

ارائه مدلی بر مبنای درخت تصمیم فازی جهت تشخیص مزاج در طب سنتی ایرانی

رؤیا رضایی اینانلو^{الف}، فرهاد سلیمانیان قره‌چپق^{الف*}، مرتضی مجاهدی^ب

^{الف} گروه مهندسی کامپیوتر، واحد ارومیه، دانشگاه آزاد اسلامی، ارومیه، ایران

^ب گروه طب سنتی، دانشکده طب سنتی، دانشگاه علوم پزشکی بابل، بابل، ایران

چکیده

سابقه و هدف: مفهوم مزاج از جمله کلیدی‌ترین مفاهیم مورد استفاده در طب سنتی ایرانی بوده و بسیاری از دستورات حفظ سلامتی، تشخیص و درمان بیماری‌ها بر اساس مزاج تعیین شده هر فرد با دیگران متفاوت است. دستیابی به شیوه‌های استاندارد تعیین مزاج از مهم‌ترین اولویت‌های پژوهش در طب سنتی ایرانی است. درخت تصمیم فازی یکی از الگوریتم‌های هوشمند برای طبقه‌بندی داده‌ها است. استنتاج درخت تصمیم فازی، فرآیند استفاده از نمونه‌های خاص و رسیدن به یک مدل عمومی است که هدف آن یادگیری نحوه طبقه‌بندی نمونه‌ها است.

مواد و روش‌ها: در این مقاله از داده‌های مربوط به تعیین مزاج افراد برای داده‌کاوی و مدل‌سازی تشخیص مزاج استفاده شد. برای این کار، درخت تصمیم فازی با الگوریتم ID3 فازی آموزش داده شد.

یافته‌ها: دو درخت تصمیم‌گیری یکی برای گرمی، سردی و دیگری برای تری، خشکی ساخته شد و قواعد تولید شده توسط پزشکان متخصص طب سنتی ایرانی از نظر بالینی و میزان صحت قوانین ارزیابی شد.

نتیجه‌گیری: باتوجه به اینکه، این مطالعه جزو اولین پژوهش‌های صورت گرفته در حوزه هوشمندسازی تشخیص مزاج در طب سنتی ایرانی است، نتایج این پژوهش می‌تواند در سطح عملیاتی توسط متخصصان طب سنتی ایرانی استفاده شود، همچنین این پژوهش می‌تواند زمینه پژوهش‌های دیگر با استفاده از بهینه‌سازی و سایر الگوریتم‌های هوشمند و حتی برای تشخیص انواع سوءمزاج با داده‌های بیشتر باشد.

کلیدواژه‌ها: طبقه‌بندی مزاج، طب سنتی ایرانی، هوش مصنوعی، درخت تصمیم فازی، یادگیری ماشین.

تاریخ دریافت: شهریور ۹۷

تاریخ پذیرش: آبان ۹۷

مقدمه:

(گرم، سرد، تر و خشک) و چهار مزاج مرکب (گرم و تر، گرم و

خشک، سرد و تر، سرد و خشک) هستند.

باتوجه به اینکه مزاج هر فرد، بسیاری از خصوصیات جسمی، روحی و فیزیولوژیک او را تحت تأثیر قرار می‌دهد، جهت تعیین مزاج هر فرد و تعیین جایگاه او در یکی از گروه‌های نه‌گانه مزاجی باید همه خصوصیات مذکور بررسی شود تا بر اساس برآیند آن‌ها، مزاج فرد به دست آید. منابع مکتوب طب سنتی ایرانی درباره شاخص‌هایی که بر اساس آن‌ها طبیب باید به تعیین مزاج پایه افراد بپردازد، توافق نظر بالایی دارند، اما اولویت یا ارجحیت خاصی برای هیچ‌کدام از شاخص‌ها قائل نشده‌اند (۱). در بین متخصصین طب سنتی ایرانی نیز

در طب سنتی ایرانی، بسیاری از دستورهای بهداشتی مربوط به سبک زندگی (تدابیر حفظ سلامتی) بر مبنای مزاج افراد تنظیم شده است. محوریت مزاج تقریباً همه حوزه‌های تشخیص و درمان این مکتب را متأثر می‌سازد. از منظر طبای طب سنتی ایرانی تعداد مزاج انسان‌ها به تعداد انسان‌های ساکن بر روی کره زمین است و در اصطلاح، مزاج هر فرد منحصر به خود اوست. بر اساس تقسیم‌بندی منابع طب سنتی ایرانی هر فرد در حین داشتن مزاج منحصر به خود در یکی از ۹ گروه اصلی مزاج جای دارد. از بین این ۹ گروه، یک گروه معتدل و هشت گروه دیگر غیر معتدل شامل چهار مزاج ساده یا مفرد

توافق بالایی بابت تعیین مزاج افراد وجود ندارد؛ لذا دستیابی به ابزارهای استاندارد که به وسیله آن بتوان با اطمینان بالایی به تعیین مزاج افراد اقدام نمود از اولویت‌های پژوهشی در طب سنتی ایرانی است (۲).

به نظر می‌رسد علم طب سنتی ایرانی برای حرکت در راستای کاربردی شدن و توسعه هر چه بیشتر، نیازمند روش‌های هوشمند و سیستم‌هایی با دقت و سرعت بالا می‌باشد. با توجه به تأکید اساتید و پژوهشگران طب سنتی ایرانی به استانداردسازی شاخص‌های تشخیصی با استفاده از آخرین دستاوردهای علمی جدید به نظر می‌رسد وجود یک سیستم واحد برای تشخیص مزاج می‌تواند موجب استانداردسازی و قابلیت اطمینان بالا شده و همچنین موجب صرفه‌جویی در زمان و هزینه مراجعه‌کنندگان و پزشکان شود و از طرف دیگر تعیین توافق میان مزاج تشخیص داده شده توسط متخصصین طب سنتی ایرانی، با استفاده از سیستم‌های خبره می‌تواند در ارزیابی بهتر و درمان بر اساس مزاج تشخیص داده شده، راهگشا باشد.

روش‌های هوش مصنوعی طیف الگوریتم‌هایی است که در زمینه‌های طبقه‌بندی، خوشه‌بندی، پیش‌بینی و بهینه‌سازی و تصمیم‌گیری عمل می‌کنند. درخت تصمیم یکی از روش‌های طبقه‌بندی است. رویکرد درخت تصمیم در بسیاری از زمینه‌ها کاربرد دارد از جمله: شناسایی الگوها، طبقه‌بندی الگوها، کلاسه‌بندی، سیستم‌های پشتیبان تصمیم‌گیری، سیستم‌های خبره و غیره؛ زیرا در مقایسه با سایر روش‌ها سریع‌ترین می‌باشد، به‌ویژه در شرایطی که فضای نمونه بزرگ است؛ علاوه بر این مهیا کردن داده آسان و درک آن نیز برای افراد غیر فنی آسان‌تر است. مزیت دیگر آن، این است که توانایی طبقه‌بندی دو نوع داده عددی و رتبه‌ای را دارد. درخت تصمیم به‌طور موفقیت‌آمیزی در تشخیص و انتخاب درمان مناسب در پزشکی نوین به‌کار می‌رود. مزیت اصلی رویکرد درخت تصمیم نشان دادن راه حل‌هاست. درخت تصمیم فازی که حاصل ترکیب الگوریتم ID3 و تئوری فازی می‌باشد، مدلی سیستماتیک ارائه می‌کند که سامانه‌های زیستی با استفاده از آن می‌توانند در

برابر تغییرات محیطی واکنش سریع و مناسب داشته باشند. مزیت دیگر این رویکرد، توانایی کار با متغیرهای کلامی است که دانش حاصل از آن برای انسان از قابلیت درک بالایی برخوردار می‌باشد. این رویکرد به دلیل فازی بودن توانایی برخورد با عدم قطعیت را داشته و با در نظر گرفتن حالت‌های مختلف در پایگاه داده، واکنش مناسب در برخورد با تغییرات محیطی دارد (۳).

یکی از عمومی‌ترین انواع سیستم‌های خبره امروزی، سیستم‌های مبتنی بر قاعده هستند. بر اساس نظرات Newell و Simon حل بیشتر مسائل توسط انسان و یا ادراک و افهام آن را می‌توان با قواعد تولیدی از نوع «اگر ... آنگاه ...» بیان کرد. با توجه به اینکه الگوریتم‌های هوشمند -که موتور استنتاج سیستم‌های خبره جدید را تشکیل می‌دهند- الهام گرفته از پردازشگر ادراکی انسان هستند، نتایج موتور استنتاج سیستم‌های خبره جدید با نتایج پردازشگر ادراکی انسان مشابهت دارند. بر خلاف دانش مربوط به تکنیک‌های حل مسائل عمومی، دانش یک فرد خبره حوزه‌مند است؛ یعنی محدود به یک دامنه خاص است. دامنه یک مسئله، نشان‌دهنده حوزه خاصی همچون حوزه پزشکی، مالی، مهندسی یا غیره است که یک فرد خبره می‌تواند مسائل آن را به‌خوبی حل کند. سیستم‌های خبره طوری طراحی شده‌اند که مثل افراد خبره در یک حوزه خاص مهارت داشته باشند.

استدلال یا استنتاج در سیستم‌های خبره از اهمیت خاصی برخوردار است؛ زیرا استدلال یک تکنیک عمومی است که سیستم‌های خبره به کمک آن مسائل را حل می‌کنند (۴). همان‌طور که از نام درختان تصمیم مشخص است، برای پشتیبانی از تصمیم به‌کار می‌روند. این ساختارها شبیه درخت بوده و اطلاعات مربوط به داده‌ها را سطح به سطح در درخت نگاشت می‌کند تا اینکه به نتیجه نهایی برسد (۵). درخت تصمیم به دلیل سادگی و قابل فهم بودن تکنیک محبوبي در داده کاوی محسوب می‌شود. به عبارت دیگر درخت تصمیم خود به تنهایی همه مطالب را توصیف می‌کند و نیاز به فرد خبره‌ای نیست تا خروجی را تفسیر کند. در واقع این یک روش

نمونه نرم افزارهای تشخیص و درمان طراحی شده بر اساس سیستم خبره پزشکی عبارتند از برنامه DENDRAL (شرح و بیان ساختار مولکولی و تحلیل شیمیایی)، نرم افزار MYCIN (تشخیص و درمان بیماری میکروبی)، نرم افزار PUFF (تشخیص بیماری‌های ریوی)، نرم افزار XBONE (تشخیص بیماری‌های استخوانی)، نرم افزار VM (نظارت بر بیماران نیازمند مراقبت‌های ویژه)، ATTENDING (آموزش در امور مدیریت بیهوشی)، CADUCEUS (تشخیص بیماری‌های مربوط به طب داخلی).

پیشینه تحقیق

اکنون جامعه پزشکی جهان به اهمیت روش‌های درمان طبیعی و مکاتب طب سنتی و مکمل پی برده است. گسترش فوق‌العاده طب سنتی چین، طب تبتی، کایروپراکتیک، آیورودا، هومیوپاتی و صدها مکتب و روش گوناگون، نشان‌دهنده این حرکت عظیم جهانی است (۹). در طب سنتی چین سال‌هاست که پژوهش‌های متعددی در زمینه هوشمندسازی تشخیص و درمان و همچنین سیستم‌های خبره تشخیص انجام شده است. از جمله پژوهش‌های انجام شده می‌توان موارد زیر را نام برد:

سیستم خبره برای پشتیبانی از تشخیص و درمان در طب سنتی چینی (۱۱)، سیستم تصمیم‌یار طب گیاهی مبتنی بر وب برای تجویز دارو (۱۰)، مقاله مروری بررسی پیشرفت‌های طب چینی در طبقه‌بندی بیماران با استفاده از یادگیری ماشین (۱۲)، سیستم خبره خودیادگیر برای تشخیص در طب سنتی چینی (۱۳)، همچنان که از عنوان این پژوهش‌ها نیز مشخص است، در زمینه طبقه‌بندی داده‌ها با استفاده از الگوریتم‌های هوشمند است. داده‌کاوی در دنیای واقعی در تشخیص و درمان بالینی طب سنتی چینی نیز از جمله پژوهش‌های مرتبط در زمینه داده‌کاوی در طب بالینی است (۱۴). پژوهش دیگری با عنوان مدل‌سازی دانش و کسب فرمول داروهای سنتی گیاهی با استفاده از متن (۱۵) و متن‌کاوی است که شاخه‌ای از داده‌کاوی محسوب می‌شود و به مدل‌سازی می‌پردازد. پژوهش کاربرد تئوری فازی در سیستم تشخیص با استفاده از معاینه زبان در طب سنتی چینی (۱۶) نیز استفاده از منطق فازی در

گرافیکی است و به این دلیل تفسیر آن شاید ساده‌تر از تکنیک‌های دیگر طبقه‌بندی باشد.

سیستم‌های خبره دارای چندین خصوصیت قابل توجه هستند (۴):

تخصص چندگانه: سطح دانش و تخصصی که از ترکیب دانش چند فرد خبره به دست می‌آید از سطح دانش یک فرد خبره بیشتر است.

افزایش قابلیت اطمینان: سیستم‌های خبره، از این بابت که دیدگاه دومی را برای خبره بشری فراهم می‌کنند و یا به موقع بروز اختلاف نظر میان افراد خبره، دیدگاه دومی را مطرح می‌کنند موجب اطمینان خاطر کاربر می‌شوند.

کاهش خطر: سیستم‌های خبره را می‌توان در محیط‌هایی که حضور در آن‌ها برای انسان خطرناک است، به کار برد. افزایش قابلیت دسترسی، دوام و بقای سیستم (که حاوی تجربه و دانش بسیاری از متخصصان است)، کاهش هزینه، پاسخ سریع (تشخیص سریع)، پاسخ کامل، ثابت و غیرحساس در همه مواقع (خستگی و حواس‌پرتی و شرایط خاص در سیستم وجود ندارد)، قابلیت توضیح استدلال و بانک اطلاعاتی هوشمند از سایر ویژگی‌های سیستم‌های خبره است.

منطق فازی برای اولین بار در جهان توسط پروفیسور لطفی‌زاده معرفی شد. همان‌گونه که منطق کلاسیک پایه‌های سیستم‌های خبره معمولی را تشکیل می‌دهد، منطق فازی نیز مبنای سیستم‌های خبره فازی است. سیستم‌های خبره فازی علاوه بر پرداختن به عدم قطعیت، قادرند استدلال عقل سلیم یا استدلال عامه‌فهم را مدل‌بندی کنند که انجام آن برای سیستم‌های معمولی بسیار دشوار است (۶، ۷). در اغلب تجزیه و تحلیل‌ها فرض می‌شود که ما با معیارهای دقیق سروکار داریم و این دقت از طریق بیان پدیده‌ها به صورت مقادیر عددی یا طبقه‌بندی شده به دست می‌آید. اگرچه استفاده از مقادیر تثبیت شده کاربردهای زیادی دارند اما طبیعت مفاهیم بشری را منعکس نمی‌کنند، زیرا این مفاهیم به انتزاعی و غیر دقیق بودن تمایل بیشتری دارند (۸).

روش بررسی

جمع‌آوری داده‌ها: باتوجه به اینکه در حال حاضر

داده‌های زیادی در طب سنتی ایرانی وجود ندارد و همچنین برای اینکه داده‌های قابل اطمینانی داشته باشیم از مجموعه داده‌های مجاهدی و همکاران (۱۸) استفاده کردیم. این مجموعه داده‌ها شامل ۵۲ رکورد از ویژگی‌های افرادی است که تشخیص مزاج شده‌اند. در واقع این مجموعه داده‌ها برای آموزش درخت استفاده می‌شوند تا قواعد تشخیص مزاج تولید شوند. این ویژگی‌ها دو دسته هستند، ۳۱ ویژگی ابتدایی متریک است که به سنجش دمای محیط و قلب و کبد و پیشانی و غیره می‌پردازد، مانند: جنسیت، دور سر، طول انگشت سبابه، سن، دور کمر، پهنای شانه، دمای محیط، قطر قدامی خلفی سر، پهنای قفسه سینه، دمای سبابه راست، پهنای سر، پهنای کف پا، دمای کف دست، شماره کفش، پهنای قفسه سینه، دمای مچ، ارتفاع سر، عمق سینه در محل نوک پستان، دمای کبد، پهنای کف دست، عمق قفسه سینه در محل چین زیر پستان، دمای قلب، طول کف دست، طول کف پا، دمای پیشانی، پهنای مچ دست، پهنای قوزک پا، وزن، پهنای مفصل سبابه، پهنای باریک‌ترین قسمت ساق پا، قد، فاصله نوک انگشتان دو دست. بقیه ویژگی‌های داده‌ها شامل ۳۹ سوال طرح شده از افراد هستند که هر سوال دارای ۵ گزینه است. این ویژگی‌ها در جدول ۱ نشان داده شده‌اند.

مدل‌سازی داده‌ها: از اهداف اصلی این کار استخراج دانش

و الگوها برای حل یک مسئله خاص است. یکی از معمول‌ترین موارد استفاده داده‌کاوی، استخراج قوانین از مجموعه داده‌ها است. در این مقاله برای مدل‌سازی داده از درخت تصمیم‌فازی استفاده شده است که با استفاده از الگوریتم ID3 آموزش می‌بیند تا قوانین مورد نظر را تولید کند.

معرفی ویژگی‌ها، کلاس‌ها و متغیرهای زبانی مجموعه داده‌های طب سنتی ایرانی بدین صورت تشریح می‌شود. مجموعه جهانی $U=\{u\}$ به وسیله مجموعه ویژگی‌ها $A=\{A_1, A_2, \dots, A_k\}$ ، تعریف می‌شود (۱۹). در مجموعه داده‌های طب

تشخیص معاینه زبان را نشان می‌دهد. این پژوهش‌ها تنها نمونه‌هایی از حجم انبوهی از پژوهش‌های صورت گرفته در این زمینه است که از گذشته تاکنون هنوز هم ادامه دارد.

در طب ایرانی، بسیاری از تشخیص‌های مزاج و سوءمزاج که از ابتدایی‌ترین نیازهای پژوهش و آموزش در این رشته می‌باشند، با استفاده از منابع طب سنتی ایرانی که به صورت کیفی بیان شده است، انجام می‌گیرد و متخصصین با معیارهای بسیار زیاد بدون معیار و شاخص‌های استاندارد شده به تشخیص و درمان می‌پردازند و این مورد موجب ایجاد تناقض بین بیماران در مراجعات مکرر به این مجموعه شده و فعلاً تنها ابزار استاندارد موجود پرسشنامه تعیین مزاج می‌باشد. علاوه بر مشکل ذکر شده به نظر می‌رسد، هوشمندسازی تشخیص مزاج و سوءمزاج می‌تواند موجب بالا رفتن سرعت و اطمینان، پرهیز از سؤالات و کارهای تکراری و صرفه‌جویی زمان پزشکان شود. پژوهش انجام شده برای پرسشنامه خوداظهاری تعیین مزاج با عنوان «بررسی پایایی و روایی پرسشنامه تعیین مزاج» شامل ده سؤال می‌باشد که جهت تعیین وضعیت مزاج از نظر گرمی-سردی و تری-خشکی به کار می‌رود. این پرسشنامه توسط مجاهدی و همکاران طراحی و روایی و پایایی آن تحت بررسی قرار گرفته است (۱۷). در این مقاله استخراج گویه‌های مرتبط با تعیین مزاج با استفاده از منابع طب سنتی ایرانی با محوریت کتاب قانون گزارش شده است. تعداد گویه‌های اولیه این پرسشنامه ۵۲ گویه بود. در مرحله بعدی روایی صوری و محتوای پرسشنامه انجام شده که با روش‌های کمی و کیفی صورت گرفته است و نهایتاً ۴۷ گویه باقی ماند. پرسشنامه به دست آمده، توسط داوطلبانی پر شده بود که قبلاً توسط پزشکان متخصص به گروه‌های (گرم/سرد) و (مرطوب/خشک) گروه‌بندی شده بودند. با استفاده از تجزیه و تحلیل رگرسیون لجستیک، نهایتاً به ۱۰ سوال منجر شده که با پاسخ دادن به آن‌ها مزاج شخص درست تشخیص داده می‌شود. ضرایب کاپای وزنی و آلفای کرونباخ برای ارزیابی و گزینش سؤالات استفاده شده است.

مواد و روش‌ها:

سنتی ایرانی مجموعه A شامل ۷۰ ویژگی است، که باتوجه به اینکه سن تمام نمونه‌ها تقریباً در یک رده سنی قرار دارد از ویژگی سن و همچنین با مرتب کردن داده‌ها بر اساس جنسیت نیز از ویژگی جنسیت صرف‌نظر شد. به این ترتیب، ۶۸ ویژگی

جدول ۱. سؤالات موجود در مجموعه داده‌ها (ویژگی‌های غیر عددی)

D1: وقتی اطرافیان، دست شما را لمس می‌کنند در مورد گرمی و سردی آن چه می‌گویند؟	D21: وضعیت زیر و بم بودن صدای شما چگونه است؟
D2: وضعیت تری و خشکی پوست شما چگونه است؟	D22: بلندی صدای شما نسبت به اطرافیان چگونه است؟
D3: وضعیت چاقی و لاغری شما نسبت به سایرین چگونه است؟	D23: هنگام صحبت کردن کلمات یک جمله را چگونه آدا می‌کنید؟
D4: نسبت عضلات و چربی در بدن شما چگونه است؟	D24: هنگام صحبت کردن چند جمله متوالی را چگونه آدا می‌کنید؟
D5: سرعت رشد موی شما نسبت به همسالانتان چگونه است؟	D25: سرعت خشم و عصبانیت شما چگونه است؟
D6: وضعیت موی سر شما از نظر پرپشتی یا کم‌پشتی چگونه است؟	D26: سرعت حرکات جسمی شما نسبت به اطرافیان چگونه است؟
D7: ضخامت موی سر شما نسبت به سایرین در چه حدی است؟	D27: سرعت پلک زدن شما نسبت به اطرافیان چگونه است؟
D8: حالت موی سر شما نسبت به سایرین چگونه است؟	D28: رنگ ادرار صبحگاهی در حالات عادی به کدامیک از رنگ‌های زیر نزدیک‌تر است؟
D9: رنگ موی سر شما در محدوده کدام گزینه است؟	D29: میزان بوی ادرار شما در حالات عادی چگونه است؟
D10: رنگ پوست بدن شما در محدوده کدامیک از گزینه‌ها می‌باشد؟	D30: رنگ مدفوع شما در حالات عادی به کدامیک از رنگ‌های زیر نزدیک‌تر است؟
D11: وضعیت کف دست شما در چه حد می‌باشد؟	D31: میزان بوی مدفوع شما در حالات عادی چگونه است؟
D12: وضعیت کوچکی و بزرگی کف پای شما در چه حد می‌باشد؟	D32: رنگ عرق بدن شما که بر روی لباس روشن به‌جای می‌ماند به کدامیک از رنگ‌های زیر نزدیک‌تر است؟
D13: وضعیت پهنای کف پای شما در چه حد می‌باشد؟	D33: قدرت اراده خود را در فعالیت‌های روزمره نسبت به اطرافیان چگونه می‌دانید؟
D14: وضعیت غضروف حنجره (برجستگی سفت جلوی گردن) شما نسبت به بافت نرم اطراف آن چگونه می‌باشد؟	D34: هنگام مطالعه، سرعت درک مطلب خود را نسبت به اطرافیان چگونه می‌دانید؟
D15: وضعیت شکل بینی شما چگونه می‌باشد؟	D35: شجاعت خود را در مقایسه با دیگران چگونه می‌دانید؟
D16: سرعت تأثیرپذیری شما از سرما و گرما چگونه است؟	D36: میزان بدبینی و خوش‌بینی خود را در مقایسه با دیگران چگونه می‌دانید؟
D17: سرعت تأثیرپذیری شما از غذاهای با طبع گرم مانند عسل، ادویه‌جات، فلفل یا غذاهای با طبع سرد مانند دوغ، ماست و خیار چگونه است؟	D37: میزان امیدواری خود را نسبت به آینده چگونه می‌دانید؟
D18: اگر محدودیت زمانی جهت خواب نداشته باشید به‌طور متوسط به چند ساعت خواب در شبانه‌روز نیاز دارید؟	D38: میزان نشاط روحی خود را نسبت به دیگران چگونه می‌دانید؟
D19: در نوجوانی سرعت رشد جسمی شما نسبت به همسالانتان چگونه بوده است؟	D39: میزان رقت قلب (دل‌نازکی) و قساوت (سنگدلی) خود را در مقایسه با دیگران چگونه می‌دانید؟
D20: قوت صدای شما نسبت به اطرافیان چگونه است؟	

- برای هر ویژگی باقیمانده که در مسیر برای بخش‌بندی استفاده نشده است، بهره اطلاعات فازی محاسبه می‌شود.
- گروه فعلی با استفاده از ویژگی که بهره اطلاعاتی را بیشینه می‌کند بسط داده می‌شود.

این درخت شبیه به درخت ID3 معمولی است با این تفاوت که نمونه‌های آموزشی می‌توانند در چندین گروه با ضرایب عضویت مختلف وجود داشته باشند.

در این الگوریتم اطلاعات زیر را داریم:

- $E = \{1, 2, \dots, N\}$ مجموعه آموزشی،

$$A = \{A_1, A_2, \dots, A_m\} \text{ مجموعه}$$

ویژگی‌ها، $(A_i) = \{A_{i1}, A_{i2}, \dots, A_{im_i}\}$ دامنه ویژگی

A_i و $(C) = \{c_1, c_2, \dots, c_m\}$ مجموعه کلاس‌ها است.

مرحله اول: برای گروه فعلی D:

- (a) با داشتن معیار $\delta (0 < \delta < 1)$ برای برگ‌ها، اگر $k' \in \{1, 2, \dots, n\}$ به‌طوری که $f_{k'}(D) \geq \delta$ ، در نتیجه D برگ است.

- (b) اگر والد گروه D در تمام شاخه‌بندی‌ها استفاده شده باشد در نتیجه D گروه برگ است و

$$f_k = \frac{M^*(D \cap c_k)}{M^*(D)}$$

مرحله دوم: اگر D برگ نیست، آنتروپی فازی را محاسبه

کن. $FE(D, A_i)$ که در آن A_i ویژگی‌هایی هستند که در فرایند شاخه‌شاخه شدن، در گروه والد استفاده نشده‌اند و سپس گروه کوچک‌ترین آنتروپی را برای گسترش استفاده می‌کنیم.

$$m_i = \sum_{j=1}^{k_j} m_{ij} \quad , \quad m_{ij} = M^*(D \cap A_{ij}) \quad \text{اگر}$$

$$m_{ij} = \sum_{k=1}^n m_{ijk} \quad \text{و} \quad m_{ijk} = M^*(D \cap A_{ij} \cap c_k)$$

$$FE(D, A_i) = - \sum_{j=1}^{k_j} \frac{m_{ij}}{m_i} E(D \cap A_{ij}) \quad (1)$$

$$E(D \cap A_{ij}) = - \sum_{k=1}^n \frac{m_{ijk}}{\bar{m}_{ij}} \log_2 \frac{m_{ijk}}{\bar{m}_{ij}} \quad (2)$$

باقی‌مانده در ساخت درخت تصمیم دخیل داده شد.

باتوجه به اینکه ۲۹ ویژگی اول متریک هستند، نیاز به فازی‌سازی دارند. در مجموعه داده‌های طب سنتی ایرانی متغیرهای زبانی برای ویژگی‌ها، دو مقدار «کمتر از میانگین» و «بیشتر از مقدار میانگین» در نظر گرفته شد و به جای ویژگی وزن، ویژگی «شاخص توده بدنی» یا همان BMI قرار گرفت، که مجموعه متغیرهای زبانی این ویژگی به صورت $18/5 < BMI < 25$ یعنی نرمال و مقدار BMI بزرگ‌تر از ۲۵ یعنی اضافه وزن می‌باشد. در واقع به جای تمام اعداد موجود در ستون‌ها دو مقدار «کمتر از میانگین» و «بیشتر از میانگین» و ستون وزن نیز به سه مجموعه تقسیم شد و به جای اعداد مربوط به وزن مقادیر «لاغر»، «نرمال» و «چاق» قرار می‌گیرد. به مراحل انجام این فرایند، فازی‌سازی می‌گویند.

برای مشخص کردن مرز بین دو متغیر زبانی «کمتر از میانگین» و «بیشتر از میانگین» در هر ستون از مقدار میانه اعداد استفاده شد که این مقدار با استفاده از روش آزمون و خطا مشخص شد. در هر ستون، میانه برای زن و مرد جدا شد تا دقت آموزش درخت بهبود یابد. در واقع با اینکه برای هر ستون دو جنسیت زن و مرد وجود دارد ولی پس از فازی‌سازی، فقط مقادیر «کمتر از میانگین» و «بیشتر از میانگین» باقی ماند. در واقع تفاوت میان زن و مرد در ستون‌ها اعمال شد بدون اینکه برای داده‌های زن و مرد درخت‌های جداگانه‌ای رسم شود. پس از فازی‌سازی، درخت تصمیم را با استفاده از الگوریتم ID3 فازی آموزش داده و برای پیاده‌سازی از زبان برنامه‌نویسی جاوا استفاده کردیم و قوانین مربوط به تشخیص مزاج تولید شدند.

ساخت درخت تصمیم فازی با استفاده از الگوریتم

ID3

برای ساخت درخت تصمیم فازی با استفاده از الگوریتم ID3 مراحل زیر را داریم:

- برای هر گروه محتوای اطلاعات فازی محاسبه می‌شود.

نتایج به دست آمده از ساخت درخت، قواعد قابل قبولی تولید کرده و مورد تأیید پزشک متخصص قرار گرفته است و به نظر می‌رسد پژوهش‌های بعدی در این زمینه می‌تواند راهگشای مشکلات موجود باشد. در مبحث یادگیری ماشین، هر قدر تعداد داده‌های آموزشی بیشتر باشد، یادگیری دقیق‌تر می‌شود.

مهم‌ترین محدودیت در این پژوهش، تعداد کم داده‌های آموزشی بود که موجب می‌شد الگوریتم با تنوع بسیار کمی از نمونه‌ها مواجه شده و یادگیری کاملی صورت نگیرد. مسئله دیگری که موجب شد این مشکل بیشتر نمود پیدا کند و تنوع داده‌ها را محدودتر کند، داده‌هایی بودند که تا حدود زیادی شبیه به هم بودند. این نمونه‌ها با وجود تعداد کم در حیطه سنی مشابه نیز بوده و در برخی ویژگی‌ها بسیار شبیه به هم و تقریباً در محدوده یکسان قرار داشتند.

نتیجه‌گیری:

در این مقاله داده‌های مزاج‌شناسی افراد، با استفاده از الگوریتم درخت تصمیم فازی طبقه‌بندی شده و افراد به گروه‌های گرم/سرد و تر/خشک تقسیم شدند. نتایج تولید شده در ارزیابی پزشک متخصص طب سنتی ایرانی تأیید شد که می‌تواند نقطه شروعی برای پژوهش‌های تکمیلی باشد تا به مرحله کاربردی شدن و توسعه در سطح کشور برسد.

پیشنهاد می‌شود باتوجه به نتایج مثبت به دست آمده، به دلیل اهمیت انتخاب ویژگی در مجموعه داده‌ها، در اولین پژوهش بعدی انتخاب ویژگی انجام و سپس درخت تصمیم فازی ساخته شود و در ادامه، ساخت درخت با سایر الگوریتم‌ها نیز انجام شده و نتایج به دست آمده با این مقاله مقایسه شود.

مرحله سوم: مراحل ۱ و ۲ را تا زمانی که شاخه جدیدی نباشد تکرار کن.

مرحله چهارم: درخت تولید شده را تبدیل به قوانین کن. قوانین تولید شده برای تشخیص گرمی و سردی در جدول ۲ و قوانین تولید شده برای تشخیص تری و خشکی در جدول ۳ نشان داده شده‌اند. در مرحله سوم یافته‌های حاصل از ترسیم درخت تصمیم توسط متخصص طب سنتی ایرانی کنترل شد و احتمال اینکه فعلاً این درخت تصمیم توانایی اولیه جهت تعیین مزاج افراد را دارد یا خیر بررسی شد.

یافته‌ها:

نتایج به دست آمده و قوانین تولید شده با استفاده از درخت تصمیم فازی به شرح زیر در جداول ۲ و ۳ آورده شده است. ارزش‌گذاری این قوانین از اعداد ۰ تا ۶ انجام شده است که ارزش «۰» نشان‌دهنده این است که قانون کاملاً نادرست و ارزش «۶» نشان‌دهنده این است که قانون تولید شده کاملاً صحیح است و اعداد بین صفر و شش میزان صحت قانون را نشان می‌دهد که هر چه به ۶ نزدیک‌تر می‌شود، صحیح‌تر است. مقادیر ۱ تا ۶ بر اساس درجه صحت قوانین ارزش‌گذاری شده‌اند.

بحث:

به نظر می‌رسد این مطالعه، اولین پژوهش صورت گرفته در زمینه هوشمندسازی تشخیص مزاج در طب سنتی ایرانی است. انجام این مقاله با تعداد داده‌های کم صورت گرفت و هدف این بود که میزان دقت الگوریتم و نتایج به دست آمده سنجیده شود تا مشخص شود این پژوهش‌ها می‌توانند مقدمه‌ای برای پژوهش‌های آینده در این زمینه و با قدرت بیشتر و حجم نمونه داده‌های بیشتر، برای حرکت در راستای کاربردی شدن و تبدیل به نرم‌افزارهای کاربردی شوند یا نه.

جدول ۲. قوانین تولید شده برای تشخیص گرمی و سردی با استفاده از الگوریتم ID3 فازی

ردیف	ارزش قانون تولید شده	شرح قوانین تولید شده
۱	۴	اگر شخص لاغر باشد دارای مزاج سرد است.
۲	۴	اگر شخص وزن نرمال داشته و هنگام صحبت کردن جملات متوالی را با فاصله بیان کند، دارای مزاج سرد است.
۳	۱	اگر شخص وزن نرمال داشته و هنگام صحبت کردن جملات متوالی را نه با فاصله و نه متصل بیان کند و دمای قلبش کمتر از میانگین بوده و پهنای کف پای باریک داشته باشد، دارای مزاج معتدل است.
۴	۰	اگر شخص وزن نرمال داشته و هنگام صحبت کردن جملات متوالی را نه با فاصله و نه متصل بیان کند و دمای قلبش کمتر از میانگین بوده و پهنای کف پای متوسط داشته باشد و پهنای شانه باریک باشد، دارای مزاج گرم است.
۵	۰	اگر شخص وزن نرمال داشته و هنگام صحبت کردن جملات متوالی را نه با فاصله و نه متصل بیان کند و دمای قلبش کمتر از میانگین بوده و پهنای کف پای متوسط داشته باشد و پهنای شانه پهن باشد، دارای مزاج سرد است.
۶	۱	اگر شخص وزن نرمال داشته و هنگام صحبت کردن جملات متوالی را نه با فاصله و نه متصل بیان کند و دمای قلبش کمتر از میانگین بوده و پهنای کف پای پهن داشته باشد، دارای مزاج گرم است.
۷	۵	اگر شخص وزن نرمال داشته و هنگام صحبت کردن جملات متوالی را نه با فاصله و نه متصل بیان کند و دمای قلبش بیشتر از میانگین باشد، دارای مزاج گرم است.
۸	۱	اگر شخص وزن نرمال داشته و هنگام صحبت کردن جملات متوالی را متصل بیان کند و سرعت خشم و عصبانیت کم داشته باشد، دارای مزاج سرد است.
۹	۱	اگر شخص وزن نرمال داشته و هنگام صحبت کردن جملات متوالی را نه متصل و نه با فاصله بیان کند و سرعت خشم و عصبانیت متوسط داشته باشد، دارای مزاج گرم است.
۱۰	۴	اگر شخص وزن نرمال داشته و هنگام صحبت کردن جملات متوالی را نه متصل و نه با فاصله بیان کند و سرعت خشم و عصبانیت زیاد داشته باشد، دارای مزاج گرم است.
۱۱	۲	اگر شخص اضافه وزن داشته باشد و دمای کبد کمتر از میانگین باشد، دارای مزاج معتدل است.
۱۲	۲	اگر شخص اضافه وزن داشته باشد و دمای کبد بیشتر از میانگین باشد، دارای مزاج معتدل است.

جدول ۳. تشریح قوانین تولید شده برای تشخیص تری و خشکی با استفاده از الگوریتم ID3 فازی

ردیف	ارزش قانون تولید شده	شرح قوانین تولید شده
۱	۶	اگر شخص لاغر باشد، دارای مزاج خشک است.
۲	۳	اگر شخص وزن نرمال داشته و شخص بدبینی باشد و همچنین پهنای قفسه سینه کمتر از میانگین باشد، دارای مزاج خشک است.
۳	۱	اگر شخص وزن نرمال داشته و شخص بدبینی باشد و همچنین پهنای قفسه سینه بیشتر از میانگین باشد و سرعت رشد موی کم داشته باشد، دارای مزاج تر است.
۴	۲	اگر شخص وزن نرمال داشته و شخص بدبینی باشد و همچنین پهنای قفسه سینه بیشتر از میانگین باشد و سرعت رشد موی نرمال داشته باشد، دارای مزاج تر است.
۵	۴	اگر شخص وزن نرمال داشته و شخص بدبینی باشد و همچنین پهنای قفسه سینه بیشتر از میانگین باشد و سرعت رشد موی زیاد داشته باشد، دارای مزاج خشک است.
۶	۱	اگر شخص وزن نرمال داشته و شخص نه بدبین و نه خوش‌بینی باشد و دمای قلبش کمتر از میانگین باشد، دارای مزاج تر است.
۷	۱	اگر شخص وزن نرمال داشته و شخص نه بدبین و نه خوش‌بینی باشد و دمای قلبش بیشتر از میانگین باشد، دارای مزاج تر است.
۸	۴	اگر شخص وزن نرمال داشته و شخص خوش‌بینی باشد و دارای پوست نرم بوده و از سرما تأثیرپذیر باشد، دارای مزاج تر است.
۹	۱	اگر شخص وزن نرمال داشته و شخص خوش‌بینی باشد و دارای پوست نرم بوده و از سرما و گرما تأثیرپذیری یکسان داشته باشد و دمای سرش کمتر از میانگین باشد، دارای مزاج خشک است.
۱۰	۱	اگر شخص وزن نرمال داشته و شخص خوش‌بینی باشد و دارای پوست نرم بوده و از سرما و گرما تأثیرپذیری یکسان داشته باشد و دمای سرش بیشتر از میانگین باشد، دارای مزاج خشک است.
۱۱	۱	اگر شخص وزن نرمال داشته و شخص خوش‌بینی باشد و دارای پوست نرم بوده و از گرما تأثیرپذیری داشته باشد، دارای مزاج تر است.
۱۲	۴	اگر شخص وزن نرمال داشته و شخص خوش‌بینی باشد و دارای پوست معمولی باشد، دارای مزاج معتدل است.
۱۳	۱	اگر شخص وزن نرمال داشته و شخص خوش‌بینی باشد و دارای پوست خشک باشد و طول کف پای کمتر از میانگین داشته باشد، دارای مزاج معتدل است.
۱۴	۱	اگر شخص وزن نرمال داشته و شخص خوش‌بینی باشد و دارای پوست خشک باشد و طول کف پای بیشتر از میانگین داشته باشد، دارای مزاج تر است.

References:

1. Mojahedi M, Naseri M, Majdzadeh SR, Keshavarz M, Ebadiani M, Nazem E, Khodadoost M, Mozaffarpur SA. A review on identification mizaj (Temperament) indices in Iranian traditional medicine (ITM). *Medical History Journal*. 2012;4(12):37-76.
2. Mojahedi M, Mozaffarpur SA, Isfahani MM, Naseri M, Kamalinejad M, Khodadoust M, Keshavarz M, Amin G, Abbassian AR, Najjari A, Zargarani A. Research priorities in traditional Persian medicine. *Journal of Research on History of Medicine*. 2016 May 31;5(2 May).
3. Jafari B, Azar A. Fuzzy decision tree; the new approach in strategy formulation. *General Management Researches*. 2013 April; 6(19): 25-39.
4. Ghazanfari M, Kazemi Z. *Osul va mabani systemhaye khebre*. Tehran: Amirkabir University of Technology Publication; 2003. [In Persian]
5. Kahraman C, Kaymak UU, Yazici A, editors. Fuzzy logic in its 50th year: new developments, directions and challenges. Springer; 2016 May 17.
6. Lotfizadeh A. Probability measures of fuzzy events. *Journal of Mathematical Analysis and Applications*. 1968 Aug 1;23(2):421-7.
7. Lotfizadeh A. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*. 1978 Jan 1;1(1):3-28.
8. Han J, Pei J, Kamber M, Esmaeili M. Data mining: concepts and techniques (The morgan kaufmann series in data management systems). Tehran: Niaz Danesh; 2015. [In Persian]
9. Ghaffari F, Naseri M, khodadoust M. Traditional Iranian medicine and the need for its revival and development. *Teb va Tazkiye*. 19(3): 63-71.
10. Ogirima SA. Web-based decision support system for prescription in herbal medicine. *Journal of Emerging Trends In Engineering And Applied Sciences*. 2015 May 1;6(7):245-54.
11. Silva P, Gago P, Ribeiro JCB, Santos MF, Portela F, Abelha A, *et al*. An expert system for supporting traditional Chinese medicine diagnosis and treatment. *Procedia Technology* .2014. 16 (Supplement C), 1487-1492.
12. Zhao C, Li GZ, Wang C, Niu J. Advances in patient classification for traditional Chinese medicine: a machine learning perspective. *Evidence-Based Complementary And Alternative Medicine*. 2015.
13. Wang X, Qu H, Liu P, Cheng Y. A self-learning expert system for diagnosis in traditional Chinese medicine. *Expert Systems With Applications*. 2004 May 1;26(4):557-66.
14. Zhang X, Zhou X, Zhang R, Liu B, Xie Q. Real-world clinical data mining on TCM clinical diagnosis and treatment: a survey. In 2012 IEEE 14th International Conference on E-Health Networking, Applications and Services (Healthcom) 2012 Oct 10 (P:88-93). IEEE.
15. Cao C, Wang H, Sui Y. Knowledge modeling and acquisition of traditional Chinese herbal drugs and formulae from text. *Artificial Intelligence In Medicine*. 2004 Sep 1;32(1):3-13.
16. Watsuji T, Arita S, Shinohara S, Kitade T. Medical application of fuzzy theory to the diagnostic system of tongue inspection in traditional Chinese medicine. In FUZZ-IEEE'99. 1999. IEEE International Fuzzy Systems. Conference Proceedings (Cat. No. 99CH36315) 1999 Aug 22 (Vol.1, P:145-148). IEEE.

17. Mojahedi M, Naseri M, Majdzadeh R, Keshavarz M, Ebadiani M, Nazem E, Isfeedvajani MS. Reliability and validity assessment of Mizaj questionnaire: a novel self-report scale in Iranian traditional medicine. Iranian Red Crescent Medical Journal. 2014 Mar;16(3).
18. Mojahedi M. Assessment of reliability of Mizaj (Temperament) criteria and validity of touch and anthropometric indices. Tehran: Shahed University;. 2014. P:312. [In Persian].
19. Li F, Jiang D. Fuzzy ID3 algorithm based on generating hartley measure. In International Conference on Web Information Systems and Mining 2011 Sep 24 (P:188-195). Springer, Berlin, Heidelberg.