:** pourmir@modares.ac.ir

[Downloaded from jhbmi.ir on 2026-11-06]

مقدمه

نارسایی مزمن کلیه یکی از بیماری‌های شایع در ایران و جهان است و به طور کلی این بیماری در جوامعی که دارای شاخص بهداشتی بالایی هستند، به دلیل افزایش طول عمر شایع است. درمان نارسایی مزمن کلیه به دو صورت دیالیز و پیوند کلیه است. پیوند کلیه روش درمانی مناسب و مؤثرترین راهبرد در بین بیماران مبتلا به ESRD(End-Stage Renal Disease) است و زندگی مطلوب‌تر و کاهش خطر مرگ و میر را برای بیماران در مرحله نهایی نارسایی کلیه به ارمغان می‌آورد [۱]. در مقابل مزایای زیادی که پیوند کلیه از نظر بهبود سلامت جسمی و روانی و در نتیجه بهبود کیفیت زندگی در بیماران پیوند کلیوی دارد، ممکن است به دلیل پاسخ ایمنی میزبان نسبت به کلیه دریافتی، دچار پس زدگی شود و عواقبی مانند نیاز به پیوندی دیگر و یا حتی مرگ در پی داشته باشد. در واقع، بیماری که می‌تواند سال‌ها با دیالیز زندگی کند با پیوندی نامناسب ممکن است زندگی خود را از دست دهد یا به اجبار در مسیر عمل‌های جراحی با ریسک بالا قرار بگیرد. با توجه به موارد گفته شده، انجام مطالعه بر روی پیش‌بینی بقای کلیه پیوندی و عوامل مؤثر بر آن و ارائه یک مدل به منظور صحت پیش‌بینی بالاتر ضرورت ویژه‌ای می‌یابد [۱].

مطالعات انجام شده در حوزه پیش‌بینی بقای پیوند کلیه شامل مطالعات آماری و هوش مصنوعی است. در مطالعه‌ای در ایران در سال ۱۳۸۸، جهت بررسی بقای کلیه پیوندی در ۳۱۶ بیمار پیوند کلیه در اصفهان، از روش شبکه عصبی مصنوعی برای تحلیل استفاده شده است [۲]. صحت به دست آمده ۷۲ درصد برای پیوندهای موفق و ۶۱ درصد برای پیوندهای ناموفق بوده است. بر اساس نتایج این مطالعه، میزان بقای کلیه پیوندی در پیوندهای زنده بیشتر از پیوندهای جسدی می‌باشد. همچنین با افزایش شاخص توده بدنی بیماران در زمان پیوند، میزان بقای کلیه پیوندی کاهش می‌یابد. در مطالعه دیگری در سال ۲۰۱۲ در آمریکا، ۵۱۴۴ مورد پیوند کلیه با سن گیرنده بالای ۱۸ سال و بدون سابقه پیوند مورد بررسی قرار گرفتند [۳]. در این مطالعه از روش شبکه بیزی استفاده شد و ۴۸ متغیر برای ورودی شبکه استفاده شد و مقدار ۰/۶۰ برای مساحت سطح زیر منحنی به دست آمد. بر اساس نتایج این مطالعه، شاخص توده بدنی، نژاد، جنسیت گیرنده و سن دهنده از عوامل مؤثر در بقای پیوند شناسایی شدند. در مطالعه‌ای در سال ۲۰۱۳ در آلمان از ۵۷ متغیر مربوط به ۷۰۷ پیوند استفاده شد تا میزان فیلتراسیون گلومولار (Glomerular filtrating rate) گیرنده یک سال پس از پیوند را پیش‌بینی کند [۴]. در این مطالعه از روش ماشین بردار پشتیبان استفاده شد و متغیرهای مؤثر سن و سطح کراتینین دهنده و جنسیت و وزن گیرنده شناسایی شد. سنتوری و همکاران، در مطالعه‌ای از شبکه عصبی مصنوعی و رگرسیون لجستیک برای ایجاد مدلی برای پیش بینی نتیجه پیوند با استفاده از ۲۰ متغیر مربوط به گیرنده و دهنده پیوند در مورد پیوند کلیه اطفال استفاده کرده‌اند [۵]. بهترین مدل شبکه عصبی ایجاد شده در این تحقیق یک شبکه عصبی چند لایه پرسپترون و دارای ۸ گره در تک لایه میانی خود و با صحت نهایی ۸۸ درصد می‌باشد. در بررسی‌های انجام شده با رگرسیون لجستیک، مدل نهایی ایجاد شده دارای صحت ۸۰ درصد می‌باشد که در مقایسه با صحت مدل شبکه عصبی از صحت کمتری برخوردار بوده است.

در تمامی مطالعات در این زمینه، محققان به دنبال ۱- شناسایی مجموعه ویژگی‌های مؤثرتر در بقای پیوند و ۲- طراحی مدل‌های پیش‌بینی با صحت بالاتر و خطای پایین‌تر بوده‌اند. در زمینه پیش‌بینی بقای پیوند کلیه، عوامل متعددی دخیل هستند که به طور خلاصه می‌توان به عوامل مربوط به بیماری زمینه‌ای و بیمار، عوامل مربوط به دهنده کلیه، عوامل مربوط به عمل جراحی پیوند کلیه و عوامل مؤثر بعد از پیوند کلیه اشاره کرد، که در این مطالعه، به منظور شناسایی متغیرهای مؤثرتر در بقای پیوند کلیه، از الگوریتم ژنتیک استفاده می‌شود. یکی از راه‌های رسیدن به پیش‌بینی با صحت بالاتر، حل مسئله با رویکرد همجوشی اطلاعات (Information fusion) است. در این راهکار، طبقه‌بندها در یک ساختار موازی قرار دارند. الگوی ورودی به همه طبقه‌بندها اعمال می‌شود و هر کدام از آن‌ها یک جواب برای مسئله ارائه می‌کنند. سپس با ترکیب مناسب این جواب‌ها پاسخ نهایی به دست می‌آید. در این حالت هدف از ترکیب، بهبود صحت حل مسئله است؛ بنابراین هدف این مطالعه، استفاده از همجوشی اطلاعات چند طبقه‌بند به منظور بالا بردن صحت پیش‌بینی بقای پیوند کلیه و شناسایی متغیرهای تأثیرگذارتر در بقای کلیه پیوندی با استفاده از الگوریتم ژنتیک است.

مطالعات انجام شده در ایران غیر از یک مورد که از شبکه عصبی برای بررسی بقای پیوند استفاده شده [۲]، بقیه مطالعات همه از نوع مطالعات آماری هستند. مطالعات آماری، به تحلیل بقاء و شناسایی متغیرهای مؤثر در بقای پیوند پرداخته‌اند، اما با توجه به پیشرفتی که روش‌های داده‌کاوی در کشف الگوهای

[Downloaded from jhbmi.ir on 2026-11-06]

پنهان در مسائل پیچیده از جمله مسائل پزشکی داشته است، به نظر می‌رسد با تحلیل مسئله بقای پیوند کلیه با روش‌های داده کاوی، می‌توان با رویکردی متفاوت به شناسایی متغیرهای مؤثر در بقای پیوند کلیه پرداخت. همچنین در این مطالعه هدف این بوده از خاصیت ذاتی بهینه‌سازی الگوریتم ژنتیک در شناسایی ویژگی‌های بهینه در مسئله بقای پیوند کلیه استفاده شود. بالاترین صحت به دست آمده از روش‌های داده کاوی در این مسئله، ۸۸٪ و با روش شبکه عصبی [۲۱] بوده است که صحت بالایی نیست؛ در این مطالعه هدف این بوده که با همجوشی اطلاعات صحت پیش‌بینی بقای پیوند را به بیش از ۹۰٪ برساند. بنا بر موارد ذکر شده، سؤالات اصلی این مطالعه موارد زیر است ۱- آیا هم جوشی اطلاعات می‌تواند صحت طبقه‌بندی را به بیش از ۹۰٪ برساند؟ ۲- آیا با استفاده از الگوریتم ژنتیک و کاهش ویژگی‌ها، می‌توان به صحت طبقه‌بندی بالاتر از ۹۰٪ رسید؟ ۳- آیا الگوریتم ژنتیک با صحت بالاتری قادر به شناسایی عوامل تأثیرگذار در بقای کلیه پیوندی است؟

روش

جمع‌آوری داده‌ها: این مطالعه از نوع توصیفی-تحلیلی بوده و داده‌های مورد استفاده در این مطالعه، با توجه به نتایج مطالعات گذشته و نظر پزشک متخصص از ۴۲۳ مورد پرونده کاغذی بیماران دریافت کننده پیوند کلیه در سال های ۹۰-۱۳۸۵ در مرکز آموزشی- درمانی افضلی‌پور شهر کرمان جمع‌آوری شده است. ۱۵ ویژگی استخراج شده در جدول ۱ آورده شده است. مرگ بیمار و یا برگشت بیمار به دیالیز به عنوان نقطه پایان در تحلیل بقاء در نظر گرفته شده است. برای تحلیل داده‌ها در این مطالعه، از نرم‌افزارهای Clementine 12 و Weka 3.5 استفاده شده است. همچنین محرمانگی داده‌های بیماران هنگام استخراج داده‌ها از پرونده‌ها و مراحل بعدی حفظ شد.

جدول ۱: توصیف ویژگی‌های استفاده شده جهت پیش‌بینی بقای پیوند کلیه

ردیف	ویژگی	ردیف	ویژگی	ردیف	ویژگی
۱	سن گیرنده	۶	قد گیرنده (سانتی متر)	۱۱	وجود دیابت در گیرنده
۲	سن دهنده	۷	وزن گیرنده (کیلوگرم)	۱۲	مدت زمان دیالیز قبل از عمل (ماه)
۳	نسبت دهنده	۸	همسانی گروه خونی دهنده و گیرنده	۱۳	نوع دیالیز
۴	جنسیت گیرنده	۹	همسانی RH خون دهنده و گیرنده	۱۴	سابقه پیوند کلیه
۵	جنسیت دهنده	۱۰	نوع دهنده	۱۵	نتیجه بقای دو ساله پیوند

آماده سازی داده‌ها: پس از شناسایی داده‌ها، مرحله پیش پردازش داده‌ها آغاز شد. روش‌های مورد استفاده در پیش پردازش داده در این مطالعه عبارت بودند از: ۱- برطرف سازی مقادیر مفقوده: مقدار مفقوده با استفاده از میانگین برای داده‌های عددی و مد برای داده‌های اسمی در رکوردهای موجود پر شده است ۲- تبدیل داده: روی داده‌ها یک نرمال سازی انجام شده، این کار صحت و کارایی الگوریتم‌ها به ویژه الگوریتم‌هایی که با فاصله کار می‌کنند را بالا می‌برد. در این مطالعه ویژگی‌های با مقادیر عددی، شامل سن دهنده و گیرنده، قد و وزن گیرنده و مدت زمان دیالیز قبل از عمل بوده است. این ویژگی‌ها که مقادیر زیادی را شامل می‌شوند، نمی‌توانند ویژگی مؤثری برای فرآیند داده کاوی باشند. در صورتی که بتوان با استفاده از نتایج مطالعات گذشته یا نظر متخصص مربوطه، چند بازه عددی تعریف کرد و شماره این بازه‌ها را به مقادیر مربوطه نسبت داد، مقدار مؤثرتری برای ویژگی ایجاد می‌شود. با توجه به نظر پزشک متخصص، تبدیل ویژگی‌های سن دهنده و گیرنده به ۴ بازه سنی، قد و وزن گیرنده به ۴ بازه شاخص توده بدنی یا BMI (Body Mass Index) و مدت زمان دیالیز قبل از عمل به دو بازه کوچکتر و بزرگتر از ۶ ماه، مطابق با جداول ۲ و ۳ انجام شده است و مقدار ۰ برای بازه مناسب و نرمال و مقدار ۱ برای بازه نامناسب و غیرنرمال اختصاص داده شده است.

جدول ۲: تقسیم بندی بازه‌های سنی مناسب و نامناسب جهت پیوند کلیه

ردیف	نوع بازه	بازه سنی
۱	مناسب	۲۰ تا ۴۰ سال و ۴۰ تا ۶۰ سال
۲	نامناسب	کوچکتر از ۲۰ سال و بزرگتر از ۶۰ سال

[Downloaded from jhbmi.ir on 2026-11-06]

جدول ۳: تقسیم بندی شاخص توده بدنی بیمار

ردیف	بازه BMI	نام گروه بر اساس شاخص بهداشت جهانی	طبقه بندی جدید
۱	کوچکتر از ۱۸/۵	وزن کم	غیر نرمال
۳	بین ۱۸/۵ و ۲۵	نرمال	نرمال
۴	بین ۲۵ و ۳۰	اضافه وزن	نرمال
۲	بزرگتر از ۳۰	چاق	غیر نرمال

پس از این مرحله، مجموعه داده مورد بررسی شامل ۱۴ ویژگی با مقادیر ۰ و ۱ مربوط به اطلاعات ۱۵۶ پرونده بیمار بوده است.

مدل‌سازی: در این مرحله با استفاده از الگوریتم‌های موجود در طبقه‌بندی مدلی برای پیش‌بینی بقای دو ساله پیوند کلیه ارائه شده است. برای مدل‌سازی از دو سطح طبقه‌بندی استفاده شده، در سطح اول با سه الگوریتم درخت تصمیم، شبکه عصبی و ماشین بردار پشتیبان مدل‌سازی انجام شده، سپس نتایج این طبقه‌بندها به منظور حصول اطلاعات قابل اطمینان‌تر و با خطای کمتر، در سطح دوم با هم ترکیب شده و مدل نهایی ایجاد شده است.

شبکه عصبی: در میان مدل‌های شبکه عصبی، مدل پرسپترون چند لایه کاربردهای زیادی پیدا کرده و در این مطالعه نیز از این مدل استفاده شده است. ساختار عمومی یک شبکه پرسپترون چند لایه از لایه ورودی، خروجی و مخفی تشکیل شده است. در این مطالعه، از یک شبکه عصبی پرسپترون با ۱۳ متغیر ورودی (۲۶ جزء محاسباتی) در لایه ورودی، یک لایه پنهان با ۳ جزء محاسباتی و لایه خروجی برای بیان بقاء یا عدم بقای کلیه پیوندی استفاده شده است [۶].

درخت تصمیم: درخت تصمیم، ساختاری شبیه درخت دارد که قابلیت یادگیری دارد. الگوریتم‌های مختلفی برای درخت تصمیم وجود دارد که این الگوریتم‌ها یا به تولید یک درخت دودویی، یا به تولید یک درخت غیر دودویی می‌پردازند [۶]. در این مطالعه، از الگوریتم C4.5 برای ساخت درخت تصمیم استفاده شده است.

ماشین بردار پشتیبان: الگوریتم ماشین‌بردار پشتیبان از یک نگاشت غیرخطی برای تبدیل داده‌های آموزشی اصلی درون یک فضا با ابعاد بزرگ‌تر استفاده می‌کند؛ به عبارت دیگر می‌توان گفت با استفاده از یک نگاشت غیرخطی داده‌های آموزشی را از فضای اصلی به فضای ویژگی منتقل می‌کند. حال در این بعد جدید، الگوریتم به دنبال ابر صفحه جداکننده بهینه است که بتواند رکوردهای یک کلاس را از کلاس دیگر جدا کند [۶]. در این مطالعه، از هسته RBF برای ماشین بردار پشتیبان استفاده شده است.

همجوشی اطلاعات: در راهکار همجوشی اطلاعات، طبقه‌بندها در یک ساختار موازی قرار دارند. الگوی ورودی به همه طبقه بندها اعمال می‌شود و هر کدام از آن‌ها یک جواب برای مسئله ارائه می‌کنند، سپس با ترکیب مناسب این جواب‌ها پاسخ نهایی به دست می‌آید. در این حالت هدف از ترکیب، بهبود صحت حل مسئله است [۷].

ارزیابی مدل: ارزیابی مدل به روش Holdout و زیرنمونه تصادفی صورت می‌گیرد. در این روش مجموعه رکوردها به دو مجموعه مستقل از هم که به داده‌های آموزشی و داده‌های آزمون معروف است، افراز می‌شود. به طور معمول دو سوم رکوردها برای آموزش و بقیه به عنوان داده‌های آزمون انتخاب می‌شود. از مجموعه داده‌های آموزشی برای ساخت مدل و از مجموعه داده‌های آزمون برای به دست آوردن صحت مدل استفاده می‌شود.

شناسایی ویژگی‌های بهینه: انتخاب ویژگی به شناسایی و انتخاب ویژگی‌های متمایز برای ساخت مدل‌ها و تفسیر بهتر داده‌ها می‌پردازد. الگوریتم‌های انتخاب ویژگی مبتنی بر دو رویکرد فیلتری و مبتنی بر یادگیری هستند. در رویکرد فیلتری، الگوریتم انتخاب ویژگی مستقل از الگوریتم طبقه‌بندی است. در این مطالعه، از الگوریتم ژنتیک برای انتخاب ویژگی‌های بهینه به روش فیلتری در مرحله پیش پردازش داده استفاده خواهد شد. سپس ویژگی‌های بهینه انتخاب شده توسط الگوریتم ژنتیک به مدل پیش‌بینی تولید شده، داده شده و صحت مدل با این تعداد ویژگی سنجیده می‌شود.

نتایج

صحت پیش‌بینی بقای پیوند کلیه: صحت حاصل از سه روش شبکه عصبی، درخت تصمیم و ماشین بردار پشتیبان به ترتیب برابر با ۹۴، ۹۲ و ۹۲ درصد و صحت به دست آمده از روش همجوشی اطلاعات برابر ۹۵/۷۴ درصد است که از سه روش ذکر شده بالاتر است. همچنین، مدل پیشنهادی با ویژگی‌های بهینه شناسایی شده توسط الگوریتم ژنتیک صحت پیش‌بینی ۹۱/۶۷ درصد را نتیجه داده است. خلاصه نتایج تحلیل پیش بینی بقای پیوند کلیه در جدول ۴ آورده شده است.

[Downloaded from jhbmi.ir on 2026-11-06]

شناسایی ویژگی‌های بهینه توسط الگوریتم ژنتیک: برای شناسایی ویژگی‌های بهینه توسط الگوریتم ژنتیک از نرم‌افزار Weka استفاده شده است. سپس ویژگی‌های تشخیص داده شده توسط این الگوریتم جهت طبقه‌بندی به مدل پیش‌بینی تولید شده، داده شده است. ویژگی‌های جنسیت و شاخص توده بدنی گیرنده، سن دهنده، همسانی گروه خونی دهنده و گیرنده و سابقه پیوند کلیه توسط الگوریتم ژنتیک به عنوان ویژگی‌های تأثیرگذارتر در بقای پیوند کلیه شناسایی شده‌اند.

جدول ٤: نتایج تحلیل بقای پیوند کلیه

روش پیش بینی	صحت (درصد)	متغیرهای تأثیرگذار
شبکه عصبی	۹۴	BMI، جنسیت گیرنده، سن بیمار، سن دهنده، سابقه پیوند کلیه، نوع دیالیز
درخت تصمیم	۹۲	
ماشین بردار پشتیبان	۹۲	
همجوشی اطلاعات	۹۵/۷۴	
همجوشی اطلاعات و الگوریتم ژنتیک	۹۱/۶۷	BMI، جنسیت گیرنده، همسانی گروه خونی، سن دهنده، سابقه پیوند کلیه

شناسایی ویژگی‌های بهینه توسط قوانین درخت تصمیم: به منظور شناسایی ویژگی‌های بهینه در نتیجه بقای پیوند کلیه و ارتباط بین این ویژگی‌ها می‌توان از الگوریتم درخت تصمیم، قوانینی استخراج کرد و با تفسیر این قوانین، ویژگی‌های بیشتر تأثیرگذار را شناسایی کرد. قوانینی که از الگوریتم درخت تصمیم به دست آمده است، در جدول ۵ نمایش داده شده است.

جدول ٥: قوانین استخراج شده از درخت تصمیم

ردیف	قانون	نتیجه
۱	اگر بیمار زن باشد و BMI غیر نرمال داشته باشد و سن دهنده کمتر از ۲۰ سال و یا بزرگتر از ۶۰ سال باشد آنگاه پیوند رد می‌شود.	رد پیوند
۲	اگر بیمار مرد باشد و BMI غیر نرمال داشته باشد و سابقه پیوند کلیه داشته باشد آنگاه پیوند رد می‌شود.	رد پیوند
۳	اگر بیمار مرد باشد و BMI غیر نرمال داشته باشد و نوع دیالیز بیمار از نوع صفاقی باشد آنگاه پیوند رد می‌شود.	رد پیوند
۴	اگر سن بیمار و دهنده در بازه سنی ۲۰ تا ۶۰ سال باشد و سابقه پیوند کلیه نداشته باشد و نوع دیالیز بیمار از نوع صفاقی باشد آنگاه پیوند سالم می‌ماند.	پیوند سالم
۵	اگر بیمار زن باشد و سن دهنده بین ۲۰ تا ۶۰ سال باشد آنگاه پیوند سالم می‌ماند.	پیوند سالم

همان‌طور که در جدول ۵ مشاهده می‌شود، در تمامی قوانینی که در رابطه با رد پیوند وجود دارد، BMI بیمار در بازه غیر نرمال قرار داشته است. همچنین در مواردی که سن بیمار یا دهنده در بازه مناسب سنی یعنی ۲۰ تا ۶۰ سال بوده است، تأثیر مثبتی در بقای پیوند کلیه داشته است و برعکس سن کمتر از ۲۰ سال و بزرگتر از ۶۰ سال به همراه عوامل دیگر زمینه مساعدی را برای رد پیوند فراهم کرده است. در قوانین فوق، ویژگی‌های جنسیت دهنده و گیرنده، سابقه پیوند کلیه و نوع دیالیز تکرار شده است که این نشان دهنده اهمیت این ویژگی‌ها در بقای پیوند کلیه است.

بحث و نتیجه‌گیری

این مطالعه با رویکرد داده کاوی و با هدف پیش‌بینی بقای پیوند کلیه و شناسایی عوامل مؤثر بر آن انجام شده است. طبقه بندهای پایه استفاده شده در این مطالعه، طبقه‌بندهای شبکه عصبی، درخت تصمیم و ماشین بردار پشتیبان بوده است. سپس نتایج این طبقه‌بندها در سطح دوم همجوشی شده و صحتی بالاتر از صحت طبقه‌بندی هر یک از طبقه‌بندهای پایه به تنهایی داده است. همچنین به منظور شناسایی ویژگی‌های با اهمیت‌تر در نتیجه بقای پیوند کلیه از الگوریتم ژنتیک بهره گرفته شده است؛ که در ادامه خروجی این روش و نتایج مطالعات گذشته در این حوزه مقایسه می‌شود.

در رابطه با ویژگی‌های بهینه شناسایی شده توسط الگوریتم ژنتیک و الگوریتم درخت تصمیم، در هر دو روش شاخص توده بدنی با درجه اهمیت ۱۰۰٪ شناسایی شده است. در مقایسه با مطالعات انجام شده در این زمینه، مشاهده می‌شود که مطالعه ای که در سال ۲۰۰۷ در کشور آلمان انجام شده، تأثیر BMI در بقای پیوند کلیه را تأیید می‌کند [۸]، در دو مطالعه اشرفی و همکاران و Meier-Kriesche و همکاران نیز مشخص شد که متغیر شاخص توده بدنی با بقای پیوند ارتباط معنی‌داری

[Downloaded from jhbmi.ir on 2026-11-06]

داشته است [۲،۹]. همچنین در مطالعه‌ای در سال ۲۰۱۱ در آلمان، وزن گیرنده به عنوان یکی از ۴ عامل مؤثر در بقای کلیه پیوندی معرفی شده است [۴].

متغیرهای همسانی گروه خونی و جنسیت گیرنده نیز، جزء ویژگی‌های با اهمیت توسط الگوریتم ژنتیک شناسایی شده‌اند. در مورد گروه خونی در مطالعه [۱۰] نشان داده شده که اختلاف معنی‌داری بین گروه‌های خونی یکسان و غیریکسان با میزان بقاء وجود ندارد، اما از طرف دیگر در مطالعه Park و همکاران [۱۱] مشخص شد که گروه‌های خونی یکسان در مقایسه با گروه‌های خونی متفاوت میزان بقای پیوند بهتری نشان می دهند. در مورد جنسیت گیرنده چند مطالعه نشان داده‌اند که ارتباط معنی‌داری بین جنسیت فرد گیرنده و دهنده پیوند با میزان بقای پیوند وجود ندارد [۱۲،۱۳]، در حالی که در مطالعات Brown و همکاران و Lasserre و همکاران جنسیت گیرنده از عوامل مؤثر بر میزان بقای پیوند شناخته شده است [۳،۴].

متغیرهای سن دهنده و سابقه پیوند کلیه دو متغیری هستند که توسط الگوریتم ژنتیک و الگوریتم درخت تصمیم در این مطالعه به عنوان ویژگی‌های مؤثر شناسایی شده‌اند. در چندین مطالعه [۱۴،۳]. سن دهنده از عوامل مؤثر در میزان بقای پیوند ذکر شده است. در رابطه با سابقه پیوند کلیه، اکثر مطالعات [۲،۴،۱۵]. نشان داده‌اند که رابطه معنی‌داری بین سابقه پیوند و میزان بقای پیوند وجود ندارد، در حالی که در این مطالعه سابقه پیوند از عوامل مؤثر بر میزان بقای پیوند شناخته شده است.

متغیر نسبت دهنده و گیرنده از ویژگی‌هایی است که توسط الگوریتم ژنتیک جزء ویژگی‌های بی‌اهمیت شناسایی شده است، در حالی که در مطالعات دیگر این ویژگی رابطه معنی‌داری با میزان بقای پیوند دارد [۱۶،۱۷]، دلیل می‌تواند این باشد که تعداد داده‌های با ویژگی فوق در این مطالعه کم است و در این صورت مدل قادر به یادگیری و تعمیم نمی‌باشد. متغیرهای مدت زمان دیالیز قبل از عمل و همسانی RH به عنوان ویژگی‌های بی اهمیت شناسایی شده‌اند که مطالعات اشرفی و همکاران و Lasserre و همکاران [۲،۴] مؤید این موضوع هستند. همچنین در این مطالعات متغیرهای جنسیت دهنده، سن گیرنده، نوع دهنده و بیماری دیابت در گیرنده جزء ویژگی‌های بی اهمیت شناسایی شده‌اند، در حالی که مطالعات [۱۵،۱۸]. به رابطه معنی‌داری بین ویژگی‌های فوق و میزان بقای پیوند اشاره کرده‌اند. در چندین مطالعه [۴،۱۶] از متغیر سطح کراتینین پس از عمل به عنوان ویژگی مهم در پیش‌بینی بقای پیوند کلیه استفاده شده است، اما در این مطالعه به دلیل این که هدف پیش بینی بقای پیوند کلیه با متغیرهای مربوط به قبل از عمل بوده، از این متغیر استفاده نشده است.

در مطالعه‌ای که در سال ۲۰۱۱ و در کشور آلمان انجام شده است [۴]، از ۵۷ ویژگی مربوط به قبل و بعد از عمل مربوط به ۷۰۷ مورد پیوند کلیه با روش ماشین بردار پشتیبان استفاده شده است که در نتیجه از این ۵۷ ویژگی، ۴ ویژگی به عنوان مؤثرترین عامل در میزان بقای پیوند معرفی شده است. چهار ویژگی به دست آمده عبارت بودند از: جنسیت گیرنده، وزن گیرنده، سن دهنده و میزان کراتینین پس از عمل؛ که سه مورد اول تقریباً ویژگی‌های خروجی الگوریتم ژنتیک در این مطالعه هستند.

همان‌طور که در جدول ۴ مشاهده می‌شود، پیش‌بینی بقای پیوند کلیه به روش همجوشی اطلاعات و الگوریتم ژنتیک، صحت ۹۱/۶۷ درصد با استفاده از تنها ۵ ویژگی BMI، جنسیت گیرنده، همسانی گروه خونی، سن دهنده و سابقه پیوند کلیه نتیجه داده است که این صحت به دست آمده تقریباً معادل با صحت درخت تصمیم با ۶ ویژگی BMI، جنسیت گیرنده، سن بیمار، سن دهنده، سابقه پیوند کلیه و نوع دیالیز است. در هیچ کدام از مطالعاتی که در زمینه تحلیل بقای پیوند کلیه بررسی شده، ویژگی نوع دیالیز به عنوان متغیر تأثیرگذار در بقای پیوند کلیه شناسایی نشده است، در حالی که ویژگی‌های شناسایی شده توسط الگوریتم ژنتیک همگی جزء متغیرهای مهم در بقای پیوند کلیه در مطالعات مرتبط گزارش شده‌اند که این نشان می‌دهد الگوریتم ژنتیک با صحت بیشتری ویژگی‌های مؤثر در بقای پیوند را شناسایی کرده است، از این روی، پیش‌بینی بقای پیوند کلیه با رویکرد همجوشی اطلاعات و الگوریتم ژنتیک، نتیجه بهتری در مقایسه با پیش‌بینی به کمک درخت تصمیم به تنهایی در پی داشته است.

صحت به دست آمده از مدل پیشنهاد شده در این مطالعه بسیار مناسب و برابر ۹۵/۷۴ درصد است که در مقایسه با مطالعات مشابه بالاتر می‌باشد. همچنین صحت به دست آمده از پیش بینی بقای پیوند کلیه با تنها ۵ ویژگی خروجی الگوریتم ژنتیک برابر با ۹۱/۶۷ درصد است که صحت بسیار مناسبی است. برای مقایسه بهتر، صحت به دست آمده از روش‌های به کار رفته در چند مورد مطالعه پیش‌بینی میزان بقای پیوند کلیه با روش‌های هوش مصنوعی، در جدول ۶ آورده شده است.

[Downloaded from jhbmi.ir on 2026-11-06]

جدول ٦: نتایج پیش‌بینی مطالعات بقای پیوند کلیه

ردیف	محل و سال انجام مطالعه	روش پیش بینی	صحت (درصد)
۱	آمریکا – ۱۹۹۸ [۵]	شبکه عصبی مصنوعی	۸۰
۲	آمریکا – ۲۰۰۳ [۲۱]	درخت تصمیم	۷۶
۳	استرالیا – ۲۰۰۵ [۲۲]	شبکه عصبی-درخت تصمیم	۶۲
۴	ایتالیا – ۲۰۰۷ [۲۵]	شبکه عصبی مصنوعی	۸۷
۵	مصر – ۲۰۰۸ [۲۸]	شبکه عصبی مصنوعی	۸۸
۶	ایران – ۲۰۰۸ [۳]	شبکه عصبی مصنوعی	۸۷

همان‌طور که در جدول ۶ ملاحظه می‌شود، بالاترین صحت به دست آمده در مطالعات برابر ۸۸٪ و در مطالعه [۲۱] با روش شبکه عصبی بوده است، بنابراین استفاده از روش همجوشی اطلاعات باعث بهبود طبقه‌بندی و افزایش صحت طبقه‌بندی می‌شود. همچنین، ویژگی‌های بهینه شناسایی شده توسط الگوریتم ژنتیک، جزء متغیرهای تأثیرگذار در بقای پیوند کلیه، مطابق با نتایج مطالعات گذشته هستند. مدل پیشنهادی در این مطالعه، با همجوشی اطلاعات و تنها ۵ ویژگی شناسایی شده توسط الگوریتم ژنتیک، قادر به پیش‌بینی بقای پیوند کلیه با ۹۱/۶۷ درصد صحت است. متغیرهای مؤثر در بقای پیوند کلیه، شاخص توده بدنی (BMI)، جنسیت گیرنده، سن دهنده، همسانی گروه خونی و سابقه پیوند کلیه شناسایی شدند.

این مطالعه به پیش‌بینی بقای پیوند کلیه و شناسایی متغیرهای تأثیرگذار در بقای کلیه پیوندی پرداخته شده است. برای پیش بینی بقای پیوند کلیه از الگوریتم‌های شبکه عصبی، درخت تصمیم و ماشین بردار پشتیبان استفاده شده، سپس نتایج این الگوریتم‌ها به روش همجوشی اطلاعات با هم ترکیب شده است. نتایج نشان می‌دهد ترکیب نتایج طبقه‌بندها به روش همجوشی اطلاعات منجر به تولید مدل پیش‌بینی با صحت بالاتر از صحت هر یک از طبقه‌بندها شده است. همچنین الگوریتم ژنتیک برای شناسایی متغیرهای تأثیرگذار در بقای کلیه پیوندی بهتر از درخت تصمیم عمل کرده و با استفاده از متغیرهای کمتر، صحت پیش بینی یکسانی با الگوریتم درخت تصمیم تولید کرده است. در نتیجه، الگوریتم ژنتیک روش مناسبی جهت شناسایی ویژگی‌های بهینه است.

لازم به ذکر است، حجم نمونه این مطالعه در مقایسه با یکی از مطالعات منتشر شده در آمریکا کمتر از ۱۰٪ است [۲۲]. یکی از علل تفاوت در حجم نمونه از وجود ساز و کاری برای جمع‌آوری داده‌های چندین مرکز و تحقیق متمرکز بر روی چندین مرکز تحقیقاتی می‌باشد؛ بنابراین مطالعات امکان‌سنجی ایجاد یک مرکز برای جمع‌آوری داده‌های مراکز مختلف فعال در این حوزه می‌تواند موضوع مناسبی برای مطالعات بعدی باشد. همچنین، استفاده از متغیرهایی که در بعضی مطالعات در این زمینه استفاده شده است، از قبیل نتیجه آزمایش‌های بیمار که به دلیل این که پرونده‌های مورد بررسی این پژوهش از نوع کاغذی و مقادیر فیلدهای مربوطه ناخوانا یا سفید بود، استفاده از این متغیرها امکان‌پذیر نبود. برای مطالعات بعدی، استفاده از سایر روش‌های همجوشی اطلاعات می‌تواند به پیشرفت این مطالعه کمک کند.

References

1. Mahdavi-Mazdeh M, Heidary Rouchi A, Norouzi S, Aghighi M, Rajolani H, Ahrabi S. Renal replacement therapy in Iran. Urol J. 2007;4(2):66-4.

2. Ashrafi M, Hamidi Beheshti MT, Shahidi S, Ashrafi F. Application of artificial neural network to predict graft survival after kidney transplantation: reports of 22 years follow up of 316 patients in Isfahan. Tehran Univ Med J. 2009, 67(5): 353-9. Persian.

3. Brown TS, Elster EA, Stevens K, Graybill JC, Gillern S, Phinney S, et al. Bayesian modeling of pretransplant variables accurately predicts kidney graft survival. Am J Nephrol. 2012;36(6):561-9.

4. Lasserre J, Arnold S, Vingron M, Reinke P, Hinrichs C. P Predicting the outcome of renal transplantation. J Am Med Inform Assoc. 2012;19(2):255-62.

5. Santori G, Fontana I, Valente U. Application of an artificial neural network model to predict delayed decrease of serum creatinine in pediatric patients after kidney transplantation. Transplant Proc. 2007;39(6):1813-9.

6. Han J, Kamber M. Data Mining. Southeast Asia: Morgan Kaufmann; 2006.

7. Ruta D, Gabrys B. An overview of classifier fusion methods. Computing and Information systems 2000;7(1):1-10.

8. Rettkowski O, Wienke A, Hamza A, Osten B, Fornara P. Low body mass index in kidney transplant recipients: risk or advantage for long-term graft function? Transplant Proc. 2007;39(5):1416-20.

9. Meier-Kriesche HU, Arndorfer JA, Kaplan B. The impact of body mass index on renal transplant outcomes: a significant independent risk factor for graft failure and patient death. Transplantation. 2002;73(1):70-4.

[Downloaded from jhbmi.ir on 2026-11-06]

10. El-Husseini AA, Foda MA, Shokeir AA, Shehab El-Din AB, Sobh MA, Ghoneim MA. Determinants of graft survival in pediatric and adolescent live donor kidney transplant recipients: a single center experience. Pediatr Transplant. 2005;9(6):763-9.
11. Park K, Kim YS, Kim MS, Kim SI, Oh CK, Han DS. A 16-year experience with 1275 primary living donor kidney transplants: univariate and multivariate analysis of risk factors affecting graft survival. Transplant Proc. 1996;28(3):1578-9.
12. Courtney AE, McNamee PT, Maxwell AP. The evolution of renal transplantation in clinical practice: for better, for worse? QJM. 2008;101(12):967-78.
13. Briganti EM, Wolfe R, Russ GR, Eris JM, Walker RG, McNeil JJ. Graft loss following renal transplantation in Australia: is there a centre effect? Nephrol Dial Transplant. 2002;17(6):1099-104.
14. Almasi Hashiani A, Rajaeefard A, Hassanzade J, Salahi H. Survival analysis of renal transplantation and its relationship with age and sex. Koomesh. 2010; 11 (4):302-6. Persian.
15. Rezaei M, Kazemnejad A, Bardideh AR, Mehmoudi M. Factors affecting survival in kidney recipients at Kermanshah. Urol J. 2009;1(3):180-7.
16. Humar A, Durand B, Gillingham K, Payne WD, Sutherland DE, Matas AJ. Living unrelated donors in kidney transplants: better long-term results than with non-HLA-identical living related donors? Transplantation. 2000;69(9):1942-5.
17. Matas AJ, Payne WD, Sutherland DE, Humar A, Gruessner RW, Kandaswamy R, et al. 2,500 living donor kidney transplants: a single-center experience. Ann Surg. 2001;234(2):149-64.
18. Song YS, Jung H, Shim J, Oh C, Shin GT, Kim H. Survival analysis of Korean end-stage renal disease patients according to renal replacement therapy in a single center. J Korean Med Sci. 2007;22(1):81-8.
19. Goldfarb-Rumyantzev AS, Scandling JD, Pappas L, Smout RJ, Horn S. Prediction of 3-yr cadaveric graft survival based on pre-transplant variables in a large national dataset. Clin Transplant. 2003;17(6):485-97.
20. Shadabi F, Cox RJ, Sharma D, Petrovsky N. A hybrid decision tree - artificial neural networks ensemble approach for kidney transplantation outcomes prediction. 9th International Conference on Knowledge-Based Intelligent Information and Engineering Systems (KES 2005); 2005 Sep 14-16; Melbourne, Australia: Springer Berlin Heidelberg; p.116-22.
21. Akl A, Ismail AM, Ghoneim M. Prediction of graft survival of living-donor kidney transplantation: nomograms or artificial neural networks? Transplantation. 2008;86(10):1401-6.
22. Port FK, Merion RM, Roys EC, Wolfe RA. Trends in organ donation and transplantation in the United States, 1997-2006. Am J Transplant. 2008;8(4 Pt 2): 911-21.

[Downloaded from jhbmi.ir on 2026-11-06]

Original Article

Journal of Health and Biomedical Informatics
Medical Informatics Research Center
2016;3(1): 1-9

The Impact of Data Mining on Prediction of Renal Transplantation Survival and Identifying the Effective Factors on the Transplanted Kidney

Mohtaram Mirzaei[1], Mohammad Firooz Abadi[2]*

• **Received:** 31 May, 2016 • **Accepted:** 16 Jun, 2016

Introduction: Chronic kidney failure is a common disease in the world and kidney transplantation is the most effective treatment in patients with chronic kidney failure. The aim of this study was to predict the survival of transplanted kidney and identify its effective factors, and also to provide a model for higher prediction accuracy.

Method: In this retrospective study, data from 423 cases of kidney transplant patients during 2006-2011 in Afzalipour Teaching Hospital in Kerman were obtained. The neural networks, decision tree and support vector machine were used to predict kidney transplantation survival and information fusion was used to combine the results of these classifiers and design a model with higher prediction accuracy. In addition, for identifying factors affecting the survival of transplanted kidney, genetic algorithm was used and for data analysis and implementation of algorithms, Clementine 12 and Weka 2.3 were used.

Results: The accuracy of neural networks, decision tree, and support vector machine were 94%, 92%, and 92%, respectively, and the accuracy of information fusion was 95.74%. Also, recipient BMI and gender, donor age, compatibility of donor and recipient blood group, and history of kidney transplantation as the effective factors on renal transplantation survival were identified by genetic algorithm. The prediction accuracy of this model was 91.67%.

Conclusion: The results show that information fusion can increase the prediction accuracy. Also, the genetic algorithm as an effective method can be used for identifying the optimal features.

Key words: Information fusion, Neural Network, Decision Tree, Support Vector Machine, Genetic Algorithm

•**Citation:** Mohtaram M, Firooz Abadi SM. The Impact of Data Mining on Prediction of Renal Transplantation Survival and Identifying the Effective Factors on the Transplanted Kidney. Journal of Health and Biomedical Informatics 2016; 3(1):1-9.

1. M.Sc in Medical Informatics, Medical Informatics Dept., Faculty of Medicine, Tarbiat Modares University, Tehran, Iran.
2. Ph.D. in Medical Engineering, Professor, Medical Informatics Dept., Faculty of Medicine, Tarbiat Modares University, Tehran, Iran.
***Correspondence:** Medical Informatics Dept., Faculty of Medicine, Tarbiat Modares University, Jalal Ahmed, Nasr, Tehran.
• **Tel:** 09121307088 • **Email:** pourmir@modares.ac.ir

[Downloaded from jhbmi.ir on 2026-11-06]