

آشنایی با مبانی داده کاوی و منطق فازی در تشخیص بیماری ها
گروه هدف این کتاب: دانشجویان، محققان و فارغ التحصیلان رشته های مختلف علوم پزشکی،
روان شناسی و علوم تربیتی می باشد.

نویسندگان

فاطمه برزگر

دانشجوی کارشناسی ارشد مهندسی نرم افزار، دانشگاه آزاد اسلامی واحد مشهد

دکتر حسن خانی

پزشک عمومی، کارمند مرکز بهداشت شهرستان بجنورد، دانشگاه علوم پزشکی خراسان شمالی

دکتر علی طالع

فوق تخصص غدد درون ریز و متابولیسم کودکان،

پژوهشگر مرکز تحقیقات متابولیک، پژوهشکده علوم سلولی - مولکولی غدد و

متابولیسم، پژوهشگاه علوم غدد و متابولیسم، دانشگاه علوم پزشکی تهران

دکتر مریم جعفرقلیان

پزشک عمومی، پژوهشگر داوطلب مرکز تحقیقات اختلالات متابولیک،

پژوهشکده علوم سلولی - مولکولی غدد و متابولیسم،

پژوهشگاه علوم غدد و متابولیسم، دانشگاه علوم پزشکی تهران

امان قربانی

کارشناس ارشد شیمی تجزیه، سازمان غذا و دارو، دانشگاه علوم پزشکی خراسان شمالی

طیبه زنا

دانشجوی کارشناسی ارشد فیزیولوژی و اصلاح گداهان دارویی،

دانشگاه آزاد اسلامی واحد بروجرد

زیر نظر

دکتر مریم رزاقی آذر

فوق تخصص غدد درون ریز و متابولیسم کودکان، استاد دانشگاه علوم پزشکی ایران

رئیس بخش غدد بیمارستان کودکان حضرت علی اصغر (ع) دانشگاه علوم پزشکی ایران،

رئیس مرکز تحقیقات اختلالات متابولیک، پژوهشکده علوم سلولی - مولکولی

غدد و متابولیسم، پژوهشگاه علوم غدد و متابولیسم، دانشگاه علوم پزشکی تهران

عنوان و نام پدیدآور: آشنایی با مبانی داده‌کاوی و منطق فازی در تشخیص بیماری‌ها/ نویسندگان
 فاطمهبرزگر... او دیگران؛ زیرنظر مریم رزاقی‌آذر.
 مشخصات نشر: قایمشهر: نشر ونداد، ۱۳۹۵.
 مشخصات ظاهری: ۱۱۰ ص.: مصور.
 شابک: ۹۷۸-۶۰۰-۷۰۱۸-۴۹-۱: ۳۴۰۰۰ ریال
 وضعیت فهرست نویسی: فیبا
 یادداشت: نویسندگان فاطمهبرزگر، حسن‌خانی، علی‌طالع، مریم جعفرقلیان، آمنه قربانی، طیبه زنده‌دل.
 موضوع: کتابنامه: ص. ۱۰۵ - ۱۱۰؛ همچنین به صورت زیرنویس.
 موضوع: تشخیص پزشکی — داده‌پردازی
 موضوع: Diagnosis -- Data processing
 موضوع: الگوریتم‌های مبتنی
 موضوع: Medical protocols
 موضوع: داده‌کاوی
 موضوع: Data mining
 موضوع: منطق فازی
 شناسه افزوده: Fuzzy logic
 شناسه افزوده: برزگر، فاطمه، ۱۳۶۴
 رده‌بندی کنگره: ۱۳۹۵ / ۵۲۱۵ / PC۷۸۷
 رده‌بندی دیویی: ۶۱۶/۰۷۵۰۲۸۵
 شماره کتابشناسی ملی: ۴۳۵۵۸۷۰

نویسندگان

فاطمهبرزگر، دکتر حسن‌خانی، دکترعلی‌طالع، دکتر مریم جعفرقلیان

آمنه قربانی، طیبه زنده‌دل

چاپ و نشر

انتشارات ونداد

صفحه‌آرایی و طراحی جلد

کریم علیزاده

شمارگان

۱۰۰۰

قیمت:

۳۴۰۰ تومان

شابک:

۹۷۸-۶۰۰-۷۰۱۸-۴۹-۱

این کتاب با مشارکت و تاییدیه علمی مرکز تحقیقات اختلالات متابولیک،

پژوهشکده علوم سلولی - مولکولی غدد و متابولیسم، پژوهشگاه علوم غدد و متابولیسم، دانشگاه علوم پزشکی تهران منتشر شده است.

کلیه حقوق مادی و معنوی این کتاب متعلق به انتشارات ونداد می‌باشد.

آدرس

استان مازندران، قائم شهر، مجتمع تجاری الماس، واحد ۸۳، انتشارات ونداد

تلفکس: ۰۱۱۴۲۲۷۶۶۸۸

پیشگفتار

در طب امروز مسئله پژوهش بسیار قابل توجه است. بخصوص طب مبتنی بر شواهد هر روز اهمیت بیشتری از تجربیات شخصی پیدا می کند. بر اساس تحقیقاتی که انجام می شود هر روز مطالب علمی تغییر می کند و روش های تشخیصی و درمان های جدیدی به منصه ظهور می رسد. بنابراین کسانی که در تولید علم نقش دارند مثل اساتید دانشگاهها، دانشجویان، دارنده های دکترای علوم پزشکی و علوم پایه باید به اصول و مبانی پژوهش آگاه باشند. انواع مطالعات پژوهشی در حوزه سلامت را بشناسند. درست است که متخصص آمار در هر تحقیق رل مهمی را بعهده دارد ولی پژوهشگر هم باید آلفبای آمار آشنا باشد و بتواند قسمت های ابتدائی آن را انجام داده و مفهوم آن ها را بشناسد. چگونه داده های صحیح را جمع آوری کند، چطور از نرم افزارهای مختصر استفاده شود، حجم نمونه را چگونه تعیین کند. حال که تحقیق انجام شد چگونه آن را منتشر کند. استفاده از اصول مرجع نگاری الکترونیک کار نویسندگان را بسیار سهل تر خواهد کرد. بر این اساس شش کتاب در مورد آشنائی با مبانی تحقیق به رشته تحریر درآمده است و نویسندگان آن نهایت کوشش خود را برای روشن کردن آسان و عملی مطالب به این داده اند. ضمن تشکر از آنها مطالعه این کتاب ها را به علاقه مندان و ذینفعان توصیه می کنم.

دکتر مریم رزاقی آذر

فوق تخصص غدد درون ریز و متابولیسم کودکان، استاد دانشگاه علوم پزشکی ایران، رئیس بخش غدد بیمارستان کودکان حضرت علی اصغر (ع) دانشگاه علوم پزشکی ایران، رئیس مرکز تحقیقات اختلالات متابولیک، پژوهشکده علوم سلولی - مولکولی - غدد و متابولیسم، پژوهشگاه علوم غدد و متابولیسم، دانشگاه علوم پزشکی تهران

پیشگفتار

به نام او که شفا دهنده است

جهان در حال گذر است و سلامتی یکی از بزرگترین نعمت های الهی و نیاز همیشگی و ثابت انسان است. در هر جامعه ای، انسان سالم است که محور توسعه پایدار است. برای ارتقای سلامت، تنها وجود منابع انسانی تخصصی و تجهیزات مناسب کافی نیست و توانمندی انسانها برای مراقبت از خود است که نقش تعیین کننده و اصلی را دارد. مراقبت های بهداشتی اولیه همواره می تواند بعنوان یک استراتژی کارساز برای ارتقای سلامت جامعه باشد، مشروط بر آنکه مقتضیات گذر زمان را از قبیل گذر دموگرافیک و گذر اپیدمیولوژیک مورد توجه قرار دهد. تحقیق و پژوهش ما را به نیازها و مسائل مرتبط با سلامت در هر زمان و مکانی آگاه می سازد و ما می توانیم با برنامه ریزی مبتنی بر دانش پژوهش به ارتقای سلامت نائل شویم. برای تحقیق باید با شیوه علمی و صحیح آشنا شویم. این کتاب به منظور آشنایی همه علاقه مندان عرصه سلامت با شیوه های تحقیق و پژوهش تهیه و تدوین گردیده است. امیدواریم مورد توجه و استفاده قرار گیرد.

گروه مولفین

فهرست مطالب

نام فصول	صفحه
مقدمه	۱۰
فصل اول	۱۵
داده کاوی	۱۷
تاریخچه	۱۸
موضوع داده کاوی چیست؟	۱۸
تعاریف داده کاوی	۲۱
تفاوت داده کاوی و آنالیز آماری	۲۳
مراحل داده کاوی	۲۴
فصل دوم	۲۷
الگوریتم های تکاملی	۲۹
علت استفاده از الگوریتم های تکاملی	۳۲
انواع الگوریتم های تکاملی	۳۲
استراتژی های تکاملی	۳۳
برنامه ریزی تکاملی	۳۵

۳۷	فصل سوم
۳۹	الگوریتم ژنتیک
۳۹	ژنتیک در طبیعت
۴۰	الگوریتم ژنتیک استاندارد
۴۳	فصل چهارم
۴۵	الگوریتم بهینه سازی ازدحام ذرات
۴۶	تعاریف اولیه الگوریتم PSO
۴۸	محدودسازی فضا
۴۹	مراحل اجرای الگوریتم PSO
۵۷	فصل پنجم
۵۹	الگوریتم کلونی زنبور عسل (ABC)
۶۰	مراحل الگوریتم کلونی زنبور عسل
۶۵	فصل ششم
۶۷	درخت تصمیم
۷۰	CART
۷۱	CHAID
۷۲	QUEST

۷۲	C5.0
۷۴	الگوریتم پایه ID3
۷۷	فصل هفتم
۷۹	شبکه های عصبی
۸۱	انواع شبکه های عصبی
۷۸	فصل هشتم
۸۹	فازی
۹۰	مجموعه های فازی
۹۲	تحلیل فرآیند استنتاج فازی
۹۵	فصل نهم
۹۷	درخت های تصمیم فازی
۱۰۰	مقایسه درخت های تصمیم فازی با درخت های تصمیم سنتی
۱۰۳	منابع و مراجع

مقدمه

از سال ۱۹۵۰ به بعد، با توجه به رونق گرفتن فناوری اطلاعات و ذخیره سازی داده ها به کار رفت، حجم اطلاعات ذخیره شده در آن پس از حدود ۲۰ سال دو برابر شد و همزمان با پیشرفت فناوری اطلاعات، حجم داده ها در پایگاه داده ها هر دو سال یک بار، در برابر خود و همچنین با سرعت بیش تری نسبت به گذشته حجم اطلاعات ذخیره شده بیش تر و بیش تر می شود. با وجود شبکه جهانی وب، سیستم های یکپارچه اطلاعاتی، سیستم های یکپارچه بانکی، تجارت الکترونیکی و ... اعظم به لحظه به حجم داده ها در پایگاه داده ها اضافه شده و باعث به وجود آمدن انبارهای (توده های) عظیمی از داده ها شده است، به طوری که ضرورت کشف و استخراج سریع و دقیق دانش از این پایگاه داده ها را بیش از پیش نمایان کرده است.

شدت رقابت ها در عرصه های علمی، اجتماعی، اقتصادی، سیاسی و نظامی نیز اهمیت سرعت یا زمان دسترسی به اطلاعات را دو چندان کرده است. بنا براین نیاز به طراحی سیستم هایی که قادر به اکتشاف سریع اطلاعات مورد علاقه کاربران با تأکید بر حداقل مداخله انسانی باشند از یک سو و روی آوردن به روش های تحلیل متناسب با حجم

داده های حجیم ازسوی دیگر ، به خوبی احساس می شود . در حال حاضر ، داده کاوی مهم ترین فناوری برای بهره وری موثر ، صحیح و سریع ازداده های حجیم است و اهمیت آن رو به فزونی است.

تشخیص پزشکی به عنوان یک وظیفه مهم و در عین حال پیچیده، که باید به دقت و به طور موثر اجرا شود، در نظر گرفته شده است. اتوماسیون این سیستم میتواند بسیار سودمند باشد. متأسفانه تمام پزشکان تجربه ای در در تخصص فرعی را ندارند و علاوه بر این کمبود منابع انسانی در مکن های خاص وجود دارد. بنابراین، سیستم خودکار تشخیص پزشکی احتمالاً با زراعت کردن همه این موارد بسیار سودمند خواهد بود.

تاریخچه داده های پزشکی تشکیل شده است از تعدادی از تستهای لازم برای تشخیص یک بیماری خاص. پایگاه داده های بالینی، عناصر حوزه ای هستند که در آن رویه داده کاوی به دلیل تمایل تدریجی داده های پژوهشی پزشکی و بالینی به سمت و سوی اجتناب ناپذیری توسعه یافته است. برای صنایع بهداشت و درمان انسان، بهره برداری از مزیت های داده کاوی با استفاده از آن به عنوان یک ابزار تشخیصی هوشمند وجود دارد. برای به دست آوردن دانش و اطلاعات در مورد یک بیماری ممکن است از اندازه گیری های خاص ذخیره شده ی بیمار نا داده های پزشکی، به هم مرتبط شوند. بنابراین، داده کاوی در حوزه حیاتی سلامت بهداشتی توسعه یافته است. با ساختن برنامه های کاربردی داده کاوی امکان پیش بینی تأثیر درمان های پزشکی وجود دارد.

داده کاوی با مقایسه و ارزیابی علل، علائم و دوره درمانها میتواند یک تخمین از تأثیر دوره های انجام شده ارائه دهد. کار بر روی پایگاه داده های بیماران نوعی از برنامه های کاربردی در حوزه زندگی واقعی است. تشخیص بیماری از علائم و عوامل مختلف، یک مشکل چند لایه ای است و ممکن است به فرضیات نادرست که اغلب با عوارض غیر قابل پیش بینی همراه است منجر شود. بنابراین سعی در استفاده از دانش و تجربه متخصصان مختلف که در پایگاه داده ها جمع آوری شده جهت کمک به روند تشخیص، حقول به نظر می رسد.

محققان در زمینه پزشکی، با کمک تکنیک های داده کاوی علاوه بر پیشنهاد مراقبت های موثر برای بیماران، بیماری را پیش بینی می کنند و تشخیص می دهند. تکنیک های داده کاوی، برای تشخیص بیماری های مختلف از جمله: دیابت، پاتییت، سرطان، بیماری های قلبی و مانند آن، توسط طیف گسترده ای از عملکردها در تحقیقات مورد استفاده قرار گرفته اند. اطلاعات مرتبط با بیماری، در فرم های الکترونیکی پرونده های بالینی، اطلاعات مربوط به درمان، حالات ژن ها، تصاویر و غیره در این عملکردها به کار گرفته شده است.

داده کاوی فرآیندی برای کشف روابط معنی دار، الگوها و گرایشات به وسیله بررسی حجم زیادی از داده های ذخیره شده در پایگاه های داده می باشد که این کار با استفاده از تکنولوژی شناخت الگوها انجام می شود. این تکنولوژی ها بر پایه تکنیک های آماری و ریاضی می باشند. داده کاوی یک فیلد میان رشته ای است که تکنیک های مختلف یادگیری ماشین، تشخیص الگو، آمار و پایگاه داده را به منظور اکتشاف در

پایگاه های داده ای حجیم، با هم درگیر می سازد. در داده کاوی، داده ها مورد کاوش قرار گرفته، دانش نهفته در آنها استخراج شده و الگو سازی صورت می گیرد. برای این منظور از روش های مختلف داده کاوی مانند الگوریتم های تکاملی، درخت تصمیم، شبکه عصبی و... استفاده می شود.

www.ketab.ir