



داده کاوی فازی و مدیریت احساسات و تفسیرپذیری

مژگان قنبرپور

استادیار دانشگاه آزاد اسلامی سروستان

چکیده

امروزه، با پیشرفت علم و تکنولوژی و ابزارهای فناوری، توانایی بازبینی و ذخیره داده‌های مهم با حجم وسیع فراهم شده و نیاز به علمی جهت جستجو در این داده‌ها و دریافت نتایج مفید، لازم و ضروری شده است. برای تسلط بر علم داده‌ها، دانستن ریاضیات و آمار ضروری است. درک صحیح مفاهیم حوزه ریاضیات و آمار در علم داده‌ها بسیار مهم است. تئوری مجموعه‌های فازی، سهم مهمی در داده‌کاوی ارائه می‌دهد که منجر به داده‌کاوی فازی می‌شود. مدیریت اطلاعات قابل تفسیر و ذهنی، هم در ورودی و هم در خروجی فرآیند داده‌کاوی، امکان‌پذیر می‌شود. در این مقاله مفهوم تفسیرپذیری و مدیریت احساسات در داده‌کاوی فازی را مورد بحث قرار می‌دهیم.

واژگان کلیدی: فازی، داده‌کاوی، تفسیرپذیری، مدیریت احساسات



مقدمه

ریشه داده‌کاوی در میان سه خانواده از علوم قابل پیگیری می‌باشد. مهمترین این خانواده ها آمار کلاسیک است. بدون آمار هیچ داده‌کاوی وجود نخواهد داشت، بطوریکه آمار، اساس اغلب تکنولوژی‌هایی می‌باشد که داده‌کاوی بر روی آنها بنا می‌شود. مطمئناً تحلیل آماری کلاسیک، نقش اساسی در تکنیکهای داده‌کاوی ایفا می‌کند. تئوری فازی توسط پرفسور لطفی عسگر زاده یا زاده معرفی شد. زاده به ناتوانی ریاضیات کلاسیک برای پرداختن به مسائل نادقیق دنیای واقعی اشاره کرد و فونداسیون چارچوب جدیدی بنام تئوری فازی را پایه ریزی و مبانی آن را معرفی کرد. همانطور که دیوید و هیکی ذکر کرده اند، به طور ساده، داده‌کاوی به معنای استخراج یا معدن کاری دانش، از مقدار زیادی داده خام است (A. David Hand And B. Heikki Mannila, ۲۰۰۱). داده‌کاوی، فرآیند یافتن دانش از مقادیر عظیم داده‌های ذخیره شده در پایگاه داده، انبار داده و یا دیگر مخازن اطلاعات است (A. J. Han And B M. Kamber, ۲۰۰۱). واژه فازی به معنای نادقیق، مبهم و گنگ است. تئوری فازی یک چارچوب جدید است که توانایی مدل کردن واقعیت را، آنچنان که هست، دارد. اساساً منطق فازی یک منطق چند منظوره است. داده‌کاوی و چندین حوزه تحقیقاتی مرتبط با آن برای القای خودکار مدل‌ها و استخراج الگوهای جالب از داده‌های تجربی، در ارتباط هستند. مشارکت مجموعه‌های فازی در داده‌کاوی متفاوت است. افزایش تفسیرپذیری، افزایش استحکام فرآیند و مدیریت اطلاعات نامشخص، به ویژه اطلاعات ذهنی و احساسی. هر دو با معرفی تئوری مجموعه‌های فازی برای ایجاد داده‌کاوی فازی ارائه می‌شوند، و به آن فرآیند ظرفیت استخراج اطلاعات پیچیده در یک محیط کلاسیک، با در نظر گرفتن حالت خاص احساسات، دشوار است. تفسیرپذیری، محور این مقاله است و در ادامه به آن پرداخته می‌شود. استحکام فرآیند، آن را قادر می‌سازد در صورت مواجهه با تغییرات جزئی در داده‌ها (مثلاً در حضور نویز) نتایج مشابهی تولید کند. استحکام سیستم‌های فازی به طور گسترده مورد مطالعه قرار گرفته است و حتی در حوزه‌های یادگیری غیر ماشینی نیز شناخته شده است. در این مقاله، ابتدا، داده‌کاوی و منطق فازی ارائه، و سپس، تفسیرپذیری و مدیریت احساسات در داده‌کاوی فازی مورد مطالعه و در پایان، نتیجه گیری ارائه شده است.

داده‌کاوی و منطق فازی

منطق فازی

منطق فازی یک مفهوم بر اساس نظریه مجموعه کلاسیک و کریسپ است. نظریه مجموعه کریسپ، یک جهان را تعریف می‌کند، که در آن، مجموعه‌ای از اشیاء شناخته شده بعنوان عناصر، در این جهان وجود دارد. اغلب این عناصر یک شباهت دارند که به آنها امکان می‌دهد تا برای سادگی یا راحتی، گروه‌بندی شوند. به عنوان مثال، در یک جهان که عناصر مشخصه آنها، شامل یک تا ده، مجموعه‌ای با عنوان "اعداد اول" وجود دارد. بنابراین، عناصر منحصر به این مجموعه، دو، سه، پنج و هفت هستند. این به عنوان یک مجموعه کلاسیک و کریسپ است، همانطور که بدون شک شناخته شده است که این چهار عدد متعلق به این مجموعه‌اند، یک سیستم منطق فازی، شامل چهار عنصر اساسی می‌باشد: فازی‌ساز، قواعد فازی، موتور استنتاج و فازی‌زدایی (غیرفازی کردن). پایگاه قواعد فازی، مجموعه‌ای از قواعدی است که در موتور استنتاج ترکیب شده است تا خروجی فازی تولید کند. ورودی‌های کریسپ را به مجموعه‌های فازی نشان می‌دهد که بعدها به عنوان ورودی‌ها به موتور استنتاج مورد استفاده قرار می‌گیرند. در حالی که فازی‌زدایی، مجموعه‌های فازی تولید شده توسط موتور استنتاج را به اعداد کریسپ تبدیل می‌کند.



داده کاوی

بطور کلی استفاده همگانی از وب و اینترنت به عنوان یک سیستم اطلاع رسانی جهانی ما را مواجه با حجم زیادی از داده و اطلاعات می‌کند. این رشد انفجاری در داده‌های ذخیره شده، نیاز مبرم به وجود تکنولوژی‌های جدید و ابزارهای خودکاری را ایجاد کرده که به صورت هوشمند به انسان یاری رسانند تا این حجم زیاد داده را به اطلاعات و دانش تبدیل کند. داده‌کاوی به عنوان یک راه حل برای این مسائل مطرح می‌باشد. در یک تعریف غیر رسمی، داده‌کاوی، فرآیندی خودکار، برای استخراج الگوهای که دانش را بازنمایی می‌کنند، می‌باشد، که این دانش به صورت ضمنی در پایگاه داده‌های عظیم، انباره داده و دیگر مخازن بزرگ اطلاعات، ذخیره شده است. داده‌کاوی بطور همزمان از چندین رشته علمی بهره می‌برد، نظیر، تکنولوژی پایگاه داده، هوش مصنوعی، یادگیری ماشین، شبکه‌های عصبی، آمار، شناسایی الگو، سیستم‌های مبتنی بر دانش، حصول دانش، بازیابی اطلاعات، محاسبات سرعت بالا و بازنمایی بصری داده. داده‌کاوی مرحله‌ای از فرایند کشف دانش می‌باشد که شامل الگوریتم‌های مخصوص است، بطوریکه تحت محدودیتهای مؤثر محاسباتی قابل قبول الگوها و یا مدلها را در داده کشف می‌کند. به بیان ساده‌تر داده‌کاوی به فرایند استخراج دانش ناشناخته درست و بالقوه مفید از داده اطلاق می‌شود. تعریف دیگر اینکه داده‌کاوی، گونه‌ای از تکنیکها برای شناسایی اطلاعات و یا دانش تصمیم‌گیری از قطعات داده می‌باشد، به نحوی که با استخراج آنها، در حوزه‌های تصمیم‌گیری، پیش‌بینی، پیشگویی و تخمین مورد استفاده قرار گیرند. داده‌ها، اغلب حجم می‌باشند که به تنهایی قابل استفاده نیستند، بلکه دانش نهفته در داده‌ها قابل استفاده می‌باشد. به این دلیل، اغلب به داده‌کاوی، تحلیل داده‌ای ثانویه گفته می‌شود (A. David Hand And B. Heikki Mannila, ۲۰۰۱, A. J. Han And B M. Kamber, ۲۰۰۱). داده‌کاوی از چند مرحله (پیش‌پردازش داده، یادگیری، تجزیه و تحلیل، انتخاب) تشکیل شده است (U.M. Fayyad et al, ۱۹۹۶). داده‌کاوی و یادگیری ماشینی دو حوزه در هم تنیده هستند. به این معنا که داده‌کاوی معمولاً شامل یک الگوریتم یادگیری ماشین به عنوان یک مرحله است. در (A. J. Han And B M. Kamber, ۲۰۰۱, L. A. Zadeh, ۱۹۸۴) تفاوت داده‌کاوی و یادگیری ماشین برجسته شده است. هدف اصلی یادگیری ماشینی توسط عملکرد در دیدگاه پیش‌بینی هدایت می‌شود، در حالی که داده‌کاوی با قابل درک بودن الگوهای کشف شده مرتبط است. هدف یادگیری ماشینی، ایجاد مدلی از مجموعه داده‌هایی است که دانش پس‌زمینه‌ای را ارائه می‌دهد. داده‌ها مجموعه آموزشی موجود را برای ایجاد مدل یا تنظیم آن تشکیل می‌دهند. در این فرآیند، مدل را میتوان به عنوان دانش جدیدی که از یادگیری تولید می‌شود، در اشکال مختلف مشاهده کرد. به عنوان مثال، تابع ریاضی، شبکه عصبی. باتوجه به اینکه الگوریتم‌های یادگیری ماشین بسیار قابل اعتماد هستند، انتخاب الگوریتمی که در داده‌کاوی مورد استفاده قرار می‌گیرد، به لطف تفسیرپذیری مدلی که می‌سازد، انجام می‌شود. داده‌کاوی فرآیندی مبتنی بر استفاده از الگوریتم یادگیری ماشین در یک کار کاربردی خاص است. یک تفاوت اصلی این است که مدل ساخته شده پس از آن نه تنها برای طبقه‌بندی موارد جدید، بلکه برای ارائه دانش اضافی نیز استفاده می‌شود. داده‌کاوی همچنین به دقت مدل به دست آمده برای ارائه اعتبار یا ارزیابی عملکرد آن مربوط می‌شود (U.M. Fayyad et al, ۱۹۹۶). در این مورد، مدل‌هایی با دقت خوب به جای بالاترین می‌توانند کافی باشند، زیرا هدف اصلی، کسب دانش قابل درک در مورد داده‌ها است.

داده کاوی فازی



داده کاوی فازی، توسعه داده کاوی است که در آن، مدلسازی مجموعه های فازی معرفی شده است. بسته به کاربرد یا مشکلی که باید با آن برخورد کرد، میتوان انواع مختلفی از داده کاوی فازی را برجسته کرد. در داده کاوی فازی، فازی میتواند در دوسطح ظاهر شود، چالش در چنین موردی، پیشنهاد الگوریتمی است که هم بتواند ورودی فازی را مدیریت کند و هم ویژگی های اصلی الگوریتم کلاسیک را برآورده کند. داده ها فازی هستند. وقتی داده ها مجموعه های فازی هستند، الگوریتم یادگیری باید آنها را مدیریت کند. این امر یا به گسترش پارامترهای رسمی الگوریتم منجر میشود (C. Marsala And B. Bouchon-Meunier, ۱۹۹۹) یا ایجاد الگوریتم های جدیدی که از نظریه مجموعه های فازی ساخته شده اند (L.-X. Wang And J. Mendel, ۱۹۹۲) تا بتواند مجموعه های فازی را مدیریت کند. به عنوان مثال، مجموعه های فازی را میتوان برای نمایش داده های عددی ارائه کرد (M. Rifqi et al, ۱۹۹۷). دانش همچنین میتواند نامطمئن باشد یا توزیع اعتقادی برای آوردن اطلاعات اضافی در مورد داده ها باشد. در این مورد، هر داده را میتوان بایک احتمال، یک باور یا هر درجه عدم قطعیت وزن کرد. معمولاً کارهای داده کاوی فازی به داده های عددی اطلاق میشود که مجموعه های فازی از آنها استنباط میشود. این مجموعه های فازی را میتوان با استفاده از یک الگوریتم یادگیری (C. Marsala, ۲۰۰۱) یا با استفاده از روش بهینه سازی که برای تنظیم پارامترهای یک پارتیشن فازی (R. Krus et al, ۱۹۹۹) استفاده میشود، ایجاد کرد. لازم به ذکر است که بر خلاف داده کاوی کلاسیک، تاکنون در داده کاوی فازی، هیچ معیار محبوبی متشکل از مجموعه داده های فازی برای مقایسه الگوریتم ها وجود ندارد. بنابراین، آنها معمولاً با استفاده از معیارهای کلاسیک، گاهی اوقات پس از فرآیند فازی سازی مصنوعی داده های عددی، مقایسه میشوند. متأسفانه، چنین مقایسه ای به ویژه برای مقایسه دقت مدل ها به جای تفسیرپذیری آنها مفید است. رویکردهایی را میتوان به منظور متعادل کردن دقت و تفسیرپذیری در مدلسازی فازی پیشنهاد کرد.

کاربرد یادگیری ماشین و آمار در داده کاوی

داده کاوی از پیشرفتهایی که در زمینه هوش مصنوعی و آمار رخ می دهد بهره می گیرد. داده کاوی جانشین تکنیک های آماری سابق نمی باشد بلکه وارث آنها بوده و در واقع تغییر و گسترش تکنیک های سابق برای متناسب سازی آنها با حجم داده ها و مسائل امروزی می باشد. تکنیک های کلاسیک برای داده های محدود و مسائل ساده مناسب بوده اند حال آنکه با پیچیده شدن مسائل و رشد روزافزون داده ها نیاز به تغییر آنها کاملاً طبیعی است. بعبارت دیگر داده کاوی ترکیب تکنیک های کلاسیک با الگوریتم های جدید می باشد. مهمترین نکته این است که داده کاوی راهکاری است برای مسائل تجاری امروز به کمک تکنیک های آماری و هوش مصنوعی برای افراد حرفه ای که قصد دارند یک مدل پیش بینی ایجاد نمایند. بسیاری از الگوریتم های یادگیری ماشین کلاسیک به موارد فازی گسترش یافته اند (S. Guillaume, ۲۰۰۱, A. Laurent, ۲۰۰۱, H. Ishibushi et al, ۲۰۰۱, I. Turksen, ۱۹۹۸). گسترش یک الگوریتم کلاسیک برای ایجاد یک الگوریتم یادگیری فازی کار جالبی است. مقالات زیادی در مورد درختان تصمیم فازی و در ساخت پایه قوانین فازی منتشر شده است (R. Weber, ۱۹۹۲, M. Ramdani, ۱۹۹۳, Y. Yuan And M. Shaw, ۱۹۹۵, T. -P. Hong And J.-B. Chen, ۲۰۰۰).

تفسیرپذیری در داده کاوی فازی

افزایش تفسیرپذیری بدست آمده با استفاده از الگوریتم داده کاوی فازی به جای الگوریتم داده کاوی (کلاسیک) برای ارزیابی، پیچیده است. در واقع، ارزیابی سطح تفسیرپذیری به مرحله ای از فرآیند داده کاوی که در آن فازی شدن رخ می دهد، بستگی دارد. از ارائه داده کاوی فازی نشان داده شده در بخش قبل، قابلیت تفسیر را میتوان در چندین مرحله از فرآیند برجسته کرد. تفسیرپذیری را میتوان با در نظر گرفتن دو نوع شخص (یا کاربران) مشا هده کرد: اپراتور و کاربر نهایی. اپراتور به ساخت فرآیند داده کاوی می پردازد و باید چندین پارامتر این فرآیند را تنظیم کند: انتخاب الگوریتم یادگیری ماشین، تهیه داده ها. کاربر نهایی نگران مدل تولید شده و استفاده از آن در

حوزه کاربردی دنیای واقعی است. این دو کاربر معمولاً می‌توانند دانش و توانایی‌های متفاوتی داشته باشند (مثلاً اپراتور معمولاً یک دانشمند رایانه است و کاربر نهایی معمولاً یک پزشک یا متخصص در حوزه برنامه است). بنابراین، تفسیرپذیری معمولاً از دیدگاه کاربر نهایی در نظر گرفته میشود. این نوع اصلی تفسیرپذیری است که در بسیاری از کاربردهای داده‌کاوی فازی جستجو میشود. همانطور که در (R.Kruse et al, ۱۹۹۹) بیان شد، یک راه حل فازی نه تنها به دلیل دقت آن، بلکه به دلیل سادگی و خوانایی آن مورد ارزیابی قرار می‌گیرد. معیارهای زیادی برای تفسیرپذیری وجود دارد. برای ارزیابی (M. Gacto et al, ۲۰۱۱) دو دسته معیار پیشنهاد می‌شود، تفسیرپذیری مبتنی بر پیچیدگی و تفسیرپذیری مبتنی بر معنایی. داده‌کاوی فازی و اطلاعات قابل تفسیر این دو کاربر معمولاً می‌توانند دانش و توانایی‌ها ی متفاوتی داشته باشند (مثلاً اپراتور معمولاً یک دانشمند رایانه است و کاربر نهایی معمولاً یک پزشک یا متخصص در حوزه برنامه است). بنابراین، تفسیرپذیری معمولاً از دیدگاه کاربر نهایی در نظر گرفته میشود. افزایش تفسیرپذیری بدست آمده با استفاده از الگوریتم داده‌کاوی فازی به جای الگوریتم داده‌کاوی (کلاسیک) برای ارزیابی پیچیده است. در واقع، ارزیابی سطح تفسیرپذیری به مرحله‌ای از فرآیند داده‌کاوی که در آن فازی شدن رخ می‌دهد، بستگی دارد. از ارائه داده‌کاوی فازی نشان داده شده در بخش قبل، قابلیت تفسیر را میتوان در چندین مرحله از فرآیند برجسته کرد. تفسیرپذیری مدل، مربوط به ساختار آن، خوانایی آن و انسجام شهودی آن است. تعداد قوانین و متغیرها در هر قانون، استفاده از مجموعه‌های فازی معنادار (R.Kruse et al, ۱۹۹۹, M. Gacto et al, ۲۰۱۱). با این حال، تفسیرپذیری ساختار مدل که باید توسط کاربر نهایی به وضوح درک شود، میتواند منجر به انتخاب یک الگوریتم خاص شود. به عنوان مثال، درختان تصمیم یا قوانین، دانش پایه در بسیاری از حوزه‌ها هستند و بنابراین برای هر نوع کاربر نهایی بسیار قابل تفسیر هستند. امروزه، بسیاری از کارها بر افزایش قابلیت تفسیر مدل‌های استنباط شده توسط داده‌کاوی فازی تمرکز دارند. در واقع، تفسیرپذیری یک مدل فازی معمولاً به لطف استفاده از اصطلاحات زبانی برای بیان مدل و روابطی که آن را برجسته میکند، انجام میشود. به عنوان مثال، برای نشان دادن این نکته، اجازه دهید یک قانون (کلاسیک) تعریف شود: «اگر قیمت بیشتر از ۱۳.۷۵ یورو باشد، کتاب با جلدسخت است». چنین قاعده‌ای را میتوان از مجموعه آموزشی متشکل از کتاب‌های موجود در فروشگاه کتاب با استفاده از الگوریتم یادگیری ماشینی برجسته کرد که همچنین میتواند به طور خودکار مرزهای قیمت‌ها را استنتاج کند. به همین ترتیب، یک قانون فازی معادل می‌تواند این باشد که «اگر قیمت بالا است، کتاب با جلدسخت است». در اینجا، عبارت "بالا" یک برجسب مرتبط با مجموعه فازی مربوطه است که از مجموعه آموزشی استنتاج میشود. تفسیرپذیری در این مورد را میتوان بهتر از قوانین قبلی در نظر گرفت: اصطلاحی که استفاده میشود از نظر معنایی برای هر شخصی مناسب است و همچنین برای هر شخصی قابل درک است حتی اگر "یورو" پول ملی او نباشد. البته، ما در اینجا از تفسیرپذیری باتوجه به کاربر نهایی صحبت کردیم. میتوان گفت که باتوجه به اپراتور، تفسیرپذیری با دانش نظریه مجموعه‌های فازی نیز مرتبط است که او را قادر می‌سازد مقادیر عددی زیربنای یک عبارت فازی را درک کند. معمولاً ابزارهای کلاسیک برای ارزیابی این تفسیرپذیری بر اساس پیچیدگی آن است: تعداد قوانین فازی که مدل از آنها تشکیل شده است، تعداد متغیرها در هر قانون، استفاده از مجموعه‌های فازی معنادار (R.Kruse et al, ۱۹۹۹, M. Gacto et al, ۲۰۱۱). با این حال، تفسیرپذیری ساختار مدل که باید توسط کاربر نهایی به وضوح درک شود، میتواند منجر به انتخاب یک الگوریتم خاص شود. به عنوان مثال، درختان تصمیم یا قوانین، دانش پایه در بسیاری از حوزه‌ها هستند و بنابراین برای هر نوع کاربر نهایی بسیار قابل تفسیر هستند. امروزه، بسیاری از کارها بر افزایش قابلیت تفسیر مدل‌های استنباط شده توسط داده‌کاوی فازی تمرکز دارند. با این حال، تفسیرپذیری نیز باید با خروجی مدل در هنگام استفاده برای داده‌های آزمایشی مرتبط باشد. نتیجه تولیدشده توسط مدل باید قابل درک باشد زیرا اغلب به کاربران نهایی غیر متخصص در علوم کامپیوتر پیشنهاد میشود (به عنوان مثال در نظریه مجموعه‌های فازی، در یادگیری ماشین، در آمار و ...). در این مورد، ارائه خروجی قابل توضیح معنایی و ارجاع به بیان کلاسیک دانش در حوزه درگیر مهم است. علاوه بر این، باید



به دانش شهودی به عنوان "عضویت در یک کلاس" اشاره کرد، که به راحتی حتی توسط کاربر ناآگاه از مدلسازی فازی قابل درک است. آخرین سطحی که قابلیت تفسیر در آن میتواند مهم باشد خود الگوریتم یادگیری است. ما در اینجا قابلیت تفسیر را با توجه به کاربر نهایی در نظر می‌گیریم زیرا اغلب برای اپراتور بسیار مهم است که به او توضیح دهد که چگونه مدل از مجموعه داده ساخته شده است.

مدیریت احساسات در داده کاوی فازی

داده کاوی فازی یکی از فعالترین زمینه‌های تحقیقاتی است که در آن بازنمایی دانش مبتنی بر مجموعه‌های فازی متمرکز است. بسیاری از حوزه‌های آن هنوز به طور گسترده مورد مطالعه قرار می‌گیرند و ما بر روی دو جنبه امیدوارکننده تمرکز خواهیم کرد، جایی که پیشرفت‌های زیادی باید انجام شود و تفسیرپذیری و مدیریت ذهنی مهم هستند. تولید احساسات مجازی و شناخت احساسات طبیعی، چالشی برای چندین سال در چارچوب نوظهور محاسبات عاطفی بوده است. مدلسازی فازی برای بیان تدریجی احساسات، که اغلب با فرآیندهای یادگیری مرتبط است، استفاده شده است. این فرآیندها یک رفتار کلی از سیستم‌هایی را ارائه می‌دهند که عوامل ذهنی مانند احساسات یا حالت‌های روان‌شناختی را با اطلاعات عینی مرتبط می‌کنند که میتواند از تصاویر دیجیتال، ویدیوها یا صداها و همچنین عناصر تعامل انسان و ماشین استخراج شود. ما بطور متوالی، تولید احساسات مجازی و شناخت احساسات طبیعی را بررسی خواهیم کرد. تولید احساسات مجازی در طراحی آواتارها، عوامل یا شخصیت‌های هوشمند برای بازی‌های ویدیویی یا کارتون در نظر گرفته میشود. این با بازتولید ویژگی‌های احساسات انسانی بر روی مصنوعات مطابقت دارد و شاخص‌های ظریف احساسات را در چهره‌ها و ژست‌های مجازی درخواست می‌کند. مدلسازی فازی کمک می‌کند تا ظرافت احساسات را به تصویر بکشد، به عنوان مثال استفاده از یک سیستم مبتنی بر قوانین فازی (A. Raouzaion et al, ۲۰۰۲, T.D. Bui et al, ۲۰۰۱) یا

شباهت‌های فازی (R. Niewiadomski And C. Pelachaud, ۲۰۰۷). در (R. Niewiadomski And C. Pelachaud, ۲۰۰۷) مدلی از احساسات برای عوامل مصنوعی که قادر به برقراری ارتباط با انسان هستند، با استفاده از قواعد استنتاج فازی برای تعیین سطوح عوامل هیجانی که باعث ایجاد احساسات با شدت‌های مختلف میشوند، پیشنهاد شده است. یک مکانیسم یادگیری برای تنوع بخشیدن به رفتارهای مرتبط با یک وضعیت عاطفی معین استفاده میشود. مدلسازی پیچیده‌تر بر اساس مدل‌های روانشناختی احساسات با هدف ارائه راهی برای بهبود تعاملات بین یک عامل هوشمند و محیط آن موجود است. مدل پیشنهادی منطق فازی تطبیقی احساسات از منطق فازی برای در نظر گرفتن شدت احساسات و انتقال آرام بین حالات احساسی با استفاده از قوانین فازی که روابط بین رویدادها، احساسات و رفتارها را ثبت میکند، استفاده میکند. این شامل یک جزء یادگیری است که ادراکات و ارتباط بین رویدادها و احساسات را مدیریت می‌کند. شناخت احساسات طبیعی در عوامل انسانی با استخراج احساسات از متون یا تصاویر و همچنین تجزیه و تحلیل فیلم‌ها یا پردازش سیگنال‌های بیولوژیکی برای تشخیص احساسات انسان با وضعیت روانی عامل انسانی سروکار دارد. چنین شناختی در تعدادی از حوزه‌های کاربردی، عمدتاً از طریق تکنیک‌های داده کاوی، دخیل است. در زمینه تعامل بین انسان و ماشین، سیستم‌های تطبیقی میتوانند در صورتی که احساسات به درستی شناسایی شوند، به احساسات کاربر خود واکنش نشان دهند. به عنوان مثال، زمانی است که گزارش‌های استفاده از وب استخراج و از طریق قوانین ارتباط فازی یا الگوهای زمانی فازی تجزیه و تحلیل میشوند تا رفتار عاطفی مصرف‌کنندگان را در نظر بگیرند. بسیاری از جنبه‌های دیگر تعامل انسان و ماشین با کمک مدلسازی فازی، به ویژه برای شناسایی حالات عاطفی یا روانی از داده‌های ارائه شده توسط انواع مختلف حسگرها، مانند حسگرهای بیولوژیکی یا دوربین‌ها، با کاربردهای حیاتی در یک محیط پزشکی یا برای کمک به افراد مسن یا ناتوان سایر کاربردهای دنیای واقعی به ارزیابی کیفیت محصول مربوط میشود. استفاده از بهره برداری ردیابی چشم، ضبط کاربر و به دنبال آن ویدئو کاوی یا استخراج در خروجی‌های حسگر به اندازه کافی با استفاده از نمایش‌های مبتنی بر مجموعه فازی مورد بررسی قرار نگرفته است، برای مثال زمانی که احساسات برای ارزیابی کیفیت ویدئو شناسایی میشوند. بازی‌ها یا کمک به طراحی آنها، حتی اگر چنین برنامه‌هایی قبلاً مورد بررسی قرار گرفته باشند. کاربرد‌های نظر کاوی، تحلیل احساسات یا شهرت الکترونیکی در هوش تجاری، روابط هوشمند با مشتریان یا پاسخگویی خودکار به ایمیل، زمینه‌ای مناسب است که در آن انعطاف‌پذیری مدل‌های فازی میتواند راه‌حلهایی را به ارمغان بیاورد که هنوز مطالعه نشده اند. تشخیص احساسات برانگیخته شده توسط تصاویر نیز موضوع



مهمی است، با کاربردهای آشکار در بازیابی تصویر، به عنوان مثال برای یافتن راه حل برای درخواستی مانند "من به دنبال یک تصویر شاد هستم"، مقابله با فقدان حاشیه نویسی احساسی واضح برای اکثر تصاویر. برخی از پیوندهای بین احساسات و رنگها با کمک بازنمایی دانش فازی مورد مطالعه قرار گرفته است. استفاده از پیوندهای بین احساسات و اشکال اشیاء شناسایی شده بر روی تصاویر نیز در طراحی صنعتی با رویکردی بسیار امیدوارکننده استفاده میشود. استفاده از مدلسازی فازی در این محیطها شناسایی احساسات در موسیقی نیز از طریق طبقه بندی فازی، مستقیماً هنوز به دست نیامده است و ظرفیت سیستم را برای مقابله با پیچیدگی احساسات تقویت می کند.

نتیجه گیری

در این مقاله، ما سهم داده کاوی فازی را برای بهبود تفسیرپذیری ارائه کردیم. میتوان نشان داد که تفسیرپذیری در مراحل مختلف فرآیند داده کاوی فازی رخ می دهد. بنابراین اشکال مختلفی به خود می گیرد و باید به ویژه در سطح مورد نظر مورد توجه قرار گیرد. علاوه بر این، برخی از منابع در مورد مدیریت احساسات به عنوان یک نوع خاص از اطلاعات ذهنی برای برجسته کردن برخی از حوزه های کاربردی فعلی و آینده ارائه شد که در آن داده کاوی فازی میتواند دیدگاه های جدیدی از تحقیق را پیشنهاد کند. در نتیجه، تفسیرپذیری خروجی، راه مهمی برای ستایش داده کاوی فازی در حوزه های کاربردی است. این نوع تفسیرپذیری میتواند برای کاربران دانشمند غیر فازی که به دنبال ابزاری برای مدیریت داده های خود هستند، بسیار جذاب باشد. با این حال، در داده کاوی فازی، هنوز کار زیادی برای انجام باقی مانده است. ابتدا باید بر تفسیرپذیری مدل، تمرکز کرد. مهم است که بتوان مجموعه ای از اقدامات را انتخاب کرد تا مدلها، ارزیابی و مقایسه شوند. علاوه بر این، این نوع اقدامات، ما را قادر می سازد تا داده کاوی فازی را با توجه به داده کاوی قرار دهیم و مزایای فازی را در این جنبه ها بهتر برجسته کنیم. ثانیاً، تفسیرپذیری الگوریتم یک نکته اصلی است که باید بیشتر مورد توجه قرار گیرد. بسیار مهم و اساسی است که ثابت کنیم یک توسعه فازی از یک الگوریتم کلاسیک یا یک الگوریتم فازی تازه متولد شده باید در داده کاوی فازی استفاده شود.

منابع

- A. David Hand And B. Heikki Mannila. (۲۰۰۱). Padhraic Smyth. Principles of Data Mining. The MIT Press
- A. J. Han And B. M. Kamber. (۲۰۰۱). Data Mining: Concepts and Techniques. San Diego Academic Press
- L. A. Zadeh. (۱۹۸۴). Making computer think like people. IEEE Spectrum. ۲۶-۳۲
- U. M. Fayyad, G. Piatetsky-Shapiro And P. Smyth. (۱۹۹۶). From data mining to knowledge discovery in databases, AI Mag. Vol. ۱۷. No. ۳. ۳۷-۵۴.
- C. Marsala And B. Bouchon-Meunier. (۱۹۹۹). An adaptable system to construct fuzzy decision trees. in: Proc. Of the North American Fuzzy Information Processing Society Conference, NAFIPS, New York, USA. ۲۲۳-۲۲۷.
- L.-X. Wang And J. Mendel. (۱۹۹۲). Generating fuzzy rules by learning from examples. IEEE Trans. Syst. Man Cybern. Vol. ۲۲. No. ۶. ۱۴۱۴-۱۴۲۷.
- M. Rifqi, S. Bothorel, B. Bouchon-Meunier And S. Muller. (۱۹۹۷). Similarity and prototype-based approach for classification of microcalcifications. in: Proc. of the ۷th IFSA World Congress, Prague. ۱۲۳-۱۲۸



C. Marsala. (۲۰۰۱). Fuzzy partitioning methods, in: W. Pedrycz(Ed.), Granular Computing: An Emerging Paradigm, Studies in Fuzziness and Soft Computing, Springer-Verlag. ۱۶۳-۱۸۶

R.Kruse, C. Borgelt And D. Nauck. (۱۹۹۹). Fuzzy data analysis: challenges and perspectives, in: Proc. of the ۸th IEEE International Conference on Fuzzy Systems, FUZZ-IEEE, Seoul, Korea. Vol. ۳. ۱۲۱۱-۱۲۱۶

I. Turksen. (۱۹۹۸). Fuzzy data mining and expert system development. in: Proc. Of the IEEE International Conference on Systems, Man, and Cybernetics, San Diego, CA, USA. Vol. ۲. ۲۰۵۷-۲۰۶۲

H. Ishibushi, T. Yamamoto And T. Nakashima. (۲۰۰۱). Fuzzy data mining: effect of fuzzy discretization, in: Proc. Of the IEEE International Conference on Data Mining, ICDM, San Jose, CA, USA. ۲۴۱-۲۴۸

A. Laurent. (۲۰۰۱). Generating fuzzy summaries from multidimensional databases, in: Proc. Of the Conference on Intelligent Data Analysis, IDA. ۲۴-۳۳.

S. Guillaume. (۲۰۰۱). Designing fuzzy inference system from data: an interpretability-oriented review, IEEE Trans. Fuzzy Syst. Vol. ۹. ۳. ۴۲۶-۴۴۳.

R. Weber. (۱۹۹۲). Fuzzy-ID^۳: a class of methods for automatic knowledge acquisition, in: Proceedings of the ۲nd International Conference on Fuzzy Logic. ۲۶۵-۲۶۸

M. Ramdani. (۱۹۹۳). Apprentissage a partir de donnees imparfaites, in: ۱^{er} congress biennal de ۱^{re} association francaise des sciences et technologies de ۱^{re} information et des systemes, AFCET'۹۳, Versailles. ۲۰۹-۲۱۸

Y. Yuan And M. Shaw. (۱۹۹۵). Induction of fuzzy decision trees, Fuzzy Sets Syst. Vol. ۶۹. ۱۲۵-۱۳۹

X. Boyen And L. Wehenkel. (۱۹۹۹). Automatic induction of fuzzy decision tree and its application to power system security assessmen, Fuzzy Sets Syst. Vol. ۱۰۲. No. ۱. ۳-۱۹

C.Olaru And L. Wehenkel. (۲۰۰۳). A complete fuzzy decision tree technique, Fuzzy Sets Syst. Vol. ۱۳۸. No. ۲. ۲۲۱-۲۵۴

T. -P. Hong And J.-B. Chen. (۲۰۰۰). Processing individual fuzzy attributes for fuzzy rule induction, Fuzzy Sets Syst. Vol. ۱۱۲. No. ۱. ۱۲۷-۱۴۰.

M. Gacto, R. Alcala And F.Herrera. (۲۰۱۱). Interpretability of linguistic fuzzy rule-based systems: an overview of interpretability measures, Inf, Sci. Vol. ۱۸۱. ۴۳۴۰-۴۳۶۰.

T.D. Bui, D. Heylen, M. Poel And A. Nijholt. (۲۰۰۱). Generation of facial expressions from emotion using a fuzzy rule based system, in: M. Stumptner, D. Corbett, M. Brooks(Eds), Proc. Of the ۱۴th Australian Joint Conference on Intelligence(AI), in: Lecture Notes in Computer Science. Vol. ۲۲۵۶. ۸۳-۹۴

A. Raouzaion, N. Tsapatsoulis, K. Karpouzis And S. Kollias. (۲۰۰۲). Parameterized facial expression synthesis based on MPEG-۴, EURASIP J. Appl. Signal Process. Vol. ۱. ۱۰۲۱-۱۰۳۸

R. Niewiadomski And C. Pelachaud. (۲۰۰۷). Fuzzy similarity of embodied agents, in: C. Pelachaud, J.-C. Martin, E. Andre, G. Chollet, K. Karpouzis, D. Pele(Eds.), Proc. Of the ۷th International Conference on Intelligent Virtual Agents, IVA, Paris, France, in: Lecture Notes in Computer Science. Vol. ۴۷۲۲. ۸۶-۹۸

H. Ushida, Y. Hirayama And H. Nakajima. (۱۹۹۸). Emotion model for life-like agent and its evaluation, in: Proc. Of the ۱۵th Conference on Artificial Intelligence, AAI. ۶۲-۶۹



Fuzzy data mining and sentiment management and interpretability

Mojgan Ghanbarpoor

Assistant Professor, Sarvestan Islamic Azad University

Abstract

Today, with the advancement of science, technology and technological tools, the ability to review and store large volumes of important data has been provided, and the need for science to search this data and obtain useful results has become necessary and essential. To master data science, it is essential to know mathematics and statistics. A correct understanding of the concepts of mathematics and statistics is very important in data science. Fuzzy set theory makes an important contribution to data mining, leading to fuzzy data mining. It enables the management of interpretable and subjective information, both in the input and output of the data mining process. In this paper, we discuss the concept of interpretability and sentiment management in fuzzy data mining. The abstract appears before the keywords.

Keywords: Fuzzy, data mining, interpretability, emotion management.