

مدل‌سازی سیستم ایمنی بدن با استفاده از شبکه‌های بیزی

فرزاد اسکندری

دریافت: ۱۳۹۵/۱۱/۱۷ / ویرایش: ۱۳۹۶/۴/۲۷ / پذیرش: ۱۳۹۶/۶/۲۲ / انتشار: ۱۳۹۷/۶/۲۹

گروه آمار، دانشکده علوم ریاضی و رایانه، دانشگاه علامه طباطبائی، تهران، ایران

*مسئول مکاتبات: askandari@atu.ac.ir

چکیده. در این مقاله براساس الگوی شبکه‌های بیزی در ساختار مارکوفی و درختی، که یکی از الگوریتم‌های معروف در یادگیری ماشین است، به مطالعه عفونت ادراری افراد، که یکی از علائم شایع در ضعف سیستم ایمنی بدن است، پرداخته شده است. در این مقاله یک نمونه در مقیاس بزرگ برای ارزیابی عملکرد الگوریتم شبکه‌های بیزی انجام شده است. در این مطالعه ۴۰۵۲ نمونه از بانک اطلاعاتی مربوط به معاونت درمان دانشگاه علوم پزشکی قزوین - مرکز بهداشت و درمان شهرستان تاکستان استخراج شد که هم آزمایش تحلیل ادرار و هم آزمایش کشت ادرار داشتند. با توجه به اهداف مطالعه و با استفاده از نظر متخصص آزمایشگاه و متخصص ارولوژی ۱۵ متغیر انتخاب شد. متغیر نوع عفونت ادراری به عنوان پاسخ و ۱۴ متغیر ورودی در مدل شبکه بیزی با الگوریتم‌های مختلف تحت بررسی قرار گرفتند. نتایج نشان می‌دهد روش شبکه بیزی با ساختار درختی و انتخاب گام به گام متغیرها تقریباً بدون خطا برای ۹۹/۷ درصد از داده‌های آموزشی، (۷۵ درصد از کل داده‌ها)، و ۹۹/۸ درصد داده‌های آزمایشی (۲۵ درصد از کل داده‌ها) دارای تشخیص درست است. براساس الگوی شبکه‌های بیزی، متغیرهای کمکی تأثیرگذار بر عفونت ادراری، افزایش باکتری‌ها و کاهش گلبول‌های سفید خون در گروه‌های سنی مختلف به دست آمده است. نتایج این مطالعه در چارچوب یادگیری ماشین و سیستم‌های هوشمند، می‌تواند برای تشخیص سریع این بیماری و درمان افراد مشکوک به این بیماری در سطح جامعه بدون حضور فیزیکی استفاده شود.

واژه‌های کلیدی. الگوریتم درختی، الگوریتم مارکوفی، باکتری، گلبول‌های سفید، یادگیری ماشین

The modeling of body's immune system using Bayesian Networks

Farzad Esakandari

Received 05.02.2017/ Revised 18.07.2017/ Accepted 13.09.2017/ Published 20.09.2018

Department of Statistic, Mathematics Faculty, Allameh Tabataba'i University, Tehran, Iran

*Correspondent author: askandari@atu.ac.ir

Abstract. In this paper, the urinary infection, that is a common symptom of the decline of the immune system, is discussed based on the well-known algorithms in machine learning, such as Bayesian networks in both Markov and tree structures. A large scale sampling has been executed to evaluate the performance of Bayesian network algorithm. A number of 4052 samples were obtained from the database of the Takestan Department of Health, a center affiliated to Qazvin University of Medical Sciences. According to the goals of the study and using the expert opinion of the laboratory and urologist, 15 variables were selected. The database included both urine analysis and culture tests. The results indicated 99.7% accuracy of the diagnosis for the training data, (75% of total data), and 99.8% accuracy of the diagnosis for testing data (25% of total data). Based on the Bayesian network model, the important covariates influencing the Urinary infection have been proved to be the increase of bacteria and the decrease of white blood cells in different age groups. The results of this study can be used in the context of machine learning and intelligent systems for rapid diagnosis of the disease and the treatment of people suspected of suffering from it.

Keywords. bacteria, machine learning, markov algorithm, tree algorithm, white blood cells

مقدمه

یکی از مکانیسم های مهمی که بدن انسان را در مقابل ویروس ها، باکتری ها و هر نوع عامل بیماری زای دیگری محافظت می کند سیستم ایمنی بدن است. معمولاً عواملی مانند تغذیه نامناسب، استرس یا عفونت ها، در عمل کردن این سیستم اختلال ایجاد می کند. بدون شک، تضعیف و اختلال در عمل کردن سیستم دفاعی بدن با ارسال علائمی همراه است. خستگی، عفونت های مداوم، سرماخوردگی و گلودرد، آلرژی ها و نیز زخم هایی که بهبودی شان مدت ها طول می کشد علائمی هستند که می توانند از ضعف سیستم ایمنی بدن خبر دهند. در این میان، عفونت ها از آن دسته از بیماری ها هستند که در ضعف سیستم ایمنی بدن نقش کلیدی دارند. باید توجه کرد یکی از عوامل مهم در تضعیف سیستم ایمنی بدن کاهش میزان گلبول های سفید خون است؛ زیرا در آن صورت بیشتر در معرض عوامل خطر ناک خارجی قرار گرفته و بیشتر بیمار می شویم.

به لحاظ علمی، ایمنی در بدن به دو صورت طبیعی یا مصنوعی ممکن است به دست آمده باشد. در هر دو مورد، ایمنی امکان دارد فعال یا غیر فعال باشد. ایمنی به دست آمده بصورت طبیعی و فعال، با توجه به استفاده سلول های حافظه ای، از ایمنی طبیعی به دست آمده فعال تر یا نسبتاً دائمی است. این ایمنی زمانی که بدن آلوده می شود یا سلول های قرمز خون عامل خارجی را دریافت می کنند به طور خود کار فعال می شود، توسعه می یابد و به تولید آنتی بادی برای غیر فعال کردن آنتی ژن می پردازد.

ایمنی به دست آمده طبیعی و منفعل، زمانی که آنتی ژن ها فعال نیستند آنتی بادی های جدیدی ایجاد نمی شود، ولی سیستم ایمنی بدن خود دارای حافظه دفاع در برابر آنتی ژن ها است ایمنی مصنوعی فعال به دست آمده، هنگامی که از موجودات مرده یا موجوداتی که ضعیف هستند به منرله عامل ویروسی استفاده شود، توسعه می یابد و بدن برای درمان با این ویروس های ضعیف به مقابله می پردازد؛ بدین مفهوم که از آنتی ژن های موجودات خاص در درمان بدون تحریک و با نگهداری واکنش بیماری در سلول های حافظه استفاده می شود (Hunt & Cooke, 1996; Kim & Bentley, 1999; Castro & Von Zuben, 2000).

زمانی از ایمنی مصنوعی منفعل به دست آمده استفاده می شود که یک آنتی بادی خاص، که توسط انسان یا حیوانات دیگر تولید شده

است، برای درمان حالت های اورژانس به بدن تزریق شود. این ایمنی از آنجاکه در سیستم ایمنی بدن فعال نیست، کوتاه مدت است. سیستم ایمنی بدن با نظارتی دائمی بر عمل کردن سلولی همواره رخ دادهای به وجود آمده را جهت کسب دانش پردازش می کند و با استفاده از این دانش به واکنش سریع در هنگام رخ دادن اتفاقات مشابه می پردازد (Kim & Bentley, 2002; Castro & Timmis, 2002). در الگوریتم سیستم های ایمنی، الگوهایی که باعث این رخداد شده اند به همراه پاسخ و نوع عملی که بدن باید هنگام مشاهده این الگو از خود نشان دهد در حافظه سلولی قرار می گیرد تا برای رخداد های مشابه از آن پیروی کند (Castro & Von Zuben, 2000). بررسی ها نشان داده است چهار مدل از سیستم ایمنی طبیعی بدن تحت ارزیابی قرار گرفته است. در مدل اول نمایش کلاسیک باعث می شود سیستم ایمنی بدن سلول های خودی و غیر خودی را تشخیص دهد (Aickelin & Cayzer, 2002; Twycross, 2007; Jerne, 1974).

درواقع، تشخیص سلول های خودی و غیر خودی با استفاده از لئوسیت های تولید شده در اندام های لئوای انجام می پذیرد. این لئوسیت ها قابلیت یادگیری اتصال به آنتی ژن ها دارند. در مدل دوم، که به نظریه انتخاب کلونال (CST) معروف است، در آن سلول های B فعال از طریق فرایند شبیه سازی آنتی بادی تولید می کنند. در این حالت، کلون ها به صورت جهش یافته آنها را تولید می کنند. مدل نظریه شبکه (NT) مدل سوم است که در آن فرض بر این است سلول های B یک شبکه را تشکیل می دهند. هنگامی که سلول B به آنتی ژنی پاسخ می دهد، سلول B فعال می شود و تمام سلول های B را، که به هم متصل هستند و یک شبکه را تشکیل می دهند، تحریک می کند. به عنوان آخرین الگو در بررسی سیستم ایمنی طبیعی بدن می توان به مدل نظریه خطر (RT) اشاره کرد. در این مدل سیستم ایمنی بدن توانایی تشخیص آنتی ژن های خطرناک و غیر خطرناک را نیز دارد. اما سیستم ایمنی مصنوعی بر اساس شبیه سازی برخی از جنبه های سیستم ایمنی طبیعی است.

یادگیری در سیستم ایمنی بدن جهت افزایش اندازه جمعیت لئوسیت ها، که اغلب آنتی ژن را تشخیص می دهد، به کار می رود (Graaff & Engelbrecht, 2006). ارزش سیستم ایمنی بدن به واسطه فرایندی شناخته شده با عنوان بلوغ میل اتصال انجام

شبکه از سلول‌ها هستند و هنگامی که یک سلول B در شبکه به یک سلول خارجی پاسخی داد، سلول B فعال شده کلیه سلول‌های B را که به صورت شبکه‌ای به هم متصل هستند تحریک می‌کند. بنابراین، یک لنفوسیت نه تنها با آنتی‌ژن تحریک می‌شود، بلکه می‌تواند لنفوسیت‌های همسایه را تحریک کند یا توسط آنها فعال‌تیش متوقف شود، یعنی زمانی یک لنفوسیت مجدد فعال می‌شود که به تحریک آنتی‌ژن واکنش نشان می‌دهد و با ترشح آنتی‌بادی و تولید کلون جهش‌یافته به تحریک لنفوسیت‌های مجاور نیز می‌پردازد. این به معنی آن است که لنفوسیت همسایه می‌تواند پس از فعال‌شدن خود دوباره به تحریک لنفوسیت آنتی‌ژن-تحریک‌شده با تولید کلون جهش‌یافته، به تحریک گروه‌های بعدی همسایگان، و ... واکنش نشان دهد. یکی از مدل‌های شبکه‌ای مهم برای تحلیل اینگونه داده‌ها شبکه‌های بیزی است. شبکه‌های بیزی یکی از مطالب کاربردی است که در بانک‌های داده‌ای که دارای بعد زیاد متغیرها و همچنین دارای حجم زیاد داده‌ها بسیار کارا و مفید است. اگر پژوهشگران علوم داده‌ای از شبکه‌های بیزی به درستی استفاده کنند، تحلیل داده‌ها بسیار دقیق صورت خواهد گرفت.

از سال ۱۹۵۰، رایانه‌ها در تحلیل و ذخیره‌سازی داده‌ها به کار گرفته شدند. پس از حدود ۲۰ سال حجم داده‌ها دوبرابر شد و بعد از آن تقریباً هر دو سال یک‌بار هم‌زمان با پیشرفت فناوری اطلاعات، حجم داده‌ها هم به دوبرابر افزایش یافت. این پیشرفت آن‌قدر زیاد بود که تعداد رکوردهای برخی از پایگاه‌های داده به چند صد میلیارد رسید. وقتی که حجم داده‌ها زیاد باشد، کاربران هر قدر هم زبردست و باتجربه باشند، قادر به تشخیص الگوهای مفید در میان حجم انبوه داده‌ها نخواهند بود، یا اگر قادر به انجام این کار هم باشند، هزینه عملیات از نظر نیروی انسانی و اقتصادی بسیار زیاد است. بنابراین، استفاده از ابزارهای داده‌کاوی برای کشف دانش از داده اجتناب‌ناپذیر به نظر می‌رسد. یکی از ابزارهای آماری در حوزه داده‌کاوی شبکه‌بیزی است که می‌تواند به ارائه دانش از مجموعه داده‌های ناکامل یا داده‌های با بعد بالا بپردازد. در واقع، در این حالت ترکیب دانش و داده تسهیل می‌شود.

مدل شبکه‌های بیزی ابزاری قدرتمند در حوزه ارائه دانش است (Nasraoui et al., 2002; Pearl, 1990; Sarafijanovic &

می‌شود. بلوغ میل اتصال را می‌توان به دو فرایند کوچک‌تر یعنی فرایند تکثیر شبیه‌سازی کلونال و فرایند طبیعی بیش‌جهش تقسیم کرد. تعامل لنفوسیت‌ها با آنتی‌ژن‌ها به فعال‌شدن لنفوسیتی منجر می‌شود که سلول در آن کلون گسترش می‌یابد و رشد می‌کند. هنگامی که آنتی‌ژن باعث تحریک لنفوسیت می‌شود، لنفوسیت نه تنها برای اتصال به آنتی‌ژن آنتی‌بادی ترشح می‌کند بلکه خود نیز به صورت جهش‌یافته‌ای تولید می‌شود و در تلاش برای تمایل اتصال بالاتر با آنتی‌ژن شناسایی شده است. در فرایند اخیر جسمی بیش از حد جهش‌یافته شناخته می‌شود. بنابراین، هنگام قرار گرفتن در معرض آنتی‌ژن‌های تکراری، سیستم ایمنی بدن اغلب متناسب با تغییرات آنتی‌ژن، آن را درحفاظه یادگیری خود قرار می‌دهد و به جای ایجاد گیرنده تصادفی به تکثیر گیرنده‌ای با مجموعه‌ای از خصوصیات می‌پردازد که آنتی‌ژن‌ها را به صورت دقیق‌تر نشان‌دهند. در صورتی که آنتی‌ژن‌ها معین شوند یا لنفوسیتی که به کلون‌شدن نیاز دارد انتخاب شود، این فرایند را انتخاب کلونال می‌نامند. بهترین کلون‌ها آنهایی هستند که تولید آنتی‌بادی آنها به بهترین آنتی‌ژن (با بالاترین) میل اتصال دارند. از آنجاکه تعداد کل لنفوسیت‌ها در سیستم ایمنی بدن معین است، افزایش در اندازه و تعداد برخی از کلون‌ها باعث کاهش اندازه دیگر کلون‌ها می‌شود. این موضوع باعث می‌شود سیستم ایمنی بدن تعدادی از آنتی‌ژن‌هایی را که قبلاً آموخته است فراموش و حذف کند. هنگامی که یک آنتی‌ژن مشخص است و شناسایی شده است، سیستم ایمنی بدن پاسخی به اندازه بهترین مقدار شبیه‌سازی ارائه می‌دهد. این پاسخ به عنوان پاسخ ایمنی ثانویه نامیده می‌شود.

اکنون، اگر بخواهیم یک مدل براساس یادگیری ماشین در تحلیل سیستم ایمنی بدن ارائه کنیم، انتظار این است که مدل تحت بررسی با تعیین درست ارتباط بین متغیرهای درگیر و تعیین سهم تأثیر متغیرها بر یکدیگر توانایی تشخیص و ارائه الگویی مناسب در دو حوزه داده‌های آموزشی و آزمایشی را با کمترین خطا داشته باشد. یکی از مدل‌های قابل اعتماد در این مورد، استفاده از مدل شبکه‌های بیزی است که در ادامه به آن پرداخته می‌شود.

مواد و روش‌ها

نظریه شبکه برای اولین بار توسط Nasraoui و همکاران (2002) ارائه شد. طبق این نظریه سلول‌های B به هم پیوسته به نمودار یک

نمی‌توان با دنبال کردن یال‌های جهت‌دار به یک گره بازگشت. چنین شبکه‌هایی در واقع یک گراف بدون دور جهت دارند. پارامترهای شبکه‌های بیزی، ساختار را از مدل کیفی به مدل کمی تبدیل می‌کند. طراحی ساختار شبکه‌های بیزی نیز براساس روابط بین متغیرها صورت می‌پذیرد. به عبارت دیگر، شبکه بیزی یک گراف بدون دور جهت‌دار است.

تعریف ۱: بردار $G = (V, E)$ را که در آن V مجموعه رئوس و E مجموعه یال‌ها در زیر مجموعه‌ای از $V \times V$ است را یک گراف می‌نامند.

هر گره گراف مربوط به یک متغیر تصادفی در دامنه است و این گراف یک خانواده از توزیع‌های احتمال را روی متغیرهای V نشان می‌دهد. یال‌های جهت‌دار E مبین وابستگی بین رأس‌های مربوط به آن یال است. اولین مرحله در ساخت یک شبکه بیزی، طراحی بخش گرافیکی آن است. پس از ایجاد ساختار شبکه، نوبت به برآورد پارامترهای شبکه بیزی می‌رسد که با مشخص شدن ساختار شبکه نسبتاً ساده است و از احتمال‌های شرطی براساس داده‌ها به دست می‌آید. شبکه‌های بیزی دارای مزایایی هستند که از مهم‌ترین آنها عبارتند از:

- روابط علی بین متغیرها را به راحتی می‌توان از شبکه‌های بیزی به دست آورد.
- با ترکیب شبکه‌های بیزی و برخی از روش‌های دیگر، راهکارهایی اصولی و کارا برای جلوگیری از بیش‌برازش داده مطرح شده است.
- شبکه‌های بیزی در ترکیب با روش‌های آمار بیزی، قادرند دانش درخور توجهی در حوزه مسئله تولید کنند و ارائه دهند. این موضوع به ویژه در زمانی که داده گران‌بها یا کمیاب است، بسیار حائز اهمیت است.
- شبکه‌های بیزی به راحتی می‌توانند مجموعه داده‌های ناقص را پردازش کنند.

شبکه‌های بیزی از اصول نظریه گراف، نظریه احتمالات، علوم رایانه و آمار استفاده می‌کنند (Heckerman, 1996; Engelbrecht, 2007). این ساختارها به علت نمایش گرافیکی داده‌ها از درک شهودی درخور توجهی برخوردارند. ایده اصلی در طراحی مدل گرافیکی استفاده از ساختار علت و معلولی است که از نظریه گراف برای ایجاد رابط مناسب استفاده می‌شود. این

(Le, 2004). در این خصوص پیشرفت‌هایی در حوزه بهبود کارایی روش‌های استنتاج در شبکه‌های احتمالاتی حاصل شد. موضوع یادگیری در شبکه‌های بیزی در هر دو حوزه ساختاری و پارامتری، به صورت هوشمند و از روی مجموعه داده‌ها ارائه شده است (Dasgupta et al., 2011). در این باره، شبکه بیزی یک نمایش فشرده از توزیع احتمال بسیار بزرگ است و زمینه‌های وسیعی از مسائل و روش‌های تحلیل داده و استدلال احتمالاتی را پوشش می‌دهد که در آن داده‌ها با تعداد زیادی متغیر جمع‌آوری شده‌اند.

در سال‌های اخیر، شبکه‌های بیزی و طرح‌های مرتبط با تشکیل چهارچوب‌های احتمالی، بیشترین کاربرد را در زمینه استخراج دانش و رمزگشایی از دانش غیرقطعی و در سیستم‌های خبره در زمینه هوش مصنوعی داشته‌اند (Shuohao et al., 2014). از دیگر کاربردهای این ابزار می‌توان به حوزه‌های کلی خوشه‌بندی، دسته‌بندی، پیش‌بینی، تخمین و مانند آن اشاره کرد. این قدرت انعطاف در حل مسائل و کاربرد فراوان در زمینه‌های مختلف باعث شده است که از زمان ابداع این شبکه‌ها، محققان متعدد برای توسعه کاربردها، رفع نواقص و افزایش کارایی و نیز تعمیم آنها تلاش کنند.

با هدف افزایش درک مطالب سه مفهوم شبکه‌های بیزی، نظریه گراف و نظریه احتمال با هم ترکیب شدند. قابلیت‌های شبکه بیزی از جمله استنتاج، یادگیری، و نمایش دانش، آنها را به ابزار مناسبی برای ساخت سیستم‌های هوشمند در حوزه مسائل مختلف تبدیل کرده است. از آنجا که یک شبکه بیزی، معرف یک مدل کامل برای متغیرها و روابط بین آنهاست، می‌تواند برای نمایش احتمالاتی بین متغیرها مورد استفاده قرار گیرد.

مدل‌های گرافیکی زیرساختی را برای معرفی مدل‌هایی ایجاد می‌کنند که در آنها تعدادی متغیر با هم تعامل می‌کنند. گره‌ها در شبکه بیزی نشان‌دهنده متغیرهای تصادفی گسسته یا پیوسته (یا ترکیب هر دو) در حوزه موردنظر هستند. یال‌های جهت‌دار بین زوج متغیرها بیانگر وابستگی‌های مستقیم آنهاست. میزان رابطه بین متغیرها با تخصیص توزیع‌های احتمال شرطی به هر گره مشخص می‌شود. همچنین، مقادیر هر متغیر نیز با استفاده از جدول‌های توزیع احتمال مشخص می‌شود. تنها محدودیت روی یال‌ها این است که نباید هیچ دور جهت‌داری ایجاد شود. به عبارت دیگر،

تعریف ۳- نسل یک رأس: مجموعه همه رأس‌هایی است که از رأس مورد نظر یک مسیر جهت‌دار به آنها وجود داشته باشد. به عبارت دیگر، نسل یک رأس در آخر زنجیره دیده می‌شود. مجموعه پدران گره X ، با عبارت $Parents(X)$ مشخص می‌شود. هر گره بدون والد ریشه نامیده می‌شود. با توجه به مفهوم علی در ساختار شبکه بیزی، گره ریشه علت اصلی است درحالی که برگ نشان‌دهنده اثر نهایی است. طبق این قرارداد، شبکه بیزی درختی است که معمولاً وارونه قرار گرفته است؛ یعنی ریشه بالا و برگ‌ها پایین هستند.

مدل‌سازی با شبکه‌های بیزی

به‌طور کلی، مدل‌سازی با شبکه‌های بیزی به فرضیه‌هایی از زنجیره مارکوف نیاز دارد. بنابر خاصیت مارکوف در توزیع احتمال شرطی، هر رأس فقط به والدین خود وابسته است. به عبارت دیگر تمام وابستگی‌های مستقیم توسط یال‌ها مشخص شده است. در واقع، هر استقلال با فقدان یک یال در سیستم مشخص می‌شود. در شبکه‌های بیزی، که واجد ویژگی مارکوف هستند، هر رأس به شرط والدینش از مجموعه رأس‌های غیرفرزند خود مستقل خواهد بود. این موضوع معادل برقراری ویژگی مارکوف در شبکه‌های بیزی است. مفهوم مفید دیگر در شبکه‌های بیزی، پوشش مارکوف یک گره است که شامل والدین گره، فرزندان آن و والدین فرزندان (همسران) است. به‌طور خاص، وقتی مقدار برخی از متغیرها مشاهده می‌شود، شرطی کردن بر مبنای اطلاعات جدید مورد نظر است. فرایند شرطی کردن (که همچنین انتشار احتمال، استنتاج یا به‌روزرسانی باور نامیده می‌شود) با استفاده از جریان اطلاعات از طریق شبکه انجام می‌شود. چهار نوع استدلال در حوزه شبکه‌های بیزی ارائه می‌دهد که به استدلال تشخیصی، استدلال پیش‌بینانه، استدلال توصیفی و استنتاج ترکیبی معروف است (Twycross, 2007). انتخاب نوع استنتاج در یک شبکه بیزی وابسته به موقعیت گره‌های درخواست‌شده و گره‌های مشاهده‌شده است. از آنجایی که شبکه‌های بیزی شامل دور جهت‌دار نیستند، هیچ رأسی جزء مجموعه فرزندان یا مجموعه والدین خود نخواهد بود. در نتیجه، با استفاده از تعریف ذیل، استدلال احتمالاتی از روی شبکه‌های بیزی به راحتی امکان‌پذیر می‌شود.

ساختارها به علت ظاهر گرافی که دارند نمایش مناسبی برای داده‌ها هستند که به صورت شهودی قابل درک است. همان‌طور که قبلاً گفته شد، مدل گرافیکی به صورت $G=(V,E)$ تعریف می‌شود که گره‌های V برای نمایش متغیرهای تصادفی به کار می‌روند که از تابع توزیع احتمال $P(V)$ استفاده می‌کنند. همچنین، یال‌های E برای تعیین توزیع احتمال وابستگی بین گره‌ها به کار می‌رود، به این صورت که وجود یک یال، بیان‌کننده وابستگی بین گره‌هاست و نبود یال، بین گراف‌ها بیان‌کننده عدم وابستگی بین گره‌هاست. نکته مهم در مدل‌های گرافیکی این است که توزیع پیوسته احتمال برای گره‌های مختلف چگونه انجام می‌گیرد و این توزیع پیوسته چگونه حساب می‌شود. پاسخ این است که برای هر گره با استفاده از قانون احتمال شرطی، مقدار احتمال هر گره به راحتی محاسبه می‌شود. سپس با استفاده از حاصل ضرب احتمالات داخلی، احتمالات سراسری تولید می‌شود. در واقع توزیع توأم احتمال کل به صورت زیر محاسبه می‌شود:

$$P(A, B, C, D, E, F) = P(A)P(B)P(C|A, B)P(D|C)P(E|C)P(F|D, E)$$

ساختار یا توپولوژی شبکه باید روابط میان متغیرها را ارائه کند. به این صورت که یک گره والد و یک گره فرزند است، اگر یک یال از گره اول به گره دوم وجود داشته باشد. برای مثال، یک یال از رأس X_i به رأس X_j مبین این مطلب است که رأس X_j با احتمالی تحت تأثیر رأس X_i است. به عبارت دیگر، مقداری که رأس X_j می‌گیرد، تحت تأثیر رأس X_i خواهد بود. ساختار یک شبکه بیزی باید رابطه کیفی بین متغیرها را به درستی نشان دهد. همان‌طور که قبلاً ذکر شد، اولین مرحله در ساخت شبکه، طراحی بخش گرافیکی آن است. در شبکه بیزی، اگر یکی از گره‌ها به صورت مستقیم از گره دیگر تأثیر بپذیرد یا روی آن تأثیر بگذارد، باید جهت یال، نحوه تأثیرگذاری، یا تأثیرپذیری را مشخص کرد. در شبکه‌های بیزی، جهت یال‌ها از گره والد (علت)، به سمت گره فرزند (معلول) است. با فرض وجود زنجیره جهت‌داری از گره‌ها و گسترش این تشبیهات، دو مجموعه زیر تعریف‌پذیر خواهد بود.

تعریف ۲- اجداد یک رأس: مجموعه همه رأس‌هایی است که از آنها یک مسیر جهت‌دار به رأس مورد نظر وجود داشته باشد. به عبارت دیگر، مجموعه اجداد یک رأس زودتر از آن رأس در زنجیره دیده می‌شوند.

است. یادگیری ساختاری شبکه های بیزی نسبت به یادگیری پارامتری آن دشوارتر است؛ زیرا با مشخص بودن ساختار شبکه، تعیین پارامترهای شبکه یا به نوعی کمی سازی آن کار دشواری نیست. الگوریتم های زیادی برای یادگیری ساختار شبکه بیزی وجود دارد که در سه دسته کلی روش های مبتنی بر قید، روش های جست و جو- امتیازدهی و روش های ترکیبی خلاصه می شوند.

وظیفه اصلی یادگیری شبکه بیزی از داده ها این است که به طور خودکار یال های میان گره ها را پیدا کند. در روش های مبتنی بر قید، از آنجا که وجود یال بین دو متغیر در شبکه بیزی وابستگی شرطی آن دو را توصیف می کند، می توان از نتایج آزمون وابستگی برای ایجاد ساختار شبکه بیزی سازگار با داده ها استفاده کرد. یادگیری ساختار با استفاده از روش های مبتنی بر قید شامل یک فرایند دو مرحله ای است: در مرحله اول کالبد که بدون جهت است یک گراف ایجاد می شود. سپس یال های ساختار ایجاد شده جهت دار می شوند. در مرحله اول، برای تعیین استقلال یا وابستگی بین متغیرها از آزمون های استقلال استفاده می کنند. اما در روش های جست و جو- امتیازدهی از یک امتیاز برای ارزیابی شبکه های کاندیدا استفاده می شود. بنابراین، مسئله یادگیری به مسئله جست و جو تبدیل می شود. در این نگرش نسبت به روش های مبتنی بر قید به جواب های دقیق تری خواهیم رسید. در واقع، خروجی این روش ساختاری منحصر به فرد خواهد بود.

تعریف ۴- تابع امتیاز g را، که برای ارزیابی ساختار شبکه در نظر گرفته می شود، تجزیه پذیر گویند، هر گاه بتواند مقداری را که برای هر ساختار در نظر می گیرد به صورت مجموع مقادیر موضعی که به هر رأس و والدین آن مربوط است بیان کند.

در فضای لگاریتمی آن را به صورت ذیل بیان می کنیم:

(۳)

$$g(G; D) = \sum_{x_i} g(x_i, Pa(x_i); D)$$

روش جست و جویی را در نظر بگیرید که در هر مرحله فقط یک یال را تغییر می دهد، اگر تابع امتیاز g تجزیه پذیر باشد، می توان برای محاسبه میزان تطابق ساختار جدید از محاسبات قبلی استفاده کرد. در حقیقت، فقط مقادیر متغیرهایی حساب می شوند که مجموعه والدین آنان تغییر کرده باشند.

تعریف ۵- تابع امتیاز g را هم ارز می نامند، هر گاه به همه گراف

در نتیجه بر اساس قانون زنجیره ای احتمال کل، توزیع احتمال توأم مجموعه متغیرهای تصادفی X_1, X_2, \dots, X_n از رابطه (۱) به دست می آید.

(۱)

$$P(X_1, \dots, X_n) = P(X_1) \cdot P(X_2 | X_1) \times \dots \times P(X_n | X_1, X_2, \dots, X_{n-1}) \\ = \prod_i P(X_i | X_1, X_2, \dots, X_{i-1})$$

در شبکه های بیزی مقدار یک گره خاص فقط به مقادیر گره های والد خود مشروط است، با فرض اینکه

$$Parent(X_i) \subseteq \{X_1, X_2, \dots, X_{i-1}\} \quad ,$$

رابطه ۱ را می توان به صورت ذیل بیان کرد.

(۲)

$P(X_1, X_2, \dots, X_n) = \prod_i P(X_i | Parent(X_i)) = \prod_i P(X_i | Pa_{X_i})$
معمولاً شبکه بیزی شامل اطلاعات مقادیر احتمالات شرطی هستند. در نتیجه، می توان برای تحلیل داده ها از طریق شبکه های بیزی گام های زیر را می توان پیشنهاد کرد:

گام اول: شناسایی متغیرهای ضروری

در این گام باید مشخص شود متغیرهای مسئله بیان کننده چه چیزی هستند، چه مقادیری می توانند بگیرند یا می توانند در چه موقعیتی قرار بگیرند. مقادیر مربوط به متغیرها، علاوه بر اینکه باید منحصر به فرد باشند، بایستی جامع و فراگیر نیز باشند؛ بدین معنی که در یک زمان دقیقاً باید یکی از این ارزش ها را بگیرند.

گام دوم: شناخت تعداد یال ها

یک مجموعه از یال های جهت دار که اگر یک یال از رأس X به رأس Y باشد، X را والد Y می نامیم.

گام سوم: شناخت تعداد گره ها

هر گره X_i ، یک توزیع احتمال شرطی $P(X_i | Pa_{X_i})$ دارد که تأثیر گره های والد بر این گره را به صورت عددی نشان می دهند.

بعد از شناخت و انجام گام های پیش گفته در شبکه های بیزی، موضوع یادگیری ماشین بسیار اهمیت دارد. یادگیری در شبکه های بیزی فرایندها و روش هایی هستند که یک شبکه، با استفاده از آنها، به صورت خودکار می تواند ایجاد شود و عمل کرد خود را بهبود بخشد. یادگیری شبکه های بیزی از روی داده ها یک فرایند دو مرحله ای شامل یادگیری ساختاری و یادگیری پارامتری

نوع داده‌ها است. باید توجه کرد که توزیع توأم متغیرهای تصادفی در شبکه‌های بیزی به صورت ذیل محاسبه می‌شود:

$$P(X_1, X_2, \dots, X_n) = \prod_i P(X_i | Pa_{X_i})$$

بنابراین، پارامترهایی که باید برآورد شوند، توزیع‌های شرطی $P(X_i | Pa_{X_i})$ هستند. اگر متغیرها گسسته باشند، بردار

پارامترها به نمودار

$$P = \{p_{ijk} | i = 1, \dots, n; j = 1, \dots, q_i; k = 1, \dots, r_i\}$$

تعریف می‌شود که در آن n تعداد متغیرها، r_i تعداد سطوح متغیر نام و q_i تعداد ترکیبات مختلف سطوح متغیر والد است. شناخت توزیع پاسخ در یک مدل شبکه بیزی بسیار مهم است. در مطالعه‌ای که برای عفونت اداری انجام گرفته است توزیع متغیر نوع عفونت یک پاسخ پنج جمله‌ای است. لذا در ادامه، ابتدا تحلیل بیزی توزیع چندجمله‌ای در مدل شبکه‌های بیزی را بررسی می‌کنیم.

تحلیل بیزی رگرسیون لوژیستیک بر اساس توزیع k جمله‌ای

از آنجا که داده‌های این مطالعه به بررسی سیستم ایمنی نوع عفونت اداری مربوط است، توزیع متغیر پاسخ یک توزیع پنج-جمله‌ای است. در این حالت فرض کنید بردار Y ، متغیر پاسخ چندجمله‌ای با $k+1$ سطح است که با $0, 1, \dots, k$ نشان داده می‌شوند، باشد. همچنین فرض کنید برای $i = 1, 2, \dots, I$ Y_i دارای توزیع چندجمله‌ای با بردارهای احتمال یعنی $p_i = (p_{i0}, \dots, p_{ik})$ باشد.

$$Y_i' \sim Mult(N_i, p_i)$$

با N_i ثابت است. ماتریس تعداد مشاهدات $Y = (y_{ij})$ برای $i = 1, 2, \dots, I$ و $j = 0, 1, \dots, k$ شامل داده‌هاست. ستون‌های Y با (y_0, y_1, \dots, y_k) نشان داده می‌شود، به طوری که $y_0 = N - \sum_{j=1}^k y_j$ که

$$N = (N_1, \dots, N_I)'$$

تابع درست‌نمایی به صورت ذیل است:

$$L(P|Y = y) = \left(\prod_{i=1}^I \frac{N_i!}{\prod_{j=0}^k y_{ij}} \right) \exp\left(\sum_{i=1}^I \sum_{j=0}^k y_{ij} \log(p_{ij}) \right)$$

زمانی که

$$\sum_{j=0}^k p_{ij} = 1, \forall i \text{ و } P = (p_{ij}), 0 < p_{ij} < 1$$

بدون دور جهت‌داری که گراف کالبد یکسان دارند مقدار یکسانی تخصیص دهد.

بدین ترتیب، نتایج ارزیابی برای یک کلاس هم‌ارزی (تمام گراف بدون دورهایی که گراف کالبد یکسان دارند) صرف‌نظر از اینکه کدام یک از اعضای آن انتخاب شده باشند، یکسان است. این امر باعث کاهش فضای تحت بررسی می‌شود. در یک تقسیم‌بندی کلی امتیازهای مربوط به شبکه بیزی در دو گروه ذیل قرار می‌گیرند.

- امتیازهای وابسته به توزیع پیشین که طبق آمار بیزی کار می‌کنند.

- امتیازهای وابسته به مفاهیم نظریه اطلاعات

یکی دیگر از مطالبی که در ادامه کانون توجه قرار می‌گیرد امتیازهای وابسته به توزیع پیشین است. در این حالت، معمولاً توابع براساس یک رابطه بیزی عمل می‌کنند، به طوری که ابتدا توابع با توزیع احتمال پیشین روی شبکه‌های ممکن شروع می‌شوند، سپس با در نظر گرفتن داده‌های موجود و توزیع پیشین روی پارامترها توزیع احتمال پسین را محاسبه می‌کنند. در نهایت، شبکه‌ای که بیشترین مقدار احتمال پسین را کسب کند، انتخاب می‌شود. در این روش، هدف پیدا کردن آن مدلی است که با توجه به امتیاز مورد نظر، احتمال رخداد G به شرط داده‌های D یا همان $P(G|D)$ را ماکسیمم کند. با توجه به قانون بیزی داریم:

(۴)

$$P(G|D) = \frac{P(D|G)P(G)}{P(D)}$$

از آنجا که احتمال $P(D)$ از ساختار گراف و انتخاب G مستقل است، می‌توان هدف بالا را معادل بیشینه کردن عبارت $P(G|D) \propto P(D|G)P(G)$ در نظر گرفت. به منظور سهولت در انجام محاسبات، به جای محاسبه $P(G|D)$ می‌توان عبارت $\log(P(G|D))$ را حساب کرد.

هدف از یادگیری پارامتری محاسبه عناصر موجود در جدول توزیع احتمال شرطی مربوط به گره‌های موجود در شبکه بیزی است که ساختار آن به دست آمده است. زمانی که ساختار شبکه بیزی روی n رأس مشخص باشد، می‌توان برآورد پارامتری موجود در جدول توزیع احتمال شرطی مربوط به رأس‌ها را به دست آورد. مساله‌ای که در یادگیری شبکه‌های بیزی وجود دارد

پارامترها می تواند کاندید مناسبی باشد. براساس توزیع نرمال، اگر بنویسیم

$$\beta_j \sim N(\mu_j, \Sigma_j) \text{ و } \Sigma_j = \text{Diag}(\sigma_{j0}, \dots, \sigma_{jq})$$

در نتیجه توزیع پسین به صورت ذیل است:

$$\pi(\beta|Y=y) \propto \prod_{i=1} \left\{ \frac{\exp(\sum_{j=1} (y_{ij} x'_i \beta_j - \frac{1}{2} ((\beta_j - \mu_j)' \Sigma_j^{-1} (\beta_j - \mu_j))))}{(1 + \sum_{j=1} \exp(x'_i \beta_j))^{N_i}} \right\}$$

اکنون، استنباط بیزی براساس تجزیه توزیع پسین است. به طور کلی، پسین بالا فرم بسته معلومی نخواهد داشت. در این حالت، تابع چگالی چندبعدی پیچیده ای خواهد داشت. روش مونت کارلوی زنجیره مارکوفی (MCMC) تکنیکی است که حل این گونه مسائل را آسان می کند. باید توجه کرد که این روش برای محاسبه توزیع پسین و امید ریاضی بعضی از توابع آن به کار برده می شود. فرم چگالی حاشیه ای به صورت ذیل است:

$$m(y) = \prod_i \prod_j \frac{N_i!}{y_{ij}!} \times \prod_i \int_{\beta_1} \dots \int_{\beta_k} \frac{\exp(\sum_{j=1} (y_{ij} x'_i \beta_j - 1/2 ((\beta_j - \mu_j)' \Sigma_j^{-1} (\beta_j - \mu_j))))}{(1 + \sum_{j=1} \exp(x'_i \beta_j))^{N_i}} d\beta_1 \dots d\beta_k$$

و توزیع پسین به نمودار ذیل است:

$$\pi(\beta|Y=y) = \frac{1}{m(y)} \prod_i \left(\frac{N_i!}{\prod_j y_{ij}!} \frac{\exp(\sum_{j=1} (y_{ij} x'_i \beta_j - 1/2 ((\beta_j - \mu_j)' \Sigma_j^{-1} (\beta_j - \mu_j))))}{(1 + \sum_{j=1} \exp(x'_i \beta_j))^{N_i}} \right)$$

حال برای برآورد بیزی β_j ، کافی است که مخاطره پسین $E(L(\beta_j, \hat{\beta}_j)|Y=y)$ را کمینه کنیم. با توجه به اینکه فرم توزیع پسین بسته نیست، لذا نمی توان میانگین توزیع پسین را محاسبه کرد. برای این منظور از شاخص دیگری که با عث بیشینه شدن توزیع پسین می شود استفاده خواهیم کرد. لذا تحت تابع زیان ۱-۰ ذیل داریم:

$$l(\beta_j, \hat{\beta}_j) = \begin{cases} 0 & |\beta_j - \hat{\beta}_j| < C \\ 1 & |\beta_j - \hat{\beta}_j| \geq C \end{cases}$$

پس $\hat{\beta}_j$ میانه فاصله $I = \{\hat{\beta}_j, |\beta_j - \hat{\beta}_j| \leq C\}$ که معادل با بیشینه کردن $P(\hat{\beta}_j \in I|y)$ است. در این مورد، برآورد بیزی مدل توزیع پسین است. برای ساده تر شدن محاسبات با لگاریتم توزیع پسین کار می کنیم:

حال فرض کنید برای $j = 0, 1, \dots, k$ و بردار پارامترهای رگرسیون بر طبق

$\beta'_j = (\beta_{j0}, \beta_{j1}, \dots, \beta_{jq})$ متغیر کمکی $x'_i = (1, x_{i1}, \dots, x_{iq})$ است. براساس لگاریتم

بخت ها $\log(\frac{P_{ij}}{P_{i0}}) = \eta_{ij}$ ساختار ذیل را برای رگرسیون لوژستیک در نظر می گیریم:

$$\eta_{ij} = x'_i \beta_j, i = 1, \dots, I$$

یا برای $\eta_j = (\eta_{1j}, \dots, \eta_{Ij})$ می توان نوشت $\eta_j = X \beta_j$ که

$$X = \begin{pmatrix} x'_1 \\ \vdots \\ x'_I \end{pmatrix}$$

تابع درست نمایی کانونی برای $\beta = (\beta_1, \dots, \beta_k)$ عبارت است از:

(۵)

$$L(\beta|Y=y) = \left(\prod_i \prod_j \frac{N_i!}{y_{ij}!} \right) \exp(\sum_i \sum_j y_{ij} x'_i \beta_j - \sum_i N_i \log(1 + \sum_j \exp(x'_i \beta_j)))$$

به سادگی، می توان نتیجه گرفت که از رابطه ۵ نمی توان یک ساختار بسته برای برآورد مقادیر مورد انتظار در هر خانه به دست آورد. برای به دست آوردن یک جواب هم گرا به مقادیر مورد انتظار در هر خانه به سراغ الگوریتم های تکراری برازش مانند الگوریتم نیوتن-رافسون می رویم.

اکنون فرض کنید بردار پارامترها را با β نمایش دهیم و $L(\beta)$ را به عنوان لگاریتم تابع درست نمایی حاصل از مدل (۵) توزیع چند جمله ای مدل (۵) در نظر بگیریم؛ پس داریم:

(۶)

$$L(\beta) = \log \left(\frac{N!}{\prod_{ijk} y_{ijk}!} \right) + \sum_{ijk} y_{ijk} (\sum_l x_{ijk l} \beta_l - \log N) - N \left(\frac{1}{N} \sum_{ijk} \exp(\sum_l x_{ijk l} \beta_l) - 1 \right)$$

با استفاده از مدل رگرسیون لوژستیک ارائه شده در (۶)، از دیدگاه بیزی از توزیع پیشین برای β_j استفاده می کنیم که احتمالی را بر مقدار β_j نزدیک μ_j طراحی می کند؛ زیرا نمی توان انتظار داشت هر متغیر بر متغیر پاسخ نفوذ کمکی خاص کند. در واقع، برای آنکه ساختار توزیع پیشین طوری باید انتخاب شود که عدم تأثیر متغیر کمکی در ابتدا مورد توجه قرار گیرد، توزیع پیشین نرمال برای

در ادامه به بررسی هر یک از مراحل اجرا شده برای داده‌ها براساس روش تحقیق پیشنهادی در جهت رسیدن به مدلی کارا برای دادن یک مدل در خصوص سیستم ایمنی بدن و برآورد احتمال ابتلا به عفونت کلیه و مجاری ادراری (دستگاه ادراری) با توجه به عوامل تأثیرگذار بر آن می‌پردازیم. در این روش سعی شده با استفاده از مدل شبکه‌بیزی و الگوریتم‌های منتسب به آن به الگویی مناسب جهت تشخیص زودهنگام عفونت دستگاه ادراری دست پیدا کنیم. این الگوریتم با استفاده از نتایج آزمایش‌های اولیه پیاده‌سازی شده است که می‌توانند نتایج را در همان زمان اولیه و بدون نیاز به انجام کشت‌های پرهزینه و زمان‌بر که تشخیص را به تأخیر می‌اندازند، صورت گیرد. الگوریتم شامل پنج گام اساسی زیر است (Rasoulzadeh et al., 2017):

الف: شناسایی سیستم شامل مطالعه بیماری عفونت کلیه و مجاری ادراری و فاکتورهای مؤثر در ابتلا به بیماری

ب: آماده‌سازی داده‌ها که شامل جمع‌آوری داده‌ها و ثبت آنان و حذف مقادیر غیرضروری

پ: مدل‌سازی مناسب با استفاده از روش‌های مختلف شبکه‌های بیزی

ج- ارزیابی نتایج حاصل از مدل پیشنهادی که شامل محاسبه توانایی مدل، میزان حساسیت مدل نسبت ده مدل و تأیید نتایج براساس نظر متخصصین

د- تعمیم نتایج به کل داده‌های موجود در جامعه آماری و ارائه یک پروتکل پیشنهادی

شیوه جمع‌آوری داده‌ها

در این گام به جمع‌آوری داده‌های حاصل از دو آزمایش UA و UC از پرونده‌های بیماران پرداخته می‌شود. این اطلاعات از بانک اطلاعاتی مربوط به وزارت بهداشت و درمان پزشکی- معاونت درمان دانشگاه علوم پزشکی قزوین- مرکز بهداشت و درمان شهرستان تاکستان جمع‌آوری شده است. کلیه داده‌های مربوط به این پژوهش از طریق فرم‌های الکترونیکی که در این مرکز ذخیره شده بوده به دست آمده است. اطلاعات مربوط به بیماران با رعایت اصل محرمانگی و بدون مشخصات شناسایی جمع‌آوری شد. در این فرآیند اطلاعات مربوط به ۴۵۵۲ مراجعه‌کننده در اختیار قرار داده شد که از این تعداد در حدود ۴۰۰ مورد به علت‌هایی همچون کم‌بودن نمونه تحویل داده‌شده و یا نداشتن نمونه ادرار و یا آلوده-

$$\log(\pi(\beta|Y=y)) = A +$$

$$\sum_{i=1}^n \sum_{j=1}^m (y_{ij} x'_i \beta_j - \frac{1}{2} ((\beta_j - \mu_j)' \Sigma_j^{-1} (\beta_j - \mu_j))) - \sum_{i=1}^n N_i \log(1 + \sum_{j=1}^m \exp(x'_i \beta_j))$$

که در آن:

$$A = \log\left(\frac{1}{m(y)} \prod_i \left(\frac{N_i!}{\prod_j y_{ij}!}\right)\right)$$

و اما برآورد بیزی به صورت ذیل است:

$$\hat{\beta}_j = \mu_j + \frac{1}{I} \sum_j \left(\sum_{i=1}^n a_{ij} x_i\right) = \mu_j + \frac{1}{I} \sum_j X'(Y_j - N_i P_j).$$

$$a_{ij} = y_{ij} - N_i P_{ij} \quad \text{که}$$

اکنون برای مقایسه بیزی مدل‌ها با عامل بیزی لازم است که چگالی حاشیه‌ای را برای دو مدل محاسبه کنیم. برای این کار، اگر بخواهیم به روش مونت کارلو عمل کنیم:

$$m(y) = E_{\pi(\beta)}(L((\beta_1, \dots, \beta_k)|Y=y)) \cong \frac{1}{n} \sum_{i=1}^n L((\beta_{i,1}, \dots, \beta_{i,k})|Y=y).$$

اما این تقریب برای چگالی حاشیه‌ای، در زمانی که توزیع پیشین پراکنده است یا زمانی که تابع درست‌نمایی از توزیع پیشین متمرکزتر است، یک تقریب استوار و محکم نیست. در این صورت، واریانس برآوردگر بزرگ خواهد بود و هم‌گرایی نیز به کندی صورت می‌گیرد. یک راه دیگر برای تقریب چگالی حاشیه‌ای تقریب لاپلاس است که برای مدل‌های d بعدی استفاده می‌شود. تقریب توزیع حاشیه‌ای با استفاده از تقریب لاپلاس برای چگالی حاشیه‌ای عبارت است از:

$$\log(m(y)) \cong \frac{1}{2} d \log(2\pi) + \frac{1}{2} \log|H^*| + \log(L(\beta^*|Y=y)\pi(\beta^*))$$

که در آن β^* برداری از مدل توزیع پسین برآورد β است و

$$H^* \text{ معکوس ماتریس هیشین } \left(\frac{\partial^2 h(\beta, y)}{\partial \beta \partial \beta'}\right)$$

در مقدار β^* است. $h(\beta, y) = \log(L(\beta|Y=y)\pi(\beta))$

پس، برای مقایسه مدل‌های M_i و M_j تقریبی از عامل بیزی را به صورت ذیل به کار می‌بریم. مقادیر بزرگ عامل بیزی بیان می‌کند که مدل M_i نسبت به مدل M_j بهتر می‌تواند به داده‌ها برازش داده شود:

$$B_{ij} \cong \frac{L(\beta_i^*|Y=y)}{L(\beta_j^*|Y=y)} \left| \frac{H^*(\beta_i^*)}{H^*(\beta_j^*)} \right|^{1/2} \left(\frac{1}{2\pi}\right)^{(d_j-d_i)/2}$$

کاربرد مدل شبکه‌های بیزی در بررسی عفونت ادراری

آموزش و آزمایش به طور تصادفی انتخاب شده‌اند. متغیرهای ورودی به ترتیب عبارتند از جنسیت (Sex)، سن افراد (Age)، رنگ (Color)، وزن مخصوص (S.G.)، شفافیت (Clarity)، خاصیت اسیدی (PH)، گلبول سفید (w.b.c)، گلبول قرمز (r.b.c)، سلول اپیتال (Ep.cell)، باکتری (Bacter)، موکوس (Mucus)، نیتريت (Nitrit)، خون (Blood)، کلنی (Colony)، و متغیر پاسخ نوع عفونت (Tashkhis) نیز است.

برای بررسی و شناخت متغیرهای مؤثر از طریق مدل شبکه بیزی عمل می‌شود، به طوری که برای پیش‌بینی از طریق شبکه بیزی، امکان تعیین میزان وابستگی متغیرهای بر یکدیگر و نیز به متغیر پاسخ امکان‌پذیر است. مدل شبکه بیزی به ما امکان ساخت یک مدل احتمالی از متغیرها را می‌دهد، تا بدین وسیله احتمال وقوع رخدادها از مجموعه ویژگی‌ها به دست آید. روابط سببی بین گره‌ها می‌تواند به وسیله شبکه بیزی نمایش داده شود، که با استفاده از آن می‌توان احتمال قرارگرفتن یک شخص را جزء دسته‌ای از بیماران محاسبه کرد. دو مدل معروف در شبکه بیزی به صورت مارکوفی و درختی است. در مدل شبکه بیزی با الگوریتم مارکوفی متغیرهای فقط به یک زمان قبل از خود وابسته هستند. اما در مدل شبکه بیزی با الگوریتم درختی، همان‌طور که از نام آن پیدا است وابستگی به صورت ریشه‌ای است و محدودیتی وجود ندارد.

نتایج

تحلیل داده‌ها با استفاده از مدل سازی شبکه بیزی الگوریتم درختی

با توجه به متغیرهای ورودی و متغیر پاسخ که در جدول ۱ آنان را تعریف کردیم و براساس مدل شبکه بیزی الگوریتم درختی، نمودار ۱ برای دو مدل شبکه بیزی درختی با ساختار گام به گام ورود متغیرها و انتخاب توأم متغیرها ارائه شده است. در هر دو مدل سهم تأثیر متغیرها بر میزان عفونت ادراری تعیین شده است. این مدل شبکه بیزی قابلیت دارد که به منرله تشخیص اولیه استفاده شود. به این منظور، ابتدا یک مدل درخت تکمیل شده ایجاد می‌کنیم و با ایجاد تغییرات لازم در پارامترها نتایج حاصل را با هم مقایسه می‌کنیم. در نهایت، دو مدل شبکه بیزی درختی قابلیت ایجاد دارد.

شدن نمونه با عوامل محیطی، درخواست تکرار مجدد داده شده بود و ۱۰۰ مورد به علت عدم تکمیل متغیرهای تحت بررسی که از نمونه‌های جمع‌آوری شده حذف شده‌اند. از نمونه باقی‌مانده در حدود ۴۰۵۲ عدد نمونه هم دارای آزمایش تحلیل ادرار و هم دارای آزمایش کشت ادرار بوده‌اند. اطلاعات موجود در بانک اطلاعاتی مرکز بهداشت و درمان دارای ۲۴ متغیر بود که از این تعداد متغیرهایی که مربوط به اهداف پروژه بود با استفاده از نظر متخصص آزمایشگاه و متخصص ارولوژی و مطالعات صورت گرفته جمع‌آوری شد. این تعداد ۱۵ شاخص است و به نحوی با فرآیند تشخیص بیماری مرتبط است که در جدول ۱ ارائه شده است. دیگر موارد به دلیل عدم کاربرد در این پژوهش حذف شده‌اند. در آماده‌سازی داده‌ها حدود ۶۸۸۸۴ رکورد مربوط به ۱۵ شاخص ایجاد شد که مربوط به بیماری عفونت کلیه و مجاری ادراری و سیستم ایمنی بدن در آن ناحیه است که به صورت فایل در محیط اکسل قابل استفاده برای انجام فرآیند داده کاوی و مدل سازی قابل استفاده است.

پیاده‌سازی مدل شبکه‌های بیزی

در این مرحله، با استفاده از روش زیر نمونه تصادفی داده‌ها به دو بخش آموزش و آزمایشی برای ارزیابی مدل ایجاد شده تقسیم می‌شوند. در مورد تعداد رکوردهای آموزش و آزمون می‌توان گفت که اگر تعداد آموزشی زیاد باشد مدل ساخته شده به واقعیت نزدیک‌تر خواهد شد ولی نمی‌توان به دقت مدل حاصل اطمینان داشت. حال، اگر تعداد رکوردهای آزمون زیاد باشد در این صورت مدل قابل اطمینان است، اما این مدل، از آنجاکه باید هوشمند باشد و وجه آموزشی لازم را در خود ایجاد کند، به نظر می‌رسد این آموزش به اندازه کافی فراهم نشده است یا به تعبیری یادگیری خوب انجام نشده است. به همین دلیل، انتخاب این درصدها با نظر متخصص آمار در حوزه یادگیری و براساس روال کارهای گذشته صورت می‌گیرد. در این پژوهش سهم بین داده‌های آموزشی و آزمایشی ۳ به ۱ است. در واقع ۷۵ درصد داده‌ها برای آموزش و ۲۵ درصد برای آزمون و اعتبار مدل استفاده شده است. داده‌های بخش آموزش با عث می‌شوند مدل به طور هوشمند خود را به روز کند و مدل بهینه را ایجاد کند. داده‌های بخش آزمون نیز مدل ایجاد شده را ارزیابی می‌کند. نکته قابل توجه این است که داده‌های انتخاب شده برای هر دو قسمت

بیماری را نداشته‌اند را بیمار و در حدود ۲/۰ درصد افراد بیمار را سالم تشخیص داده است همان‌طور که در جدول ۱ مشخص است الگوریتم توانسته است تمام رکوردها را دسته‌بندی کند. بررسی‌ها نشان داد که در مدل شبکه بیزی الگوریتم درختی با انتخاب توأم متغیرها، متغیر افزایش باکتری و میزان کاهش گلبول‌های سفید (wbc) و سن افراد مراجعه‌کننده موجود در این قسمت با توجه به نتایج ارائه‌شده در جدول ۱ درصد تشخیص درست در شرایط با انتخاب گام به گام متغیرها برای مدل شبکه بیزی با ساختار الگوریتم درختی بهتر از با انتخاب توأم متغیرها عمل می‌کند.

تحلیل داده‌ها با استفاده از مدل شبکه بیزی با الگوی مارکوفی

به‌طور کلی، مدل‌سازی با شبکه‌های بیزی به فرضیه‌هایی از زنجیره مارکوف نیاز دارد. بنابر خاصیت مارکوف در توزیع احتمال شرطی، هر رأس فقط به والدین خود وابسته است. به عبارت دیگر تمام وابستگی‌های مستقیم توسط یال‌ها مشخص شده است. در واقع، هر استقلال با فقدان یک یال در سیستم مشخص می‌شود. در شبکه‌های بیزی، که واجد ویژگی مارکوف هستند، هر رأس به شرط والدینش از مجموعه رأس‌های غیرفرزند خود مستقل خواهد بود. این موضوع معادل برقراری ویژگی مارکوف در شبکه‌های بیزی است. مفهوم مفید دیگر در شبکه‌های بیزی، پوشش مارکوف یک گره است که شامل والدین گره، فرزندان آن و والدین فرزندانش (همسران) است. برای ایجاد مدلی مانند مدل بخش قبل، ابتدا باید گره جریان را ایجاد کنیم و مدل شبکه بیزی با الگوی مارکوفی را فراخوانی کرده و پارامترهای مد نظر در مدل را تنظیم کنیم.

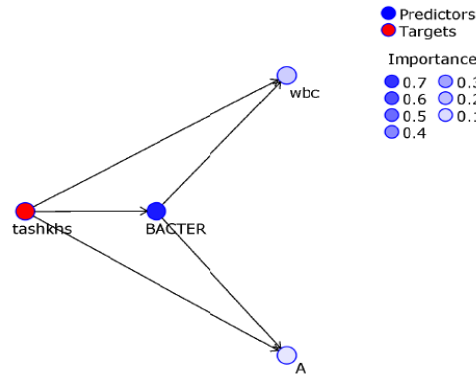
به دلیل اینکه می‌توان چندین نوع مختلف از شبکه بیزی پوشش مارکوفی را با تنظیم پارامترها ساخت، باید بررسی کرد کدام روش بهترین خروجی را دارد. این موضوع در نمودار ۲ ارائه شده است. به این منظور، ابتدا یک مدل شبکه بیزی با الگوریتم مارکوفی ایجاد و مشخص شد دو مدل شبکه بیزی با الگوریتم مارکوفی با انتخاب گام به گام متغیرها و با انتخاب توأم متغیرها تایید شد. در این حالت معیاری لازم است تا با استفاده از آن بتوان مدل‌ها را ارزیابی کرد و کاراترین مدل را انتخاب کرد. لازم به گفتن است برای محاسبه دقت رده‌بندی‌کننده پس از ساخت مدل، اگر از داده‌های

همان‌طور که درگراف‌های حاصل از شبکه بیزی درختی مشاهده می‌شود، متغیرهایی با گره آبی پررنگ تا کم رنگ سهم اثرگذاری هر پارامتر را در متغیر پاسخ تشخیص عفونت نشان می‌دهد. بعد از ایجاد مدل شبکه بیزی، به مرحله ارزیابی مدل ایجادشده می‌رسیم. در روش پیش‌گویانه می‌توان دقت مدل را برآورد کرد. در واقع، دقت مدل معیاری است که با استفاده از آن می‌توان مدل‌ها را ارزیابی کرد و کاراترین مدل را انتخاب کرد برای محاسبه دقت یک رده‌بندی‌کننده، اگر پس از ساخت مدل شبکه بیزی از داده‌های آموزشی استفاده شود، نتایج بسیار خوش‌بینانه خواهد شد. برای رفع این مشکل از داده‌های ازمونی نیز برای برآورد دقت مدل استفاده می‌کنیم. دقت مدل ایجادشده به صورت درصد بیان می‌شود.

ملاک تشخیص مدل بوتو: فرض کنید در مدل پیشنهادی شبکه بیزی تعداد رکوردهای مربوط به افراد سالم، که توسط مدل شبکه بیزی پیشنهادی از میان افراد سالم درست تشخیص داده شده است (a) و تعداد رکوردهای مربوط به افراد بیمار، که توسط مدل شبکه بیزی پیشنهادی از میان افراد بیمار درست تشخیص داده شده است (b) باشد. همچنین تعداد کل افراد در نمونه n باشد، در آن صورت، دقت مدل شبکه بیزی عبارت است از:

$$= \frac{a + b}{n}$$

از ملاک پیش‌گفته برای مدل شبکه بیزی شبیه‌سازی شده می‌توان استفاده کرد. اکنون برای ارزیابی مدل پیشنهادی می‌توان از ملاک تشخیص استفاده کرد. با توجه به آنچه که ارائه شده است می‌توان گفت مدل شبکه بیزی الگوریتم درختی با انتخاب توأم متغیرها در حدود ۷/۹۹ درصد افرادی را که به بیماری عفونت دچار نبوده‌اند و ۹۹ درصد افراد بیمار را درست پیش‌گویی کرده است و تنها حدود ۳/۰ درصد از افرادی را که این بیماری را نداشته‌اند بیمار و در حدود ۱ درصد افراد بیمار را سالم تشخیص داده است. همان‌طور که در جدول ۱، براساس تعریف ۴ مشخص است، الگوریتم توانسته تقریباً تمام رکوردها را به درستی دسته‌بندی کند. در جدول ۱ همچنین مدل شبکه بیزی الگوریتم درختی با انتخاب گام به گام متغیرها در حدود ۷/۹۹ درصد افرادی را که به بیماری عفونت دچار نبوده‌اند و ۸/۹۹ درصد افراد بیمار را درست پیش‌گویی کرده است و تنها حدود ۳/۰ درصد از افرادی که این



شکل ۱- گراف ایجاد شده به وسیله دو مدل شبکه بیزی با استفاده از الگوریتم درختی.
Fig. 1. Graph for both Bayesian Network using Tree Algorithm.

همچنین، در مدل شبکه بیزی الگوریتم درختی با انتخاب گام به-گام، متغیر باکتری و wbc و سن افراد مراجعه کننده موجود در جامعه آماری به ترتیب بیشترین اثر را در فرایند تشخیص نسبت به باقی متغیرها داشته‌اند و باقی متغیرها از نظر آماری معنی دار نیستند. بقیه متغیرها از فرایند مدل‌سازی شبکه بیزی مارکوفی کنار گذاشته شده‌اند. در این قسمت با توجه به نتایج ارائه شده در جدول ۲ درصد تشخیص درست در شرایط با انتخاب گام به گام متغیرها برای مدل شبکه بیزی با ساختار الگوریتم مارکوفی بهتر از انتخاب توأم متغیرها عمل می‌کند.

بحث

مقایسه بین مدل‌های شبکه بیزی با الگوریتم درختی و مارکوفی

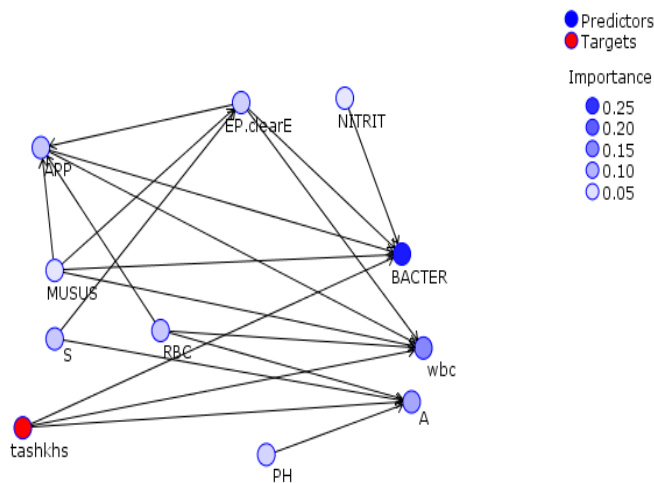
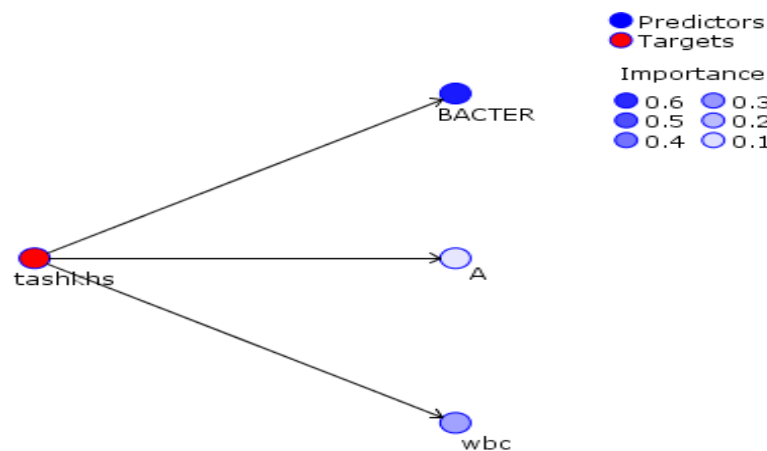
برای بررسی و مقایسه مدل‌های شبکه‌های بیزی ایجاد شده با الگوریتم‌های درختی و مارکوفی با یکدیگر و تشخیص اینکه برای رسیدن به یک مدل بهتر کدام یک می‌تواند در اولویت انتخاب قرار بگیرد، از طریق ابزارهای مهمی صورت می‌گیرد. یکی از این ابزار رسم نمودار مربوط به روند تغییرات بیماری در خصوص تغییرات داده‌های آموزشی و آزمایشی در مدل‌های شبکه بیزی براساس دو الگوریتم درختی و مارکوفی است. نمودار ۳ این موضوع را براساس الگوریتم درختی مورد توجه قرار می‌دهد. درصد برازش بین نمودار حاصل از مدل و داده‌های واقعی در این نمودارها به خوبی دیده می‌شود و می‌توان بیان کند که الگوی درختی می‌تواند به‌عنوان یک الگوی مناسب مورد استفاده قرار گیرد.

آموزشی استفاده شود، نتایج بسیار خوش‌بینانه خواهد بود. برای حل این موضوع، نموداری از داده‌های آزمون برای برآورد دقت مدل استفاده می‌کنیم. نتایج نشان می‌دهد مدل شبکه بیزی مارکوفی با انتخاب توأم متغیرها در حدود ۳/۹۹ درصد افرادی را که به بیماری عفونت دچار نبوده‌اند و ۲/۹۵ درصد افراد بیمار را درست پیش‌گویی کرده است و تنها حدود ۴/۰ درصد از افرادی را که این بیماری را نداشته‌اند بیمار و در حدود ۲/۳ درصد افراد بیمار را سالم تشخیص داده است. همان‌طور که در جدول ۲ مشخص است الگوریتم منطقه غیرقابل رده‌بندی دارد زیرا مقدار مثبت درست و منفی نادرست برابر جمع تعداد افراد بیمار نیست؛ یعنی الگوریتم نتوانسته تمام رکوردها را دسته‌بندی کند. همچنین می‌توان بیان کرد مدل شبکه بیزی مارکوفی با انتخاب گام به گام متغیرها در حدود ۷/۹۹ درصد افرادی را که به بیماری عفونت دچار نبوده‌اند و ۹۹ درصد افراد بیمار را درست پیش‌گویی کرده است و تنها حدود ۳/۰ درصد از افرادی که این بیماری را نداشته‌اند را بیمار و در حدود ۱ درصد افراد بیمار را سالم تشخیص داده است. همان‌طور که در جدول ۲ مشخص است، الگوریتم منطقه غیر قابل رده‌بندی ندارد زیرا مقدار مثبت درست و منفی نادرست برابر جمع تعداد افراد بیمار است یعنی الگوریتم نتوانسته تمام رکوردها را دسته‌بندی کند. بررسی‌ها نشان داد در مدل شبکه بیزی الگوریتم مارکوفی متغیر باکتری و wbc و سن و app و جنسیت افراد مراجعه کننده موجود در جامعه آماری به ترتیب بیشترین اثر را در فرایند تشخیص نسبت به باقی متغیرها داشته است. یک‌سری از متغیرها نیز از فرایند تشکیل مدل کنار گذاشته شده‌اند.

جدول ۱ - ماتریس توافقی برای مدل شبکه بیزی الگوریتم درختی.

Table 1. Contingency Matrix for Bayesian Network using Tree Algorithm.

با انتخاب توأم متغیرها	کلاس تشخیص پیش‌بینی شده مدل شبکه بیزی الگوریتم درختی با انتخاب توأم متغیرها						
		عدم تشخیص	فراوانی	سالم	فراوانی	بیمار	فراوانی
کلاس تشخیص براساس داده‌های واقعی	سالم	۰	%۰	۳۴۱۹	%۷/۹۹	۱۱	۰/۳%
	بیمار	۰	%۰	۶	%۰/۱	۶۱۶	%۹۹
با انتخاب گام به گام متغیر	کلاس تشخیص پیش‌بینی شده الگوریتم درختی با انتخاب گام به گام متغیرها						
		عدم تشخیص	فراوانی	سالم	فراوانی	بیمار	فراوانی
کلاس تشخیص براساس داده‌های واقعی	سالم	۰	%۰	۳۴۲۰	%۷/۹۹	۱۰	۰/۳%
	بیمار	۰	%۰	۱	۰/۲%	۶۲۱	%۸/۹۹



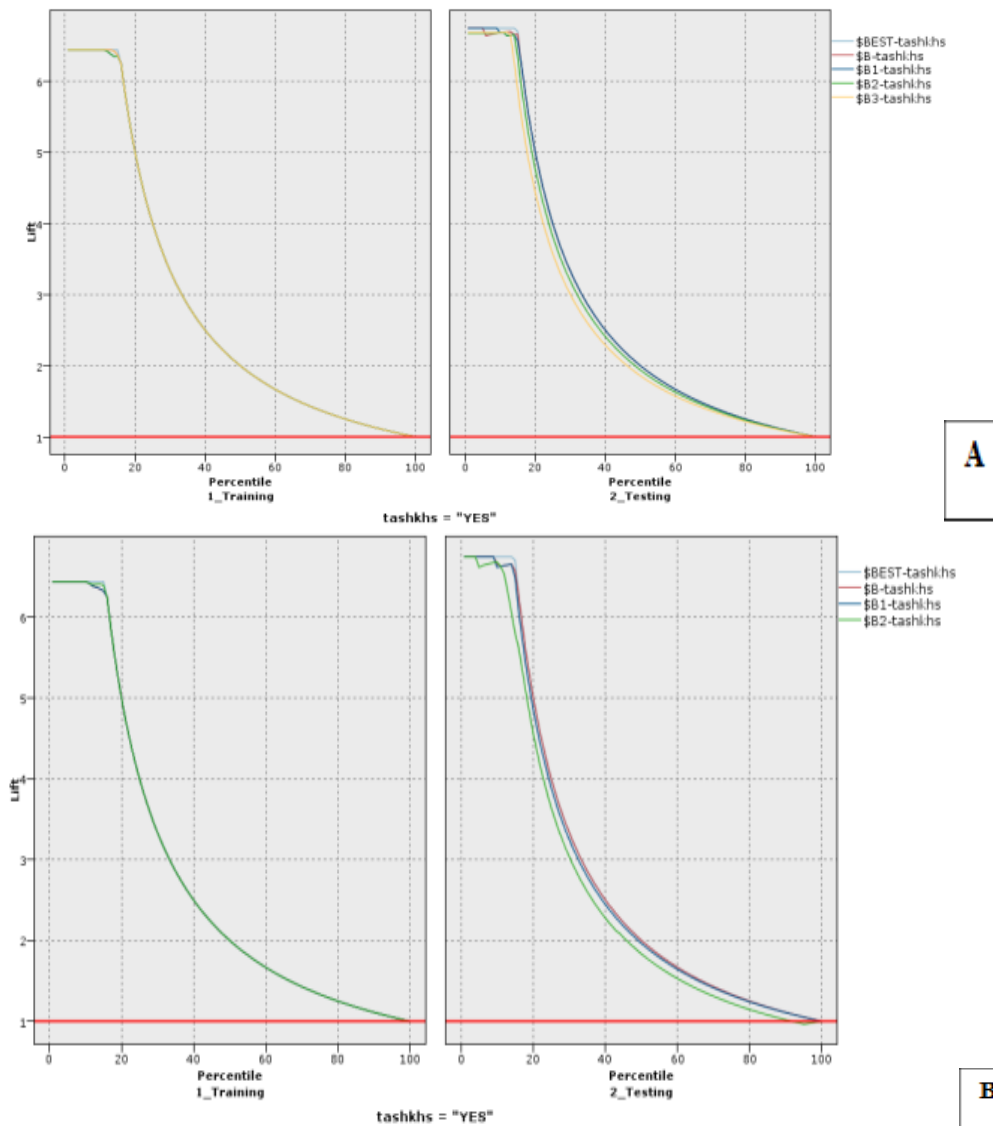
شکل ۲ - گراف ایجادشده به وسیله دو مدل شبکه بیزی با استفاده از الگوریتم مارکوفی.

Fig. 2. Graphs generated by two Bayesian Networks using Markov Algorithm.

جدول ۲- ماتریس توافقی برای مدل شبکه بیزی مارکوفی.

Table 2. Contingency matrix for Bayesian Network using Markov Algorithm.

با انتخاب توأم متغیرها		کلاس تشخیص پیش بینی شده مدل شبکه بیزی مارکوفی با انتخاب توأم متغیرها					
		عدم تشخیص	فراوانی	سالم	فراوانی	بیمار	فراوانی
کلاس تشخیص براساس داده های واقعی	سالم	۸	۰/۲٪	۳۴۰۷	٪۳/۹۹	۱۵	۰/۴٪
	بیمار	۱۰	٪۶/۱	۲۰	٪۲/۳	۵۹۲	٪۲/۹۵
با انتخاب گام به گام متغیرها		کلاس تشخیص پیش بینی شده مدل شبکه بیزی مارکوفی با انتخاب گام به گام متغیرها					
		عدم تشخیص	فراوانی	سالم	فراوانی	بیمار	فراوانی
کلاس تشخیص براساس داده های واقعی	سالم	۰	٪۰	۳۴۲۰	٪۷/۹۹	۱۰	۰/۳٪
	بیمار	۰	٪۰	۶	٪۱	۶۱۶	٪۹۹



شکل ۳- برازش داده های A: آموزشی و B: آزمایشی مدل های شبکه های بیزی با الگوریتم مارکوفی و درختی نسبت به داده های واقعی.

Fig. 3. Fitting for A: training data and B: testing data Baesian Network using both Markov Algorithm and Tree Algorithm with real data.

REFERENCES

- Aickelin, U. and Cayzer, S.** 2002. The danger theory and its application to artificial immune systems. – ICARIS 2002, 1st int. Conf. on Artificial Immune Systems 141-148.
- Castro, L.N. and Von Zuben, F.J.** 2000. An evolutionary immune network for data clustering. – IEEE Brazilian Symposium on Artificial Neural Networks. pp 84-89.
- Castro, L.N. and Timmis, J.** 2002. Artificial immune systems; a new computational approach. – Springer-Verlag, London, UK.
- Castro L.N. and Von Zuben, F.J.** 2000. The clonal selection algorithm with engineering applications. – GECCO 2000, Genetic and Evolutionary Computation Conference 36-37.
- Dasgupta, D., Yua, S. and Nino, F.** 2011. Recent advances in artificial immune systems: models and applications. – App. Soft Comp. 11: 1574-1587.
- Engelbrecht, A.P.** 2007. Computational intelligence an introduction, artificial immune systems. Second edition. Vol. 5, pp 413-448.
- Graaff, A.J. and Engelbrecht, A.P.** 2006. Optimised coverage of non-self with evolved lymphocytes in an artificial immune system. – Int. J. Comp. Intel. Res. 2: 127-150.
- Heckerman, D.A.** 1996. Tutorial on learning with Bayesian networks. Technical report. – MSR-TR. 95-96.
- Hunt, J.E. and Cooke, D.E.** 1996. Learning using an artificial immune system. – J. Net. Com. App. 19: 189-212.
- Jerne, N.K.** 1974. Towards a network theory of the immune system. – Ann. Immunol. 2: 373-389.
- Kim, J. and Bentley, P.J.** 1999. Negative selection and niching by an artificial immune system for network intrusion detection. – GECCO 1999, Genetic and Evolutionary Computation Conference, pp 149-158.
- Kim, J. and Bentley, P.J.** 2002. Immune memory in the dynamic clonal selection algorithm. – 1st International Conference on Artificial Immune Systems, pp 59-67.
- Nasraoui, O., Dasgupta, D. and Gonzalez.** 2002. The promise and challenges of artificial immune system based web usage mining. – Second International Conference on Data Mining. pp 29-39.
- Pearl, J.** 1990. Bayesian network, UCLA, CL.
- Rasoulzadeh, M., Golmakani, N., Ebrahimzadeh, Zagmi, S. and Nasiri, S.** 2017. Cranberry Effects on Prevention of Urinary Tract Infection. – J. St. Res. Comm. Sabzevar Univ. Med. Sci. 22: 16-22.
- Sarafijanovic, S. and Le Boudec, J.** 2004. An artificial immune system for misbehavior detection in mobile ad-hoc networks with virtual thymus. – 3rd International Conference on Artificial Immune Systems, pp 342-356.
- Shuohao, L., Zhang, J. and Boliang, A.** 2014. An incremental structure learning approach for Bayesian network. – The 26th Chinese Control and Decision Conference. DOI: 10.1109/CCDC.2014.6853036.
- Twycross, J.** 2007. Integrated innate and adaptive artificial immune systems applied to process anomaly detection. – Ph.D. thesis, School of Computer Science. University of Nottingham, UK.

اما رسم نمودار به تنهایی کافی نیست. بعد از انجام این مراحل برای مقایسه میزان درستی تمام مدل‌های شبکه‌های بیزی ساخته شده توسط دو الگوریتم مارکوفی و درختی مورد نظر ایجاد کرده و با توجه به نتایج حاصل از نمودارها و روش‌های اعتبارسنجی که بررسی شد بهترین مدل را می‌توان انتخاب کرد. نتایج ارائه شده در جداول ۱ و ۲ اعلام می‌دارد بهترین عمل کرد را در بین داده‌های آموزشی و داده‌های آزمایشی مدل شبکه‌های بیزی با الگوریتم درختی و انتخاب گام به گام متغیرهای ورودی دارا است.

نتیجه گیری

باتوجه به نزدیکی نتایج حاصل از مدل‌های تحت بررسی، می‌توان آن را به عنوان مکمل در کنار تشخیص آنی بیماری توسط پزشک استفاده کرد، به طوری که باعث افزایش قدرت تصمیم‌گیری پزشک در تشخیص بیماری و تجویز درمان مؤثر برای بیماری و همچنین باعث کاهش خطاهای پزشکی در فرایند تشخیص و جلوگیری از ایجاد آسیب احتمالی در بیمار شود. همچنین نتایج این مطالعه در چارچوب یادگیری ماشین و سیستم‌های هوشمند، می‌تواند برای تشخیص سریع این بیماری و درمان افراد مشکوک به این بیماری در سطح جامعه استفاده شود. بنابراین، با در نظر گرفتن نتایج مشابه و مطابق با مدل شبکه‌های بیزی، می‌توان این روش را برای درمان افراد بدون حضور فیزیکی آنان استفاده کرد. به عنوان یک پیشنهاد پایانی که در این مطالعه می‌توان از آن نام برد این است که به دلیل پیچیده بودن اینگونه مطالعات گاهی اوقات می‌توان با ترکیب الگوهای جاری به ترکیبی از مدل‌های شبکه‌های بیزی دسترسی پیدا کرد که بهتر از انتخاب انفرادی یک مدل عمل می‌کند. این موضوع می‌تواند زمینه تحقیق آینده را فراهم آورد.

سپاسگزاری

از هسته پژوهشی یادگیری آماری دانشگاه علامه طباطبایی و پژوهشگاه دانش‌های بنیادی که با حمایت مادی و معنوی خود در رسیدن به نتایج طرح پژوهشی این فعالیت به نویسنده کمک نمودند سپاسگزاری می‌کنم. از آقای عادل و آقای طاهرخانی که در امور نرم‌افزاری کمک کردند کمال تشکر را دارم.

How to cite this article:

Eskandari, F. 2018. The modeling of body's immune system using Bayesian Networks. – Nova Biologica Rep. 2018: 205-220.

اسکندری، ف. ۱۳۹۷. مدل‌سازی سیستم ایمنی بدن با استفاده از شبکه‌های بیزی. – یافته‌های نوین در علوم زیستی ۱۳۹۷: ۲۲۰-۲۰۵.

