Archive of SID

طبقه‌بند همباشی ادراکی مبتنی بر منطق فازی توسعه یافته

مهناز کدخدا[۱]، دانشجوی دکترای مهندسی کامپیوتر؛ محمدرضا اکبرزاده توتونچی[۲]، استاد؛ فرناز صباحی[۳]، استادیار

۱- قطب علمی رایانش نرم و پردازش هوشمند اطلاعات-گروه مهندسی کامپیوتر- دانشکده مهندسی- دانشگاه فردوسی مشهد مشهد- ایران- ma.kadkhoda@um.ac.ir

۲- قطب علمی رایانش نرم و پردازش هوشمند اطلاعات-گروه مهندسی کامپیوتر- دانشکده مهندسی- دانشگاه فردوسی مشهد مشهد- ایران-akbazar@um.ac.ir

۳- دانشکده برق و کامپیوتر- دانشگاه ارومیه- ارومیه- ایران- f.sabahi@urmia.ac.ir

چکیده: پایگاه قوانین طبقه‌بند فازی همباشی (FAC)، مجموعه‌ای از قوانین فازی همباشی است که اغلب مبتنی بر داده‌های کمّی سیستم می‌باشد. درحالی‌که در دنیای واقعی- با پیچیدگی‌ها و عدم قطعیت‌های موجود- طبقه‌بندی، یک مسئله تصمیم‌گیری است که تحت تأثیر شدید دانش، تجربه، و دیدگاه شخصی افراد می‌باشد. در این مقاله، ساختار کلی ف-طبقه‌بند فازی همباشی(*f*-FAC) را در چارچوب منطق فازی توسعه‌یافته معرفی می نماییم که بیش از پیش به شیوه تفکر و استنتاج آدمی نزدیک می‌باشد. در ساختار پیشنهادی، دانش و تجربه انسانی در قالب مفهوم اعتبار فازی در هر دو مرحله تشکیل پایگاه قوانین و استنتاج طبقه بندهای فازی همباشی لحاظ شده است. در این طبقه‌بند، اعتبار مشخصه‌ها و قوانین با تلفیق نظرکارشناسان براساس هوش‌جمعی و با استفاده از محاسبات ادراکی تعیین می‌گردد. به منظور ارزیابی روش پیشنهادی، *f*-FARC-HD به عنوان توسعه‌ای از طبقه‌بند FARC-HD پیاده‌سازی شده و با تعدادی از طبقه‌بندهای دیگر- فازی همباشی و غیرفازی همباشی- مقایسه می‌شود. همچنین، کارآیی دو طبقه‌بند *f*-FARC-HD و FARC-HD در سطوح مختلف اغتشاش بررسی می‌گردد. آزمایش‌ها بر روی یک مجموعه داده واقعی ازاطلاعات بیماران بخش سوختگی بیمارستان‌های اهواز اجرا شده است. نتایج نشان می‌دهد که با در نظر گرفتن مفهوم اعتبار در *f*-FARC-HD، طبقه‌بندی کارا با پیچیدگی بسیار کم‌تر بدست می‌آید که حساسیت آن نسبت به تغییرات اغتشاش کمتر از FARC-HD می‌باشد.

واژه‌های کلیدی: اعتبار، منطق فازی توسعه‌یافته، طبقه‌بند فازی همباشی، عدم قطعیت، محاسبات ادراکی.

Perceptual Associative Classifier based on Extended Fuzzy Logic

M.kadkhoda[1], PhD student; M.-R Akbarzadeh-T[2], Professor; F.Sabahi[3], Assistant professors

1- Center of Excellence on Soft Computing and Intelligent Information Processing, Department of Computer Engineering, Ferdowsi University of Mashhad, Mashhad, Iran, Email: ma.kadkhoda@um.ac.ir

2- Center of Excellence on Soft Computing and Intelligent Information Processing, Department of Computer Engineering, Ferdowsi University of Mashhad, Mashhad, Iran, Email: akbazar@um.ac.ir

3- Faculty of Electrical and Computer Engineering, Urmia University, Urmia, Iran, Email: f.sabahi@urmia.ac.ir

Abstract:

Rule base of a Fuzzy Associative Classifier (FAC) is a collection of fuzzy associative rules that are often based on the system's quantitative data. However, due to the real-world complexities and uncertainties, classification in many practical circumstances remains a matter of art of decision-making that is strongly influenced by the knowledge, experience, and personal perspective of individuals. In this paper, we introduce the *f*-associative fuzzy classifier (*f*-FAC) in the framework of Extended Fuzzy Logic (FLe), which is more closely related to the way of thinking and reasoning of human beings. In the proposed structure, human knowledge and experience are considered by fuzzy validity concept in both phases of construction and deduction of FACs. In this classifier, the validity of the items and rules is determined by integrating the opinion of experts on the basis of wisdom of crowds and using perceptual computing. To evaluate the proposed approach, a real dataset of burn patients in Ahwaz are considered. *f*-FARC-HD is then implemented as an extension of FARC-HD associative classifier and is compared with the other approaches (associative classifier and non-associative classifier). Also, *f*-FARC-HD and FARC-HD are compared in different levels of noise. Results indicate that considering the concept of validity in the proposed extended approach, *f*-FARC-HD, leads to comparable accuracy, but at a considerably less complexity. Also, *f*-FARC-HD is less sensitive against noise.

Keywords: Validity, Extended Fuzzy Logic, Fuzzy Associative Classifier, Uncertainty, Perceptual Computing.

تاریخ ارسال مقاله: ۱۳۹۶/۱۱/۱۴
تاریخ اصلاح مقاله: ۱۳۹۷/۱۱/۰۶و ۱۳۹۸/۰۳/۱۳
تاریخ پذیرش مقاله: ۱۳۹۸/۰۵/۰۱
نام نویسنده مسئول: دکتر محمدرضا اکبرزاده توتونچی
نشانی نویسنده مسئول: ایران - مشهد- میدان آزادی- دانشگاه فردوسی مشهد- دانشکده فنی مهندسی

www.SID.ir

Archive of SID

۱- مقدمه

طبقه‌بند فازی همباشی[۱]، سیستمی طبقه‌بند مبتنی بر قوانین فازی [۳-۱] است که پایگاه قوانین آن با استفاده از روش *کشف قوانین فازی همباشی*[۲] بدست می‌آید. کشف قوانین فازی همباشی، یکی از تکنیک‌های مهم داده‌کاوی است که به استخراج قوانین بر اساس رابطه همباشی بین مشخصه‌ها می‌پردازد. اولین بار، مفهوم قانون همباشی توسط اگراوال [۴] درسال ۱۹۹۴ ارائه شد. از آن زمان، تحقیقات بسیار زیادی در مورد قوانین فازی همباشی صورت گرفته است [۷-۵]. اما اولین طبقه‌بند فازی همباشی در سال ۲۰۰۸ ارائه شد [۸]. از موارد کاربرد طبقه‌بندهای فازی همباشی می‌توان به مواردی همچون تصمیم‌گیری [۹، ۱۰]، سیستم‌های توصیه‌کننده [۱۱, ۱۲]، تجارت الکترونیک [۱۳] و کاربردهای مالی [۱۴] اشاره‌کرد.

پایگاه طبقه‌بند فازی همباشی در دو مرحله بصورت زیر بدست می آید: در مرحله اول، با استفاده از یکی از الگوریتم‌های کشف قوانین همباشی، قوانین اولیه از نمونه داده‌های یادگیری بدست می‌آید. در اکثر موارد، تعداد قوانین بدست آمده در این مرحله بسیار زیاد می‌باشد. بدین منظور در مرحله دوم، قوانین اضافی و غیرمهم حذف شده و مجموعه محدودی از قوانین به عنوان پایگاه قوانین انتخاب می‌شوند. از پایگاه قوانین بدست‌آمده، به‌منظور طبقه‌بندی نمونه داده‌های آزمایش استفاده می‌گردد. هرچه تعداد قوانین و طول آنها کمتر باشد، پایگاه قوانین خواناتر و قابل فهم‌تر بوده و پیچیدگی طبقه‌بند نیز کم‌تر است.

در طبقه‌بندهای فازی همباشی، جهت گسسته‌سازی مقادیر مشخصه‌های کمّی پیوسته از توابع عضویت فازی استفاده شده است. با توجه به تأثیری که نحوه تقسیم‌بندی فازی و دانه‌بندی هر مشخصه در طبقه‌بندی دارد، اکثر تحقیقات جدید به این موضوع پرداخته‌اند. در روشهای ارائه شده، سعی شده که تا حد امکان، بهترین دانه‌بندی برای هر مشخصه با توجه به مجموعه داده‌ها تعیین گردد. نویسندگان در [۱۵] روشی جهت تنظیم افقی[۳] برچسب‌های زبانی در قوانین طبقه‌بند با استفاده از الگوریتم ژنتیک ارائه نمودند. الگوریتم FARC-HD که در [۱۶] ارائه شده است، علاوه بر تنظیم افقی توابع عضویت، با اضافه کردن یک مرحله پیش‌پردازش قوانین اولیه، سعی در کاهش تعداد قوانین دارد. در [۱۷] الگوریتم MO-FARCG ارائه شده‌است که در واقع توسعه‌ای از الگوریتم FARC-HD است. در این الگوریتم، تعداد تقسیمات فازی مربوط به هر مشخصه، قبل از شروع مرحله تولید قوانین، بصورت خودکار تعیین می‌گردد. الگوریتم D-MOFARC نیز ساختاری درختی جهت تقسیم‌بندی فازی و دانه‌بندی هر مشخصه ارائه می‌نماید [۱۸]. در این روش، نه تنها تعداد تقسیمات فازی هر مشخصه تعیین می‌گردد، بلکه شکل هر مجموعه فازی نیز بدست می‌آید. آنتونیو سنز و همکارانش، توسعه‌ای از FARC-HD ارائه نمودند که در آن از مجموعه‌های فازی با مقادیر بازه‌ای به منظور گسسته‌سازی مشخصه‌ها استفاده نمودند [۱۲]. نسخه دیگری از این الگوریتم جهت مجموعه داده‌های نامتوازن در [۱۴] ارائه گردید.

با مروری بر طبقه‌بندهای فازی همباشی موجود، مشخص می‌گردد که در اغلب روش‌های ارائه شده، تنها از داده‌های سیستم جهت تشکیل طبقه‌بند استفاده شده است. دو چالش مهم در طبقه‌بندهای موجود را می‌توان بدین صورت بیان نمود: الف) عدم قطعیت، جزئی لاینفک از دنیای واقعی است و به صورت‌های مختلفی از قبیل ابهام، تقریب، فازی بودن، نادقیق بودن، ناسازگاری، نامشخص بودن، تصادفی بودن ظاهر می شود. بنابراین تعیین قوانین یک سیستم طبقه‌بند واقعی تنها با تکیه بر داده‌های آماری، منطقی به نظر نمی‌رسد. ب) پیچیدگی اکثر طبقه‌بندهای موجود- علیرغم الگوریتم‌های هرس و انتخاب متعدد - زیاد می‌باشد. این امر منجر به عدم خوانایی پایگاه قوانین و همچنین کندتر شدن طبقه‌بند می‌شود. این مسئله در سیستم‌های واقعی با تعداد مشخصه‌های زیاد، حادتر نیز می‌گردد. چراکه در اغلب این سیستم‌ها، مشخصه‌هایی وجود دارند که از نظر کارشناسان، تاثیری در طبقه‌بندی ندارند و یا تأثیر آن‌ها بسیار کم است، اما از آن‌جا که این مشخصه‌ها به دفعات زیاد در مجموعه داده تکرار شده‌اند، بر اساس الگوریتم کشف قوانین فازی همباشی، تعداد زیادی قانون با طول زیاد تولید می‌گردند که منجر به پیچیدگی سیستم می‌شوند.

تحقیقات نشان داده‌اندکه علیرغم وجود انواع پیچیدگی و عدم قطعیت در دنیای واقعی، انسان‌ها دارای توانایی تصمیم‌گیری منطقی بدون انجام محاسبات دقیق و فقط بر اساس درک مفاهیم هستند. انسان بر اساس تجربه و دانش خود، تنها تعدادی از مشخصه‌های تأثیرگذار در سیستم را انتخاب می‌کند و قوانین تصمیم‌گیری با توجه به میزان تأثیر آن‌ها تعیین می‌گردد. به همین دلیل نیز مشاهده می‌کنیم که اکثر تصمیمات انسان بر اساس مجموعه محدودی از قوانین ساده صورت می گیرد.

امروزه محاسبات ادراکی، شیوه جدیدی از حل مسائل را با استفاده از منطق فازی ارائه نموده است. یکی از خصوصیات منحصر به فرد این نوع محاسبات، دخالت دادن نحوه تفکر و درک انسان در حل مسائل دنیای واقعی با استفاده از منطق فازی است. اما در این منطق، انتظار می رود که درستی نتایج، اثبات پذیر باشند. در حالی‌که بسیاری از مسائل دنیای امروز مخصوصاً در حوزه تصمیم‌گیری، اثبات پذیر نیستند.

پروفسور زاده در سال ۲۰۰۹، برای حل معضل اثبات پذیری، مفهوم اعتبار[۴] را در قالب منطق فازی توسعه یافته[۵] معرفی کرد [۱۹] و پس از آن، مقالات متعددی به این موضوع پرداخته‌اند [۲۲-۲۰]. منطق فازی توسعه یافته با ارائه مفهوم اعتبار می‌تواند به مدل‌سازی جنبه دیگری از عدم قطعیت در مسائل دنیای واقعی کمک کند. در این منطق، هم راه حل و هم پاسخ مسئله می‌توانند تا *درجه‌ای معتبر* باشند.

در این تحقیق برآنیم که سیستمی طبقه‌بند برای مسائل پیچیده دنیای واقعی با فرض عدم‌قطعیت‌های موجود، طراحی نماییم که بیش از پیش به شیوه تفکر و استنتاج آدمی نزدیک باشد. بدین‌منظور، با وارد نمودن مفهوم اعتباربه ساختار طبقه‌بندهای فازی همباشی موجود، ف-طبقه‌بندی درچهارچوب منطق فازی توسعه یافته (*f*-FAC) پیشنهاد

www.SID.ir

Archive of SID

داده‌ایم. در طبقه‌بند پیشنهادی، مشخصه‌ها و قوانین از میزان اعتبار یکسانی برخوردار نیستند، بلکه اعتبار آنها بر اساس هوش جمعی و با استفاده از محاسبات ادراکی تعیین می‌گردد. ایده اساسی در هوش جمعی[۶] این است که عقلانیت موجود در یک جمعیت، نهایتاً به نتایج درست همگرا می‌شود [۲۳]. بنابراین در این روش، نه یک فرد، بلکه جمعی از کارشناسان با تجربیات، پیش‌زمینه و دانش متفاوت، در تعیین اعتبار سهیم هستند و نظر خود را با توجه به ماهیت کیفی مفهوم اعتبار، با استفاده از کلمات بیان می‌کنند.

فلسفه شکل‌گیری این تحقیق، پرداختن به عدم قطعیت در فضای ف- انتقالی[۷] می‌باشد. این دیدگاه فلسفی جدید، نوآوری‌های زیر را به همراه داشته است:

۱. ارائه ساختار کلی یک ف-طبقه‌بند فازی همباشی براساس منطق فازی توسعه یافته.
۲. مدل‌کردن اعتبار به عنوان بعد دیگری از عدم قطعیت که با مروری که بر طبقه‌بندهای همباشی موجود داشته‌ایم، تاکنون به آن پرداخته نشده است. در این سیستم، نه تنها مشخصه‌ها، بلکه قوانین نیز دارای اعتبار می‌باشند.
۳. طبقه‌بند پیشنهادی را می‌توان یک سیستم تلفیقی از انسان/ماشین دانست. اعتبار هر مشخصه که از اجماع نظر کارشناسان تعیین می شود، ریشه در ادراک، دانش و تجربه بشر دارد و به صورت کلامی نشان داده می‌شود.

به منظور ارزیابی روش پیشنهادی، ف-طبقه‌بند *f*-FARC-HD که ف-انتقالی طبقه‌بند فازی همباشی FARC-HD است، پیاده‌سازی شده و بر روی مجموعه داده واقعی BURN اجرا می‌شود. این مجموعه داده حاوی مشخصات ۴۵۵۶ بیمار است که در بازه زمانی یک سال در بخش سوختگی بیمارستانهای اهواز جمع‌آوری شده‌است. ارزیابی در دو مرحله انجام می‌گیرد. ابتدا کارآیی طبقه‌بند *f*-FARC-HD با چند طبقه‌بند دیگر- فازی همباشی و غیر فازی همباشی- مقایسه می‌شود. در مرحله بعد، با توجه به عدم قطعیت موجود در دنیای واقعی، رفتار طبقه‌بند *f*-FARC-HD و نسخه اولیه‌اش FARC-HD در سطوح مختلف اغتشاش[۸] بررسی می‌شود.

در ادامه، مقاله به‌صورت زیر سازمان‌دهی شده است. در بخش ۲، مروری داریم بر مفاهیم اولیه منطق فازی توسعه که در این تحقیق به آن‌ها اشاره شده است. ساختار کلی ف-طبقه بند فازی همباشی پیشنهادی در بخش ۳ تعریف می‌شود. در بخش ۴، الگوریتم *f*-FARC-HD به عنوان یک ف-طبقه‌بند فازی همباشی نمونه ارائه شده است. بخش ۵ نیز به بررسی نتایج حاصل از آزمایش‌ها بر روی مجموعه داده BURN می‌پردازد. در پایان، در بخش آخر، مطالب جمع بندی می شود.

۲- منطق فازی توسعه یافته

منطق فازی با توجه به انعطاف پذیری بالای آن، در حل بسیاری از مسائل استفاده شده است [۲۴, ۲۵]. اما این منطق، برخلاف آنچه تصور می شود، خودش فازی نیست. اساساً منطق فازی، منطقی دقیق درباره نادقیقی است [۲۶]. دقیق بودن منطق فازی، به‌خاطر ارتباط مجموعه‌های فازی با توابع عضویت معین و ارتباط دانه‌بندی با قیود تعمیم‌یافته معین است که مبین خصلت معیارپذیری منطق فازی است. به بیان دیگر، نتایج در محدوده منطق فازی، اثبات‌پذیر می‌باشند. درحالی‌که در دنیای واقعی، بسیاری از مسائل، مخصوصاً در حوزه تصمیم‌گیری این شرط را نقض می‌کنند. مسأله پیدا کردن سریع‌ترین مسیر، نمونه‌ای از این مسائل است. زیرا سریع‌ترین مسیر، الزاماً کوتاه‌ترین مسیر نیست. این مسیر را راننده بر اساس تجربه‌اش انتخاب می‌کند. اینکه این مسیر، سریع‌ترین است، نه قابل اثبات است و نه قابل نفی. بنابراین دراین حالت نمی‌توان به یک پاسخ معین قابل اثبات رسید و برای حل آن نیاز به نوعی متفاوت از محاسبات می‌باشد که بیش‌تر بر پایه مفاهیم استوار است [۱۹].

منطق فازی توسعه‌یافته برای اولین بار توسط پروفسور زاده به منظور پرداختن به مسائل معیارناپذیر ارائه گردید. با ارائه مفهوم اعتبار در این منطق، امکان استنتاج اطلاعات نادقیق و ناکامل در مسائل معیارناپذیر فراهم شده‌است. به عبارت دیگر، منطق فازی توسعه‌یافته از اضافه شدن منطق فازی معیارناپذیر FLu به منطق فازی معیارپذیر FLp بدست می آید [۲۷]. FLu منطقی است که نتایج در محدوده آن لازم نیست اثبات‌پذیر باشند ولی حداقل اثبات‌پذیر درجه‌ای یعنی تا درجه‌ای معتبر باشند.

مفهوم اعتبار، کاملاً متفاوت ازمفاهیم امکان و احتمال است. امکان و احتمال از نظر ساختاری با هم متفاوت هستند اما هر دو مربوط به بعد محتوایی مسئله می‌باشند، در حالیکه اعتبار مربوط به عمق مسئله است. در واقع، با اعمال مفهوم اعتبار می‌توان عدم قطعیت بیشتری را مدل کرد. به عنوان مثال، بطور کلی، نمی‌توان یک عدد را بر صفر تقسیم نمود اما می‌توان صفر را به عنوان یک عدد بسیارکوچک مانند ε در نظر گرفت و تقسیم را انجام داد، اما اعتبار نتیجه کاهش می‌یابد. به عبارت دیگر، با انتشار اعتبار، می‌توان منعکس کننده اثر تغییر صفر به ε در فرآیند تقسیم بود [۲۸].

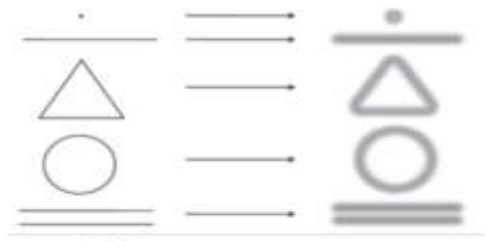
در این قسمت، به طور مختصر به چند مفهوم اساسی در منطق فازی توسعه‌یافته می‌پردازیم. توضیحات بیشتر در [۱۹, ۲۹] آمده است.

۲-۱- ف-انتقالی

ف-انتقالی، نگاشتی از " یک" به "زیاد " است که در آن اصالت "یک" حفظ شده باشد [۱۹]. به تعبیری دیگر، یک ف-انتقالی مربوط به مفهوم قطعی C، یک مفهوم فازی است که توسط *f*-C نشان داده

www.SID.ir

Archive of SID



شکل ۱: ف-انتقالی اشکال هندسی [۱۹]

می‌شود و به آن ف-انتقالی C گفته می‌شود. *f*-C را می‌توان قیاسی از رسم C با دست توسط اسپری دانست. همان‌طور که در شکل ۱[۱۹] دیده می‌شود، زاده، ف-انتقالی را با اشکال هندسی نشان داده است، اما ف-انتقالی بر روی هر مفهومی می‌تواند اعمال شود. مفاهیمی مانند ف-محدب، ف-قانون، ف-الگوریتم و ف-قضیه مثال‌هایی از این دسته‌اند.

۲-۲- فرض تطبیق پذیری

توانایی مدل کردن اطلاعات ناقص معیارناپذیر در منطق فازی توسعه یافته با این فرض ایجاد می‌شود که ف-انتقال‌ها تطبیق‌پذیر در نظر گرفته می‌شوند [۱۹]. درحقیقت، فرض می‌شود *f*-C ها دارای قابلیت انطباق زیاد با Cها باشند، مگر اینکه خلافش گفته شود. به این فرض، فرض تطبیق‌پذیری گفته می‌شود.

۲-۳- اصل اعتبار

پروفسور زاده در [۱۹] اصل اعتبار را ارائه می‌نماید که در مورد مقدار اعتباری است که نتیجه استنتاج منطق فازی توسعه‌یافته خواهد داشت. این اصل، به صورت زیر بیان می‌شود:

" فرض کنید p یک نتیجه $p\text{-valid}$ از زنجیره‌ای گزاره $p_1, p_2, ..., p_n$ باشد. در این صورت $f\text{-}p$ نتیجه‌ای با اعتبار $f\text{-valid}$ از گزاره‌های $f\text{-}p_1, f\text{-}p_2, ..., f\text{-}p_n$ خواهد بود و $f\text{-}p$ دارای اعتبار بالایی است."

۲-۴- اصل P/I

فرض کنید h یک تابع یا یک عمل‌گر باشد. برای محاسبه $h(f\text{-}c)$ زاده اصل P/I را ارائه می‌نماید [۱۹]. بر اساس این اصل، برای محاسبه $h(f\text{-}c)$ می‌توان $f\text{-}h(c)$ را محاسبه نمود که یک تقریب ف-اعتبار از $h(f\text{-}c)$ است. به عبارت دیگر، بر اساس این اصل داریم:

$$f\text{-}h(c)\ f\text{-} = h(f\text{-}c) \qquad (1)$$

که در آن $f\text{-} =$ (ف- تساوی) به معنای تقریباً مساوی است [۳۰].

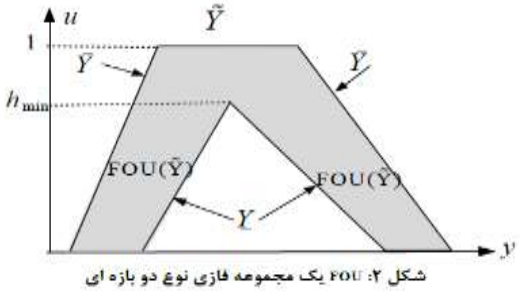
۳- تعیین اعتبار مشخصه‌ها

در این بخش، روشی جهت تعیین اعتبار مشخصه‌ها با استفاده از نظر کارشناسان ارائه می‌گردد. مبنای این روش، بر پایه هوش جمعی و ادراک انسان است. در هوش جمعی، فرض بر این است که عقلانیت موجود در یک جمعیت، نهایتاً به نتایج درست همگرا می‌شود. چرا که نظر هر فرد شامل هم اطلاعات درست و هم نادرست، هر کدام به میزانی می‌باشد. در صورتی که این نظرها تجمیع گردد، اطلاعات درست، هم جهت بوده و تقویت می‌گردند، در حالی‌که اطلاعات نادرست، تأثیر یکدیگر را خنثی می‌نمایند [۲۱].

تعیین اعتبار یک مشخصه، قضاوتی است که هر فرد خبره در مورد میزان اهمیت آن مشخصه انجام می‌دهد و عموماً تحت تأثیر دیدگاه شخصی، تجربه و یا پیش زمینه شخص می‌باشد. بنابراین اعتبار یک مفهوم کیفی است. به همین‌دلیل هم در این تحقیق، کارشناسان، نظر خود را با کلمات زبان طبیعی مانند "کم"، "زیاد" و یا "نسبتاً زیاد" بیان می‌نمایند. تعیین اعتبار مشخصه‌ها در دو مرحله انجام می‌شود که در ذیل به آنها می‌پردازیم:

- **مدلسازی کلمات**

یکی از چالش‌های موجود در مبحث محاسبات ادراکی این است که کلمات از نظر افراد مختلف ممکن است بار معنایی متفاوتی داشته باشند، بنابراین غیرقطعی هستند. مجموعه‌های فازی نوع دو بازه‌ای[۹]، از آنجا که با FOU [۱۰] مشخص می‌شوند، قادر به نمایش این عدم قطعیت می‌باشند. در شکل ۲، FOU مربوط به یک مجموعه فازی نوع دو بازه‌ای نشان داده شده است که در آن، $\tilde{Y}$ یک مجموعه فازی نوع ۲ بازه‌ای، $\bar{Y}$ تابع عضویت بالایی و $\underline{Y}$ تابع عضویت پایینی است. همان‌طور که در شکل مشاهده می‌شود، FOU فضای بین $\bar{Y}$ و $\underline{Y}$ است. روش‌هایی جهت مدل‌سازی کلمات بصورت FOU با استفاده از تلفیق نظر افراد مختلف ارائه شده‌اند [۳۱]، [۳۲]. ما در این مقاله، برای مدلسازی کلمات از کتابچه رمزی[۱۱]استفاده کرده‌ایم که توسط مندل و همکارانش به روش داده‌های بازه‌ای[۱۲] تهیه شده است. جزئیات روش مندل در [۳۳] آمده است.



شکل ۲: FOU یک مجموعه فازی نوع دو بازه ای

- **تجمیع نظرات کارشناسان**

به منظور جمع‌بندی نظر کارشناسان که در بخش قبلی به‌صورت کلمات تهیه شده است، از یکی از عملگرهای محاسبات ادراکی استفاده می‌کنیم. این عملگر که مندل آن را متوسط وزنی زبانی[۱۳] می‌نامد، عمل میانگین وزنی را بر روی کلمات انجام داده و نتیجه آن نیز یک کلمه می باشد [۳۳]. بنابراین، اعتبار مشخصه r به صورت زیر تعریف می‌شود:

تعریف ۲: فرض کنید $\tilde{Y}_{ri}$, $(i=1,2,...\sigma)$ نظرات σ کارشناس در مورد مشخصه r می‌باشد که به صورت کلامی بیان شده‌اند و $\tilde{\omega}_i$ وزنی است که به هریک از کارشناسان تخصیص داده شده است. اعتبار مشخصه r که خود نیز یک مجموعه فازی نوع دو بازه‌ای است، بصورت زیر بدست می‌آید:

www.SID.ir

Archive of SID

$$\tilde{v}_r = \frac{\sum_{i=1}^{\sigma} \tilde{\Upsilon}_{ri}\tilde{\omega}_i}{\sum_{i=1}^{\sigma} \tilde{\omega}_i} \qquad (۲)$$

به منظور محاسبه $\tilde{v}_r$، کافی است که $\overline{v}_r$ و $\underline{v}_r$ محاسبه شوند. هنگامی که $\overline{v}_r$ و $\underline{v}_r$ محاسبه شوند، FOU مربوط به $\tilde{v}_r$ که فضای بین $\overline{v}_r$ و $\underline{v}_r$ است، بدست می‌آید. جزئیات این الگوریتم نیز در [۳۳] آمده است. لازم به ذکر است که از آنجا که میزان تخصص و تجربه کارشناسان متفاوت است، در این تحقیق به هر یک از آنان وزنی اختصاص داده شده تا در محاسبات لحاظ گردد. در غیر این صورت می‌توان وزن همگی را یکسان در نظر گرفت.

۴- ساختار کلی ف-طبقه‌بند فازی همباشی

در این بخش، ف-طبقه‌بند فازی همباشی پیشنهادی تعریف می‌گردد. این طبقه‌بند، یک ف-انتقالی از طبقه‌بند فازی همباشی است که معیارهای مهم در این طبقه‌بند، با توجه به مفهوم اعتبار، در چارچوب منطق فازی توسعه یافته بازنگری شده و مجدداً تعریف می‌گردند.

روال کار در طبقه‌بندهای فازی همباشی، بطور خلاصه بدین صورت است: به‌منظور تعیین پایگاه قوانین ابتدا الگوهای تکراری در مجموعه داده‌ها استخراج می‌گردند. الگوی تکراری، الگویی است که *پشتیبان*[۱۴] در مجموعه‌داده‌ها بیشتر از *حداقل پشتیبان* تعیین شده باشد. سپس، از روی الگوهای تکرار، قوانین پیشنهادی برای هر طبقه بدست‌می‌آید. بدین صورت که الگوی تکراری در مقدم قانون و برچسب طبقه در تالی قانون قرارداده می‌شود. در صورتی‌که *اطمینان*[۱۵] قانون بیشتر از *حداقل اطمینان* تعیین شده باشد، این قانون، به پایگاه قوانین اضافه می‌شود. لازم به ذکر است که در اکثر موارد، اطمینان قانون به عنوان وزن قانون در نظر گرفته می‌شود. در پایان، عمل انتخاب قوانین بهتر و هرس قوانین اضافی به منظور ایجاد یک پایگاه قوانین فشرده و کارا انجام می‌شود. در هنگام طبقه‌بندی، با توجه به پایگاه قوانینی که در مرحله قبل تشکیل شده است، طبقه نمونه ورودی $\hat{x}$ تعیین می‌گردد. بدین‌منظور از روش‌های متعددی مانند *بیشترین درجه همباشی*[۱۶] و *رأی وزن‌دار*[۱۷] جهت استنتاج استفاده می‌شود که هسته اصلی اکثر آنها، معیار درجه *همباشی* می‌باشد. در ادامه به ارائه ف-طبقه‌بند فازی همباشی پیشنهادی و تعریف معیارها در این طبقه‌بند می‌پردازیم.

فرض کنید مجموعه داده D از N نمونه داده برچسب دار از k طبقه به‌صورت $x_p = (x_{p1},...,x_{pi},...,x_{pm},C_p)$ موجود است. بطوری‌که $p = \{1,2,...,N), i = (1,2,...,m)$ می‌باشد و $C_p \in \{C_1,C_2,...,C_k\}$ برچسب طبقه x_p است. x_{pi}، iامین مشخصه از p امین نمونه در مجموعه داده می‌باشد. هر یک از مشخصه‌ها دارای یک میزان اعتبار می‌باشند که به صورت $\tilde{v} = (\tilde{v}_1,...,\tilde{v}_i,...,\tilde{v}_m)$ نشان داده می‌شوند. بطوریکه $\tilde{v}_i$ اعتبار فازی iامین مشخصه در مجموعه داده‌هاست و توسط کارشناسان تعیین می‌گردد. یک طبقه بند فازی همباشی مبتنی بر منطق فازی توسعه‌یافته یا به اختصار ف-طبقه‌بند فازی همباشی بصورت زیر تعریف می‌شود:

تعریف ۱: یک ف-طبقه‌بند فازی همباشی، مجموعه‌ای از ف- قانون است که با استفاده از الگوریتم ف-کشف قوانین فازی همباشی از مجموعه داده‌های یادگیری و نظرات کارشناسان بدست آمده است. j امین ف-قانون در این طبقه‌بند بصورت زیر تعریف می‌گردد:

$$f\text{-}R_j : \text{if } (x_{j1},\tilde{v}_1) \text{ is } \tilde{A}_{j1} \text{ and} \ldots (x_{jr},\tilde{v}_r) \text{ is } \tilde{A}_{jr} \text{ and} \ldots \\ (x_{jm},\tilde{v}_m) \text{ is } \tilde{A}_{jm} \Rightarrow \text{class}=(C_j,V_{R_j}) \text{ with } f\text{-}RW_{R_j} \qquad (۳)$$

که در آن $\tilde{A}_{jr}$ ها کلمات زبانی، $\tilde{v}_r$ اعتبار مشخصه x_{jr}، $f\text{-}RW_{R_j}$ ف-وزن قانون و V_{R_j}، اعتبار قانون و به معنای میزان درستی قانون می‌باشد و $C_j \quad C_j \in \{C_1,C_2,...,C_k\}$, نیز برچسب طبقه قانون است.

از آن‌جا که اعتبار، بر روی مقدار یک مشخصه تأثیر می‌گذارد، رابطه (۳) را به صورت زیر می‌توان نوشت:

$$f\text{-}R_j : \text{if } x_{j1} \text{ is } f\text{-}\tilde{A}_{j1} \text{ and } \ldots x_{jr} \text{ is } f\text{-}\tilde{A}_{jr} \text{ and } \ldots x_{jm} \text{ is } \\ f\text{-}\tilde{A}_{jm} \Rightarrow \text{class= } f\text{-}C_j \text{ with } f\text{-}RW_{R_j} \qquad (۴)$$

که به‌طور خلاصه، آن‌را به‌صورت $f\text{-}R_j : f\text{-}A_j \Rightarrow f\text{-}C_j$ نشان می‌دهیم.

$\tilde{v}_r$ ها در رابطه (۳) با استفاده از فرمول (۲) در بخش قبل بدست می‌آید. اعتبار هر ف- قانون (V_{R_j}) در این تحقیق، بصورت میانگینی از اعتبار مشخصه‌ها در مقدم آن ف-قانون در نظرگرفته شده است. داریم:

$$V_{R_j} = \text{avg}(\xi(\tilde{v}_{j1}),...,\xi(\tilde{v}_{jr}),...\xi(\tilde{v}_{jm})), \qquad (۵)$$

بنابراین، اعتبار قوانین نیز بطور غیر مستقیم با استفاده از نظر کارشناسان تعیین شده است. پشتیبان و اطمینان (وزن) هر ف-قانون نیز به‌صورت زیر تعریف می‌گردد:

$$supp_{f\text{-}R_j} = f\text{-}supp_{R_j} = \frac{\sum_{x_p \in C_j} \Pi_{f\text{-}A_j}(x_p)}{|N|} = \frac{\sum_{x_p \in C_j} f\text{-}\Pi_{A_j}(x_p)}{|N|} \qquad (۶)$$

$$RW_{f\text{-}R_j} = f\text{-}RW_{R_j} = f\text{-}conf_{R_j} = \frac{\sum_{x_p \in C_j} \Pi_{f\text{-}A_j}(x_p)}{\sum_{x_p \in D} \Pi_{f\text{-}A_j}(x_p)} = \frac{\sum_{x_p \in C_j} f\text{-}\Pi_{A_j}(x_p)}{\sum_{x_p \in D} f\text{-}\Pi_{A_j}(x_p)} \qquad (۷)$$

که در آن، D مجموعه داده‌ها، $|N|$ تعداد نمونه ورودی در D، C_j برچسب طبقه قانون و $f\text{-}\Pi_{A_j}(x_p)$ ، ف-درجه انطباق[۱۸] نمونه ورودی x_p با مقدم قانون $f\text{-}R_j$ می‌باشد. ف-درجه انطباق با توجه به قضیه زیر، قابل محاسبه است:

قضیه ۱: ف-درجه انطباق نمونه داده x_p با مقدم قانون $f\text{-}R_j$ بصورت زیر می‌باشد:

$$f\text{-}\Pi_{A_j}(x_p) = \min\big((\mu_{\tilde{A}_{j1}}(x_{p1}) \times \xi(\tilde{v}_1)),...,(\mu_{\tilde{A}_{jr}}(x_{pr}) \times \\ \xi(\tilde{v}_r)),...,(\mu_{\tilde{A}_{jm}}(x_{pm}) \times \xi(\tilde{v}_m))\big) \qquad (۸)$$

که در آن $\times$ یک عملگر *T-norm* (مانند ضرب و یا مینیمم) است و عملگر $\xi(.)$، centroid فازی نوع ۲ بازه‌ای را محاسبه می‌کند.

اثبات: طبق تعریف درجه انطباق داریم:

$$\Pi_{A_j}(x_p) = \min_{r=1,...,m}(\mu_{\tilde{A}_{jr}}(x_{pr}))$$

www.SID.ir

Archive of SID

بنابراین، می‌توانیم درجه انطباق x_p با قانون $f\text{-}R_j$ را به صورت زیر بنویسیم:

$$\Pi_{f\text{-}A_j}(x_p) = \min_{r=1,\dots,m}(\mu_{f\text{-}\tilde{A}_{jr}}(x_{pr})) \quad (۹)$$

که با توجه به اصل P/I برابر است با:

$$f\text{-}\Pi_{A_j}(x_p) = \min_{r=1,\dots,m}(f\text{-}\mu_{\tilde{A}_{jr}}(x_{pr})) \quad (۱۰)$$

$$f\text{-}\Pi_{A_j}(x_p) = \min(f\text{-}\mu_{\tilde{A}_{j1}}(x_{p1}),\dots, f\text{-}\mu_{\tilde{A}_{j1}}(x_{p1}),\dots, f\text{-}\mu_{\tilde{A}_{jm}}(x_{pm})) \quad (۱۱)$$

براساس اصل اعتبار داریم:

$$f\text{-}\Pi_{A_j}(x_p) = \min((\mu_{\tilde{A}_{j1}}(x_{p1}),\tilde{v}_1),\dots,(\mu_{\tilde{A}_{jr}}(x_{pr}),\tilde{v}_r)\dots, (\mu_{\tilde{A}_{jm}}(x_{pm}),\tilde{v}_m)) \quad (۱۲)$$

به منظور هم‌مرتبه کردن اعتبار و امکان، از آن‌جا که اعتبار از فازی مرتبه دو است، ابتدا مرکز اعتبارتوسط عملگر ξ محاسبه می‌شود. این عمل تاثیری روی اعتبار نخواهد داشت فقط فضای کاری آن را از حوزه فازی نوع ۲ به فازی نوع یک می‌نگارد. عملگر $\times$ نیز یک عملگر $norm\text{-}T$ مانند ضرب یا می‌نیمم می‌باشد که باعث اعمال تأثیر اعتبار بر عضویت فازی یک عنصر می‌شود. بنابراین داریم:

$$f\text{-}\Pi_{A_j}(x_p) = \min((\mu_{\tilde{A}_{j1}}(x_{p1})\times\xi(\tilde{v}_1)),\dots,(\mu_{\tilde{A}_{jr}}(x_{pr})\times \xi(\tilde{v}_r)),\dots,(\mu_{\tilde{A}_{jm}}(x_{pm})\times\xi(\tilde{v}_m))) \quad (۱۳)$$

و قضیه اثبات می‌شود.

مثال ۱: فرض کنید مجموعه داده D شامل ۵ نمونه داده و سه مشخصه می‌باشد. در جدول ۱ این مجموعه داده با مقادیر توابع عضویت هر مشخصه نشان داده شده است. میزان اعتبار هر مشخصه، بعد از غیرفازی سازی به‌صورت زیر می‌باشد:

$$\xi(\tilde{v}_1) = ۰/۳ \quad \xi(\tilde{v}_2) = ۰/۸ \quad \xi(\tilde{v}_3) = ۰/۴$$

پشتیبان و اطمینان ف-قانون زیر

$$f\text{-}Rule: if\ (x_1\ is\ متوسط\ and\ x_2\ is\ زیاد) \Rightarrow C1$$

را بدست می‌آوریم. توجه کنید که به منظور خوانایی بیشتر، در ف-قانون، فقط نام مشخصه‌ها و طبقه ذکر شده است.

$$supp_{f\text{-}Rule} = (\min(۰/۶\times۰/۳,۰/۷\times۰/۸) + \min(۰/۳\times۰/۳,۰/۵\times۰/۸))/\ ۵ = ۰/۰۵۴$$

$$RW_{f\text{-}Rule} = conf_{f\text{-}Rule} = ۰/۰۵۴\ /\ ((\min(۰/۶\times۰/۳,۰/۷\times۰/۸)+ \min(۰/۹\times۰/۳,۰/۸\times۰/۸) +\min(۰/۷\times۰/۳,۰/۴\times۰/۸) + \min(۰/۳\times ۰/۳,۰/۵\times۰/۸) +\min(۰/۴\times۰/۳,۰/۶\times۰/۸)) = ۰/۰۶۲$$

جدول ۱: مجموعه داده D با مقادیر توابع عضویت

طبقه	<x3, کم>	<x2, زیاد>	<x1, متوسط>
C1	۰/۳	۰/۷	۰/۶
C2	۰/۴	۰/۸	۰/۹
C2	۰/۱	۰/۴	۰/۷
C1	۰/۸	۰/۵	۰/۳
C2	۰/۹	۰/۶	۰/۴

اگر می‌نیمم اطمینان، در این الگوریتم ۰/۵ تعیین شده باشد، از آن‌جا که ۰/۵ > ۰/۰۶۲ است، این ف-قانون به پایگاه قوانین اضافه نمی شود.

همان‌طور که قبلاً نیز اشاره شد، یک معیار دیگر درطبقه‌بندهای فازی همباشی، معیار درجه همباشی است که در طبقه‌بندی نمونه‌داده‌ها بکار می‌رود. درجه همباشی بین قانون R_j و نمونه ورودی $\hat{x}$، در طبقه‌بندهای فازی همباشی بصورت زیر تعریف می‌شود:

$$\delta_{R_j}(\hat{x}) = \Pi_{A_j}(\hat{x}).\ RW_{R_j} \quad (۱۴)$$

با توجه به رابطه ۱۴، درجه همباشی بین قانون $f\text{-}R_j$ و نمونه ورودی $\hat{x}$، را به صورت زیر می‌توان نوشت:

$$\delta_{f\text{-}R_j}(\hat{x}) = \Pi_{f\text{-}A_j}(\hat{x})\ .\ RW_{f\text{-}R_j}.\ V_{R_j} \quad (۱۵)$$

و طبق اصل P/I داریم:

$$f\text{-}\delta_{R_j}(\hat{x}) = f\text{-}\Pi_{A_j}(\hat{x})\ .\ f\text{-}RW_{R_j}.\ V_{R_j} \quad (۱۶)$$

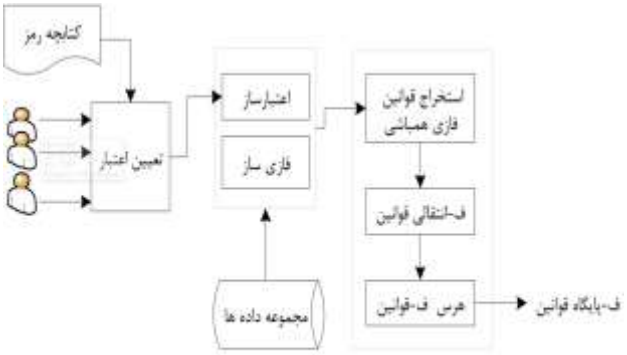
که در آن، V_{R_j} و $f\text{-}RW_{R_j}$ به ترتیب، اعتبار و وزن ف-قانون می‌باشند که از روابط ۵ و ۷ بدست می‌آیند و $f\text{-}\Pi_{A_j}(\hat{x})$ ، ف-درجه انطباق نمونه داده $\hat{x}$ با قانون $f\text{-}R_j$ است که در رابطه ۱۳ تعریف شده است. بعد از تعیین ف-درجه همباشی هر نمونه ورودی با قوانین موجود، با یکی از روش‌های استنتاج، طبقه نمونه ورودی مشخص می‌شود.

۵- الگوریتم ف-طبقه‌بند f-FARC-HD

هدف اصلی این مقاله، معرفی یک دیدگاه جدید با بکارگیری مفهوم اعتبار در طراحی الگوریتم‌های طبقه‌بند می‌باشد. لذا، در این بخش، یک طبقه‌بند فازی همباشی معتبر (FARC-HD) جهت توسعه و آزمودن این دیدگاه انتخاب شده است. لازم به ذکر است که FARC-HD یکی از الگوریتم‌های پایه‌ای و مطرح در حوزه طبقه‌بندهای فازی همباشی است که بسیاری از طبقه‌بندهای فازی همباشی [۱۲, ۱۴, ۱۸]، با تغییرات کمی، توسعه‌ای از آن هستند. شکل ۳، مراحل تشکیل طبقه‌بند f-FARC-HD که یک ف-انتقالی از FARC-HD است را نشان می‌دهد. در ادامه، به طور مختصر به چگونگی تشکیل ف-طبقه‌بند f-FARC-HD می پردازیم.

www.SID.ir

Archive of SID



شکل۳ : تشکیل ف-طبقه بند *f*-FARC-HD

۵-۱- تعیین اعتبار مشخصه‌ها

اعتبار زبانی هریک از مشخصه‌ها با استفاده از نظر کارشناسان با توجه به بخش ۳-۱ بدست می‌آید.

۵-۲- استخراج ف-قوانین اولیه

همان‌طور که در شکل ۳ نشان داده شده، نمونه داده‌های ورودی ابتدا وارد فازی‌ساز و اعتبارساز می‌گردند. سپس ف-قوانین اولیه از روی ف-نمونه‌های ورودی بدین صورت بدست می‌آیند: برای هر طبقه، یک درخت جستجو از تمامی ف-الگوهای تکراری[۱۹] آن طبقه تشکیل می شود. ف-الگوی تکراری، الگویی است که ف-پشتیبان آن در مجموعه‌داده‌ها بیشتر از *مینیمم پشتیبان* باشد. سپس، از روی درخت تشکیل شده، ف-قوانین اولیه برای هر طبقه بدست‌می‌آید. بدین صورت که ف-الگو در مقدم قانون و برچسب طبقه در تالی قانون قرارداده می شود. ف-قانونی که ف-اطمینان آن بیشتر از *مینیمم اطمینان باشد*، این ف-قانون انتخاب شده و به لیست قوانین اولیه اضافه می‌گردد. در پایان این مرحله، اعتبار هر قانون بر اساس رابطه (۵) تعیین می‌گردد.

۵-۳- پیش انتخاب ف-قوانین

از آنجا که تعداد ف-قوانین در مرحله قبل زیاد است، در مرحله دوم با استفاده از الگوریتم *اکتشاف زیرگروه* [۳۴] ف-قوانین بهتر پیش‌انتخاب می‌شوند. بدین منظور، به هر نمونه‌داده در مجموعه‌داده‌ها، یک وزن به‌صورت $w(i)=1/(i+1)$ اختصاص داده می‌شود. مقدار i نشان‌دهنده این است که یک نمونه‌داده چند بار توسط یک ف-قانون پوشش داده شده‌است. در ابتدا، وزن همگی یکسان $w(0)=1$ در نظر گرفته می‌شود. برای هر طبقه، بهترین قانون انتخاب می‌شود. سپس وزن نمونه‌داده‌هایی که توسط قانون، پوشش داده‌شده‌اند، کم می‌شود. با این روش، نمونه‌داده‌هایی که هنوز پوشش داده‌نشده‌اند، در تکرارهای بعدی شانس بیشتری دارند که پوشش داده‌شوند. زمانی که شمارنده i یک نمونه‌داده به مقدار آستانه k برسد، نمونه‌داده حذف می‌گردد. ف-قوانین باقیمانده، مرتب می‌شوند و عملیات ادامه پیدا می‌کند تا زمانی که یا تمامی نمونه‌داده‌ها حذف شوند و یا اینکه دیگر ف-قانونی باقی نماند. به منظور ارزیابی کیفیت ف-قوانین، معیار $wWRAcc''$ بصورت زیر تعریف شده‌است:

$$wWRAcc''(f\text{-}A\rightarrow f\text{-}C_j)=\frac{f\text{-}n''(A.C_j)}{n'(C_j)}.\left(\frac{f\text{-}n''(A.C_j)}{f\text{-}n''(A)}-\frac{n(C_j)}{N}\right) \quad (۱۷)$$

در این معیار، $f\text{-}n''(A)$ برابر است با مجموع حاصل‌ضرب وزن‌های تمامی ف-نمونه‌داده‌های پوشش‌داده‌شده توسط ف-قانون در ف-درجه انطباق ف-نمونه‌داده با مقدم ف-قانون. $f\text{-}n''(A.C_j)$، مجموع حاصل‌ضرب وزن‌های تمامی ف-نمونه‌داده‌های درست پوشش داده شده در ف-درجه انطباقشان با بخش مقدم ف-قانون می‌باشد. $n(C_j)$، تعداد ف-نمونه داده‌های طبقه C_j و $n'(C_j)$ مجموع وزن‌های نمونه‌داده‌های طبقه C_j است در مجموعه داده‌ها است. N نیز تعداد کل ف-نمونه‌داده‌ها در مجموعه داده می‌باشد.

۵-۴- تنظیم و انتخاب ف-قوانین با استفاده از الگوریتم تکاملی چند هدفه

در این مرحله، با استفاده از یک الگوریتم تکاملی دو هدفه، به طور همزمان، عمل انتخاب ف-قوانین و تنظیم افقی پارامترهای ف-پایگاه قوانین صورت می‌گیرد. منظور از پارامترها، پارامترهای توابع عضویت مثلثی مشخصه‌ها می‌باشد و هدف الگوریتم تکاملی این است که بهترین وضعیت قرارگرفتن توابع عضویت در محور افقی را مشخص کند. درادامه، مشخصات اصلی الگوریتم تکاملی، به‌صورت خلاصه بیان می‌شود. برای جزئیات بیشتر می‌توانید به [۱۶] مراجعه نمائید.

- **اهداف الگوریتم تکاملی**

هر کروموزوم با یک بردار دو بعدی مشخص می‌گردد. هر کدام از عناصر این بردار، میزان برآورده‌شدن یکی از اهداف زیر را بیان می‌کنند:

۱) **کمینه‌کردن خطای طبقه‌بندی**: تابع زیر به منظور محاسبه خطای طبقه‌بندی استفاده می‌شود.

$$Fitness(C)=1-\frac{\#Hits}{N} \quad (۱۸)$$

که $\#Hits$ تعداد ف-نمونه‌داده‌هایی است که به درستی طبقه‌بندی شده‌است و N تعداد کل ف-نمونه‌داده‌ها می‌باشد.

۲) **کمینه‌کردن پیچیدگی:** پیچیدگی با تعداد ف-قوانین انتخابی نشان داده می‌شود.

- **رمزگذاری**

در این الگوریتم، از یک رمزگذاری دوبخشی به فرم $C^p=C_S^pC_T^p$ برای انتخاب (C_S) و تنظیم (C_T) ف-قوانین استفاده می‌شود. C^p، کروموزوم p ام در جمعیت است. بخش $C_S^p=(c_{S1},...,c_{Sm})$ با یک رشته باینری از m ژن رمزگذاری می‌شود به‌طوری‌که m تعداد ف-قوانین اولیه است. اگر یک ف-قانون انتخاب شود، ژن مربوط به آن یک می‌گردد، در غیر این صورت صفر است. بخش C_T^p به‌صورت رشته‌ای از اعداد حقیقی تعریف می‌شود که پارامترهای توابع عضویت مثلثی مشخصه‌ها را رمزگذاری می‌نماید:

www.SID.ir

Archive of SID

$$C_T^p = C_1 C_2 ... C_n$$
$$C_i = (a_1^i, b_1^i, c_1^i, ..., a_{m_i}^i, b_{m_i}^i, c_{m_i}^i) \quad ,i = 1,...n \qquad (۱۹)$$

که m_i، تعداد برچسب‌های (توابع عضویت) هر یک از n مشخصه است. هر تابع عضویت مثلثی به‌صورت $MF_j = (a_j, b_j, c_j)$ نشان داده می‌شود. هر یک از پارامترهای MF تنها مقادیری از بازه‌های زیر را می‌توانند بگیرند.

$$[I_{a_j}^l, I_{a_j}^r] = [a_j - (b_j - a_j)/2,\ a_j + (b_j - a_j)/2]$$
$$[I_{b_j}^l, I_{b_j}^r] = [b_j - (b_j - a_j)/2,\ b_j + (c_j - b_j)/2] \qquad (۲۰)$$
$$[I_{c_j}^l, I_{c_j}^r] = [c_j - (c_j - b_j)/2,\ c_j + (c_j - b_j)/2]$$

اولین فرد[۲۰] از اولین جمعیت، پایگاه قوانین مرحله قبل را رمزگذاری می‌کند. بقیه افراد اولین جمعیت، به‌صورت تصادفی با توجه به بازه‌های مربوطه تولید می‌شوند.

- **جهش و تبادل ژنی:**

هر فرزند با استفاده از عملگرهای جهش و تبادل ژنی به‌صورت زیر بدست می‌آید: ابتدا، بخش C_T فرزند با انجام تبادل ژنی ترکیبی[۲۱] [۳۶] بر روی بخش C_T والدین بدست می‌آید. سپس، بخش باینری C_S بر اساس بخش C_T والدین و فرزند تولید می‌شود. برای هر ژن در بخش C_S، مراحل زیر انجام می‌شود:

- هر ژن بخش C_S که نشان‌دهنده توابع عضویت قانون است، برای والدین و فرزند در نظر گرفته می‌شود. توابع عضویت این سه قانون استخراج می‌شود.
- بین ف-قانون فرزند و هر ف-قانون والد، فاصله‌های اقلیدسی نرمال شده با توجه به نقاط مرکزی توابع عضویت این قوانین بدست می‌آید. قانون والدی که به قانون فرزند نزدیک‌تر است، انتخاب می‌شود و مقدارش در بخش C_S فرزند کپی می‌شود.

این عملیات تکرار می‌شود تا زمانی که تمامی ژن‌ها در بخش C_S فرزند مقدار بگیرند. در هر مرحله، چهار فرزند تولید می‌شود. بعد از عمل جهش، تنها دو تا از بهترین فرزندان انتخاب می‌گردند. در ضمن، نوع عملگر تبادل ژنی به‌گونه‌ای است که از تولید مجدد قانون حذف‌شده جلوگیری می‌کند. استفاده از این عملگرها مزیت‌هایی دارد: تبادل ژنی بین افراد با قوانین بسیار متنوع، به الگوریتم اجازه می‌دهد که قسمت‌های متنوعی از فضای جستجو را کشف کند. عملگر جهش نیز با حذف قوانین اضافی، به بهبود استخراج قوانین کمک می‌نماید. به‌منظور نشان دادن روند کلی طبقه‌بند *f*-FARC-HD، شبه‌کد این طبقه‌بند در شکل ۴ ارائه گردیده است.

Algorithm LinguisticValidities (CBook)
Input: CBook: Mendel Code book;
Output: $\tilde{v}$: list of linguistic validities of items;
$\tilde{v} \leftarrow \{\ \}$; //list of items' validities
1. $\tilde{\Upsilon}_r \leftarrow (\tilde{\Upsilon}_{r1}, ..., \tilde{\Upsilon}_{r\sigma})$, $\tilde{\Upsilon}_r \subset CBook$ // linguistic validities indicated by σ experts for item r.
2. $\tilde{w} \leftarrow (\tilde{w}_1, ..., \tilde{w}_\sigma)$, $\tilde{w} \subset CBook$ // linguistic weight of experts.
3. For each item r, r= (1,.., m)
$\tilde{v}_r = \sum_{i=1}^{\sigma} \tilde{\Upsilon}_{ri}\tilde{\omega}_i \Big/ \sum_{i=1}^{\sigma} \tilde{\omega}_i$ // Compute LWA
End for.
End Algorithm.

Algorithm *f*-CreatClassifier (D, N, m, C, $\tilde{v}$, *minsup*, *minconf*)
Input: D:a dataset with N samples, m item, and C class.
$\tilde{v}$: list of items' validities.
minsup : predefined minimum support.
minconf : predefined minimum confidence.
Output: an *f*-Fuzzy Associative Classifier.

f-RB $\leftarrow \{\ \}$; // Rule base of f-classfier
1. For each class in D do
- Generate a tree of frequent patterns using Appriori, such that $f\text{-}\Pi_{A_{ij}}(x_p)$, $f\text{-}supp_{Rj}$ are defined according to the relations (7) and (5), respectively.
- Generate f-rules from the tree in the following form:
f-R$_j$: if $(x_{j1}, \tilde{v}_1)$ is $\tilde{A}_{j1}$ and... $(x_{jr}, \tilde{v}_r)$ is $\tilde{A}_{jr}$ and... $(x_{jm}, \tilde{v}_m)$ is $\tilde{A}_{jm}$ then class= (C_j, V_{Rj}) with $f\text{-}RW_{Rj}$.
where
$RW_{f\text{-}R_j} = \sum f\text{-}\Pi_{A_j}(x_p) \Big/ \sum f\text{-}\Pi_{A_j}(x_p)$
$V_{Rj} \leftarrow \text{avg}(\xi^{\tilde{w}}(\tilde{v}_1), ..., \xi^{\tilde{w}}(\tilde{v}_r), ..., \xi^{\tilde{w}}(\tilde{v}_m))$.
End for.
2. *f-RB* $\leftarrow$ Prescreening *(f-RB)*; //Select the *f*-rules with maximum covering of patterns
3. *f-RB* $\leftarrow$ Lateral tuning and selection *(f-RB)*;
End Algorithm.

Algorithm *f*-Classification (*f*-RB, $\hat{x}$)
Input: *f*-RB: Rule base of *f*-classifier;
$\hat{x}$: A pattern to be classified;
Output: class of pattern;
for each class C_p, $C_p \in \{C_1, C_2, ..., C_k\}$
$\Gamma_{Cp}(\hat{x}) = \sum_{f\text{-}Rj \in f\text{-}RB, Cj = Cp} f\text{-}\Pi_{A_j}(\hat{x}) \cdot f\text{-}RW_{R_j} \cdot V_{R_j}$;
End for;
Class($\hat{x}$)= the class that has maximum $\Gamma(\hat{x})$;
End Algorithm.

شکل ۴: شبه کد ف-طبقه‌بند *f*-FARC-HD

۶- آزمایش‌ها

ارزیابی روش پیشنهادی در دو بخش صورت می‌گیرد. در بخش اول، کارایی طبقه‌بند *f*-FARC-HD با FARC-HD و تعدادی از طبقه‌بندهای دیگر- طبقه‌بندهای همباشی و غیر همباشی- مقایسه شده است. این طبقه‌بندها عبارت‌اند از: طبقه‌بندهای همباشی کلاسیک مانند CBA [۳۵]،CBA2 [۳۶] و CMAR [۳۷]، طبقه‌بند فازی همباشی FPAR [۳۸]، درخت تصمیم C45 [۳۹] و طبقه‌بند مبتنی بر

www.SID.ir

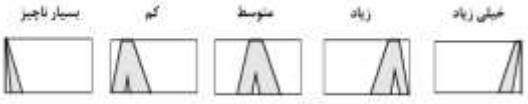
Archive of SID

جدول ۲: مقایسه طبقه‌بندها

طبقه‌بند	داده‌های یادگیری					داده‌های آزمایش					تعداد قوانین	طول قوانین
	صحت	دقت		بازخوانی		صحت	دقت		بازخوانی			
		OK	Dead	OK	Dead		OK	Dead	OK	Dead		
CBA	۰/۹۸۰	۰/۹۸۷	۰/۹۱۶	۰/۹۶۷	۰/۹۶۸	۰/۹۶۷	۰/۹۸۳	۰/۸۸۹	۰/۹۵۶	۰/۹۵۵	۳۸/۳	۴/۵۱۲
CBA2	۰/۹۷۴	۰/۹۹۱	۰/۹۱۰	۰/۹۶۴	۰/۹۷۷	۰/۹۶۴	۰/۹۸۵	۰/۸۸۸	۰/۹۵۶	۰/۹۶۱	۵۶	۴/۸۵۷
CMAR	۰/۹۱۸	۰/۹۹۴	۰/۷۸۱	۰/۹۲۳	۰/۹۸۱	۰/۹۲۳	۰/۹۱۹	۰/۷۷۹	۰/۹۲۱	۰/۹۷۷	۱۲۳/۵	۳/۰۶۳
FURIA	۰/۹۵۶	۰/۹۶۵	۰/۹۳۳	۰/۹۷۴	۰/۹۳۷	۰/۹۷۴	۰/۹۷۲	۰/۹۳۵	۰/۹۷۶	۰/۹۵۹	۱۰۵	۳/۸۸۸
FPAR	۰/۹۷۲	۰/۹۷۵	۰/۹۸۸	۰/۹۴۶	۰/۹۷۸	۰/۹۴۶	۰/۹۷۵	۰/۹۲۲	۰/۹۶۹	۰/۹۴۸	۲۸۳/۵	۳/۱۵۳
C45	۰/۹۵۷	۰/۹۷۵	۰/۹۳۷	۰/۹۷۴	۰/۹۳۷	۰/۹۷۴	۰/۹۷۵	۰/۹۳۷	۰/۹۷۴	۰/۹۳۷	۱۱۰۶	۳/۵۴۱
FARC-HD	۰/۹۶۱	۰/۹۸۸	۰/۹۳۶	۰/۹۷۵	۰/۹۷۰	۰/۹۷۵	۰/۹۸۱	۰/۹۱۲	۰/۹۶۵	۰/۹۵۱	۵۸/۲	۲/۴۲۸
f-FARC-HD	۰/۹۵۳	۰/۹۸۰	۰/۹۱۲	۰/۹۶۵	۰/۹۵۰	۰/۹۶۵	۰/۹۷۶	۰/۸۹۸	۰/۹۵۹	۰/۹۳۷	۱۶/۴	۱/۶۹

قوانین فازی FURIA [۴۰]. در بخش دوم، ازآنجاکه عدم قطعیت، جزئی لاینفک از دنیای واقعی است، عملکرد طبقه‌بند *f*-FARC-HD با FARC-HD در سطوح اغتشاش مختلف باهم مقایسه می‌شود.

آزمایش‌ها بر روی یک کامپیوتر شخصی با مشخصات پردازشگر Intel(R)Core(TM)2DuoT7250(2GHz) و 4GB RAM و بر روی مجموعه داده واقعی BURN اجراشده است. این مجموعه‌داده، از اطلاعات ۴۵۵۶ بیمار تشکیل‌شده است که در بازه زمانی یک سال، از بخش‌های سوختگی بیمارستان‌های اهواز جمع‌آوری‌شده است. اطلاعات هر بیمار شامل ده مشخصه سن، جنسیت، درصد سوختگی، درجه سوختگی، علت سوختگی، نتیجه آزمایش خون، نتیجه آزمایش ادرار و زخم، تعداد روزهای اقامت در بیمارستان و تاریخ بستری می‌باشد و وضعیت هر بیمار در پایان معالجات نیز با یکی از دو برچسب طبقه «*بهبودیافته*» و «*فوت‌شده*» مشخص‌شده است. ۷۱ درصد نمونه داده‌ها در طبقه «بهبودیافته» و ۲۹ درصد در طبقه «*فوت‌شده*» می‌باشند.



شکل ۵. کلمات استفاده شده در پرسشنامه

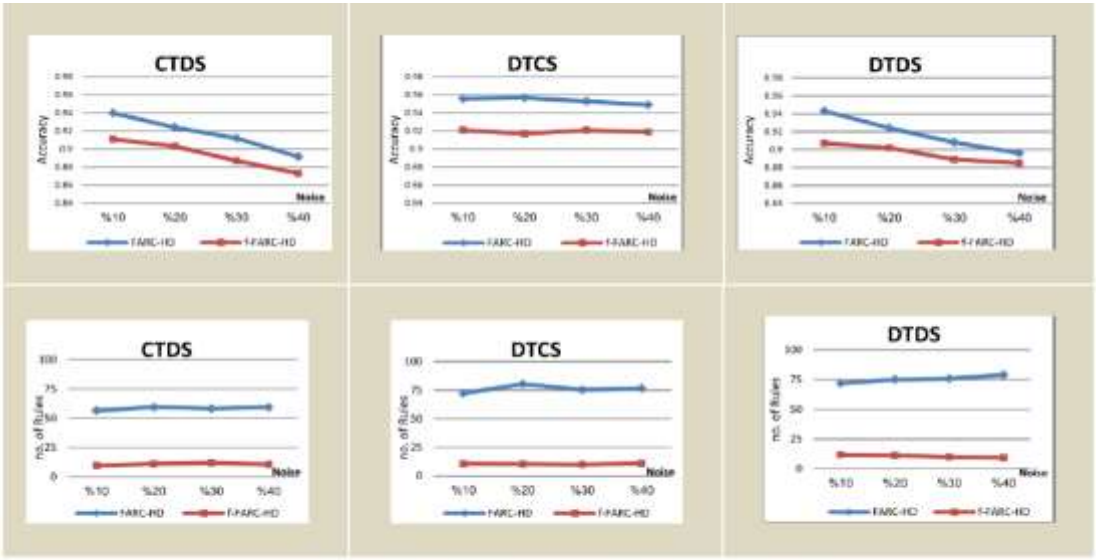
به‌منظور تعیین میزان اعتبار هر مشخصه، ابتدا کلماتی از کتابچه رمز مندل انتخاب گردید که تمامی فضای دامنه را پوشش دهند. در شکل ۵، این کلمات نشان داده‌شده است. سپس، با استفاده از یک پرسش‌نامه، از ۲۰ نفر از پزشکان عمومی، متخصصین سوختگی، دانشجویان رزیدنت و پرستاران شاغل در بخش سوختگی باتجربه و سابقه کارهای متفاوت (از ۱ تا ۱۵ سال) خواسته شد که نظر خود را درباره میزان اهمیت هر یک از مشخصه‌ها با استفاده از کلمات تعیین‌شده، بیان نمایند. درنهایت، با روشی که در بخش ۳-۱ توضیح داده شد، اعتبار هر یک از مشخصه‌ها به دست آمد.

به‌منظور حصول نتایج مطمئن‌تر، از روش آماری *ارزیابی* استفاده‌شده است. در این شیوه، مجموعه‌داده به‌صورت تصادفی به ۱۰ قسمت تقسیم می‌شود و ۱۰ بار الگوریتم اجرا می‌شود. در هر بار اجرای الگوریتم، یکی از قسمت‌ها به‌عنوان داده آزمایش و ۹ قسمت دیگر به‌عنوان داده یادگیری مورداستفاده قرار می‌گیرد. در این روش، مطمئن می‌شویم که هر قسمت داده، یک‌بار به‌عنوان داده آزمایش در نظر گرفته شود. درنهایت میانگین نتایج این ۱۰ بار اجرا به‌عنوان نتیجه نهایی لحاظ می‌گردد.

در بخش اول، کارایی سیستم پیشنهادی ازنظر متوسط تعداد قوانین، متوسط طول قوانین، صحت طبقه‌بندی داده‌های یادگیری و داده‌های آزمایش، میزان دقت و بازخوانی[۲۳] با طبقه‌بندهای دیگر مقایسه می‌شود. نتایج در جدول ۲ نشان داده‌شده است و بهترین نتایج در هر ستون، تیره شده‌اند. همان‌طور که در جدول مشاهده می‌شود، ازنظر صحت، دقت و بازخوانی تفاوت چندان محسوسی بین *f*-FARC-HD و سایر طبقه‌بندها مشاهده نمی‌شود. تفاوت در حد صدم واحد است. این موضوع هم در مورد داده‌های یادگیری و هم در داده‌های آزمایش صدق می‌کند. اما تفاوت در میزان پیچیدگی طبقه‌بند *f*-FARC-HD و بقیه طبقه‌بندها کاملاً واضح است. به‌عنوان نمونه، بیشترین میزان صحت در داده‌های آزمایش، متعلق به طبقه‌بند FPAR با مقدار ۰/۹۷۵ است. اما همان‌طور که در جدول دیده می‌شود، تعداد قوانین در این طبقه‌بند ۲۸۳/۵ با متوسط طول ۳/۱۵۳ می‌باشد. پیچیدگی این طبقه‌بند بسیار بالا است. اما طبقه‌بند *f*-FARC-HD با ۴/ ۱۶ قانون که طول متوسط آن‌ها ۱/۶۹ است، دارای میزان صحت ۰/۹۶۱می‌باشد. با توجه به تفاوت فاحشی که در پیچیدگی این دو طبقه‌بند وجود دارد، این مقدار کاهش صحت، منطقی است. چراکه همواره ناگزیر به پذیرفتن تعدیلی[۲۴] بین پیچیدگی و میزان صحت یک طبقه‌بند هستیم.

www.SID.ir

Archive of SID



شکل ۶- مقایسه کارایی دو طبقه‌بند از نظر صحت و تعداد قوانین در سطوح نویز متفاوت

در بخش دوم، به بررسی تأثیر اغتشاش در کارایی دو طبقه‌بند FARC-HD و *f*-FARC-HD می‌پردازیم. اغتشاشی که در اینجا بررسی می‌شود، اغتشاش مربوط به مشخصه‌هاست و به‌صورت دستی به روشی که در [۴۱] ارائه‌شده است، ایجاد می‌گردد. به‌منظور ایجاد اغتشاشی در سطح x، به $x \times ٪۱۰۰$ مقادیر هر مشخصه، یک مقدار تصادفی داده می شود. این مقدار تصادفی، در مشخصه‌های عددی، مقداری بین کمینه و بیشینه مشخصه است. در این روش، فرض بر این است که احتمال این‌که مشخصه‌ای نویز داشته باشد، هیچ ارتباطی با سایر مشخصه‌ها ندارد. سطوح اغتشاش در نظر گرفته‌شده در این تحقیق عبارت‌اند از ۱۰٪، ۲۰٪، ۳۰٪ و ۴۰٪ که در سه حالت بر روی مجموعه‌داده BURN اعمال‌شده است:

الف) اغتشاش تنها بر مجموعه‌داده آزمایش اعمال می‌شود و مجموعه‌داده یادگیری بدون اغتشاش می‌ماند (DTC).

ب) اغتشاش تنها بر مجموعه‌داده یادگیری اعمال می‌شود و مجموعه‌داده آموزش بدون اغتشاش می‌ماند (DTC).

ج) اغتشاش بر هر دو مجموعه‌داده آزمایش و یادگیری اعمال می‌شود (DTDS).

الگوریتم‌های FARC-HD و *f*-FARC-HD در سه وضعیت بالا به روش ارزیابی ده‌قسمتی اجراشده و متوسط صحت طبقه‌بندی، متوسط تعداد قوانین و متوسط طول آن‌ها در شکل‌های ۶ و ۷ نشان داده‌شده است. موارد زیر را در نمودارها می‌توان مشاهده نمود:

- بیشترین مقدار کاهش صحت در هر دو طبقه‌بند در حالت CTDS و کمترین مقدار کاهش در DTCS می‌باشد. روند کاهش صحت در طبقه‌بند *f*-FARC-HD در هر دو حالت DTCS و DTDS کمتر از FARC-HD است. مخصوصاً در DTCS، میزان صحت *f*-FARC-HD با افزایش اغتشاش، تقریباً ثابت مانده است.
- تعداد قوانین طبقه‌بند *f*-FARC-HD در هر سه حالت CTDS، DTCS و DTDS تقریباً ثابت مانده است. اما در طبقه‌بند FARC-HD در هر دو حالت DTCS و DTDS تعداد قوانین افزایش‌یافته است.
- طول قوانین در هر دو طبقه‌بند با افزایش اغتشاش، تغییر چندانی نداشته است.

به‌طور خلاصه، می‌توان گفت که بهترین وضعیت برای هر دو طبقه‌بند، حالتی است که داده‌های یادگیری حاوی اغتشاش باشند اما در داده‌های آزمایش، اغتشاشی وجود نداشته که همان وضعیتDTCS است. شاید بدین علت که در هنگام آموزش، علیرغم وجود اغتشاش، بخشی از واقعیت در قوانین لحاظ می‌گردد و در هنگام طبقه‌بندی داده‌های عاری از اغتشاش، به استنتاج صحیح منجر می‌گردد. در ضمن، به نظر می‌رسد حساسیت طبقه‌بند *f*-FARC-HD نسبت به اغتشاش در مقایسه با FARC-HD کمتر می‌باشد. زیرا علاوه بر این‌که روند کاهش صحت با افزایش اغتشاش، کندتر است، تعداد قوانین نیز در *f*-FARC-HD تقریباً ثابت مانده است. بنابراین، می‌توان گفت که *f*-FARC-HD، طبقه‌بندی مقاوم‌تر نسبت به تغییرات اغتشاش است.

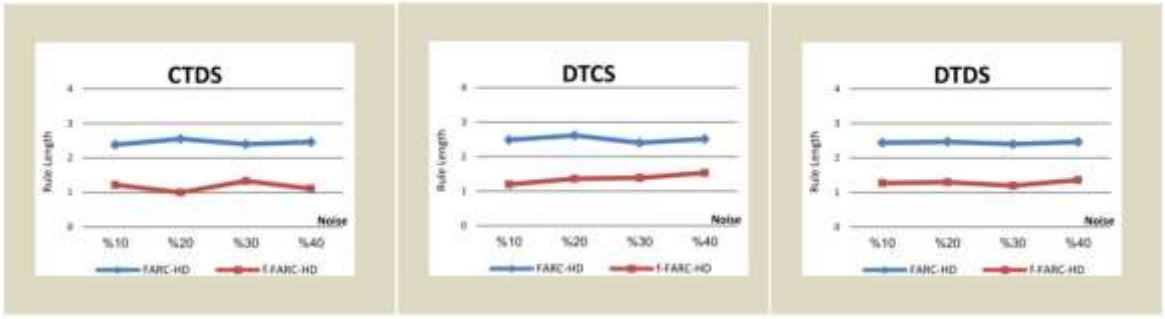
۷- نتیجه

در این مقاله، به معرفی یک دیدگاه جدید با به‌کارگیری مفهوم اعتبار در طراحی الگوریتم‌های طبقه‌بند می‌پردازیم. در این دیدگاه، هر یک از مشخصه‌ها و قوانین در طبقه‌بند، تا درجه‌ای معتبر می‌باشند و این میزان اعتبار، مستقیم یا غیرمستقیم، توسط کارشناسان تعیین می‌گردد. بر این اساس، ساختار کلی یک ف-طبقه‌بند فازی همباشی در چارچوب منطق فازی توسعه‌یافته ارائه می‌شود.

جهت آزمودن این ایده جدید، یک الگوریتم طبقه‌بند فازی همباشی معتبر (FARC-HD) با توجه به مفهوم اعتبار، توسعه داده‌شده است (*f*-FARC-HD). در این طبقه‌بند، اعتبار مشخصه‌ها از تلفیق اعتبار مشخصه‌ها در مقدم آن قانون تعریف‌شده است. ارزیابی در

www.SID.ir

Archive of SID



شکل ۷:- مقایسه طول قوانین در سطوح نویز متفاوت

دومرحله انجام می‌شود. در مرحله اول، کارایی طبقه‌بند *f*-FARC-HD با تعدادی از طبقه‌بندهای دیگر ازنظر صحت طبقه‌بندی، میزان دقت، بازخوانی و پیچیدگی مقایسه می‌شود. در مرحله دوم، دو طبقه‌بند *f*-FARC-HD و FARC-HD با در نظر گرفتن درجه‌های مختلفی از اغتشاش باهم مقایسه می‌شوند. لازم به ذکر است که تمامی آزمایش‌ها بر روی یک مجموعه‌داده واقعی (BURN) از اطلاعات بیماران سوختگی بیمارستان‌های اهواز اجراشده است.

نتیجه آزمایش‌ها بر روی مجموعه‌داده BURN نشان می‌دهد که با در نظر گرفتن مفهوم اعتبار در *f*-FARC-HD، طبقه‌بندی کارا با پیچیدگی بسیار کم‌تر به دست می‌آید که برای کاربران بسیار خواناتر و قابل‌فهم‌تر است. در ضمن، در شرایطی که اغتشاش وجود دارد، به نظر می‌رسد که حساسیت طبقه‌بند *f*-FARC-HD نسبت به اغتشاش کمتر است. بدین معنا که علاوه بر ثابت ماندن تقریبی پیچیدگی طبقه‌بند، روند کاهش صحت طبقه‌بند نیز کندتر می‌باشد.

در این مقاله، مجبور شدیم به اجرای ایده خود بر روی به یک مجموعه‌داده اکتفا کنیم. چراکه فرآیند تهیه اعتبار مشخصه‌ها با استفاده از پرسش‌نامه، فرآیند زمان‌بری می‌باشد. اما با توجه به پیشرفت‌هایی که مخصوصاً در حوزه شبکه‌های اجتماعی صورت پذیرفته، امیدواریم در کارهای آتی، بتوانیم عملکرد ف-طبقه‌بند پیشنهادی را بر روی مجموعه‌داده‌های بیشتر، مخصوصاً با ابعاد بالا بررسی نماییم.

مراجع

[1] M. Antonelli, D. Bernardo, H. Hagras, and F. Marcelloni, "Multiobjective evolutionary optimization of type-2 Fuzzy rule-based systems for financial data classification", IEEE Transactions on Fuzzy Systems, vol. 25, no. 2, pp. 249–264, 2017.

[2] A. Ferranti, F. Marcelloni, A. Segatori, M. Antonelli, and P. Ducange, "A distributed approach to multi-objective evolutionary generation of fuzzy rule-based classifiers from big data ", Information Sciences, vol. 415, pp. 319–340, 2017.

[3] M. I. Rey, M. Galende, M. J. Fuente, and G. I. Sainz-Palmero, "Multi-objective based fuzzy rule based systems (FRBSs) for trade-off improvement in accuracy and interpretability: A rule relevance point of view", Knowledge-Based Systems, vol. 127, pp. 67–84, 2017.

[4] R. Agrawal, R. Srikant, and others, "Fast algorithms for mining association rules", in Proceeding of VLDB'94, 1994, vol. 1215, pp. 487–499.

[5] M. Kadkhoda, M.-R. Akbarzadeh-T, and S. M. Taheri, "Mining fuzzy temporal itemsete within various time intevals in

[6] quantitative dataset", Iranian Journal of Fuzzy Systems, vol. 13, no. 7, 2016.

[7] Z. Davarzani, M. Kadkhoda, and M. R. Akbarzadeh Totonchi, "A GA-based fuzzy mining algorithm using selections pressure technique for extracting membership functions", in 3rp joint congress on fuzzy and intelligent systems, 2009.

[8] Z. Zhang, W. Pedrycz, and J. Huang, "Efficient mining product-based fuzzy association rules through central limit theorem', Applied Soft Computing, vol. 63, pp. 235–248, 2018.

[9] Z. Chen and G. Chen, "Building an associative classifier based on fuzzy Association rules", International Journal of Computational Intelligence Systems, vol. 1, no. 3, pp. 262–273, 2008.

[10] S. Saraswathi and N. Kannan, "A hybrid associative classification model for software development effort estimation", Circuits and Systems, vol. 7, no. 06, pp. 824–834, 2016.

[11] V. G. Giannoglou, D. G. Stavrakoudis, J. B. Theocharis, and V. Petridis, "Genetic fuzzy rule based classification systems for coronary plaque characterization based on intravascular ultrasound images", Engineering Applications of Artificial Intelligence, vol. 38, pp. 203–220, 2015.

[12] J. P. Lucas, A. Laurent, M. N. Moreno, and M. Teisseire, "A fuzzy associative classification approach for recommender systems", International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, vol. 20, no. 04, pp. 579–617, 2012.

[13] J. A. Sanz, A. Fernández, H. Bustince, and F. Herrera, "IVTURS: A linguistic fuzzy rule-based classification System Based On a new interval-valued fuzzy reasoning method with tuning and rule selection.", IEEE Transaction on Fuzzy Systems, vol. 21, no. 3, pp. 399–411, 2013.

[14] Y. Ma, G. Chen, and Q. Wei, "A novel business analytics approach and case study – fuzzy associative classifier based on information gain and rule-covering", Journal of Management Analytics, vol. 1, no. 1, pp. 1–19, 2014.

[15] J. Sanz, D. Bernardo, F. Herrera, H. Bustince Sola, and H. Hagras, "A compact evolutionary interval-valued fuzzy rule-based classification system for the modeling and prediction of real-world financial applications with imbalanced data", IEEE Transactions on Fuzzy Systems, vol. 23, no. 4, pp. 973–990, 2015.

[16] R. Alcalá, J. Alcalá-Fdez, and F. Herrera, "A proposal for the genetic lateral tuning of linguistic fuzzy systems and its interaction with rule selection", IEEE Transactions on Fuzzy Systems, vol. 15, no. 4, pp. 616–635, 2007.

[17] J. Alcala-Fdez, R. Alcala, and F. Herrera, "A fuzzy association rule-based classification model for high-dimensional problems with genetic rule selection and lateral tuning", IEEE Transactions on Fuzzy Systems, vol. 19, no. 5, pp. 857–872, 2011.

[18] M. Fazzolari, R. Alcala, Y. Nojima, H. Ishibuchi, and F. Herrera, "Improving a fuzzy association rule-based classification model by granularity learning based on heuristic measures over multiple granularities", IEEE International Workshop on in Genetic and Evolutionary Fuzzy Systems (GEFS), pp. 44–51, 2013.

[19] M. Fazzolari, R. Alcalá, and F. Herrera, "A multi-objective evolutionary method for learning granularities based on fuzzy discretization to improve the accuracy-complexity trade-off of fuzzy rule-based classification systems", Applied Soft Computing, vol. 24, pp. 470–481, 2014.

[20] L. A. Zadeh, "Toward extended fuzzy logic—A first step", Fuzzy Sets and Systems, vol. 160, no. 21, pp. 3175–3181, 2009.

www.SID.ir

Archive of SID

[۲۱] فرناز صباحی و محمد رضا اکبرزاده توتونچی، «شناسایی سیستم‌های غیرخطی بر اساس منطق فازی توسعه یافته»، مجله مهندسی برق دانشگاه تبریز، جلد ۴۴، شماره ۱، صفحه ۳۲-۲۳، ۱۳۹۳.

[22] F. Sabahi and M.-R. Akbarzadeh-T, "Introducing validity in fuzzy probability for judicial decision-making", International Journal of Approximate Reasoning, vol. 55, no. 6, pp. 1383–1403, 2014.

[23] V. A. Niskanen, "A concept map approach to approximate reasoning with fuzzy extended logic", in Fuzzy Technology, Springer, 2016, pp. 47–70.

[24] J. Surowiecki, "The wisdom of crowds. 2004 Anchor", Reprint edition (August 16, 2005), 2005.

[۲۵] الناز زعفرانی معطر, محمدرضا فیضی درخشی و آزاده روحانی, «تشخیص هوشمند و خودکار غلط‌های تایپی در پایگاه‌داده‌های بزرگ بدون استفاده از لغت‌نامه», مجله مهندسی برق دانشگاه تبریز، جلد ۴۷، شماره ۱، صفحه ۹۱-۸۱، ۱۳۹۶.

[۲۶] سید هادی حسینی, بابک نجار اعرابی, بهزاد مشیری و اشکان رحیمی کیان, «الگوریتم ترکیب فازی مدل‌های پیش بین جریان ترافیک در حضور داده‌های اغتشاشی», مجله مهندسی برق دانشگاه تبریز، جلد ۴۶، شماره ۱، صفحه ۱۳۲-۱۲۱، ۱۳۹۵.

[27] L. A. Zadeh, "Fuzzy logic and approximate reasoning", Synthese, vol. 30, no. 3, pp. 407–428, 1975.

[28] L. A. Zadeh, "From fuzzy logic to extended fuzzy logic-A first step", in NAFIPS 2009-2009 Annual Meeting of the North American Fuzzy Information Processing Society, 2009, pp. 1–2.

[29] F. Sabahi and M. R. Akbarzadeh-T, "Extended fuzzy logic: sets and systems", IEEE Transactions on Fuzzy Systems, vol. 24, no. 3, pp. 530–543, 2016.

[30] F. Sabahi and M.-R. Akbarzadeh-T, "A qualified description of extended fuzzy logic", Information Sciences, vol. 244, pp. 60–74, 2013.

[31] F. Sabahi and M. Akbarzadeh Tootoonchi, "A framework for analysis of extended fuzzy logic", Journal of Zhejiang University Science, vol. 15, no. 17, pp. 584–591, 2014.

[32] F. Liu and J. M. Mendel, "Encoding words into interval type-2 fuzzy sets using an interval approach", IEEE Transactions on Fuzzy Systems, vol. 16, no. 6, pp. 1503–1521, 2008.

[33] D. Wu, J. M. Mendel, and S. Coupland, "Enhanced interval approach for encoding words into interval type-2 fuzzy sets and its convergence analysis", Fuzzy Systems, IEEE Transactions on, vol. 20, no. 3, pp. 499–513, 2012.

[34] J. Mendel and D. Wu, *Perceptual Computing: Aiding People in Making Subjective Judgments*, vol. 13. John Wiley & Sons, 2010.

[35] B. Kavšek and N. Lavrač, "APRIORI-SD: Adapting association rule learning to subgroup discovery", Applied Artificial Intelligence, vol. 20, no. 7, pp. 543–583, 2006.

[36] B. Liu, W. Hsu and Y. Ma, "Integrating classification and association rule mining", in Proceeding of KDD"98, New York, NY, 1998, pp. 80–86.

[37] B. Liu, Y. Ma, and C.-K. Wong, "Classification using association rules: weaknesses and enhancements", in Data mining for scientific and engineering applications, Springer, pp. 591–605, 2001.

[38] W. Li, J. Han, and J. Pei, "CMAR: Accurate and efficient classification based on multiple class-association rules", in Proceeding of ICDM'01, pp. 369–376, 2001.

[39] M. Nandhini, M. Rajalakshmi, and S. N. Sivanandam, "Experimental and statistical analysis on the performance of firefly based predictive association rule classifier for health care data diagnosis", Journal of Control Engineering and Applied Informatics, vol. 19, no. 2, pp. 101–110, 2017.

[40] J. R. Quinlan, *C4. 5: programs for machine learning*. Elsevier, 2014.

[41] J. Hühn and E. Hüllermeier, "FURIA: An algorithm for unordered fuzzy rule induction", Data Mining and Knowledge Discovery, vol. 19, no. 3, pp. 293–319, 2009.

[42] X. Zhu and X. Wu, "Class noise vs. attribute noise: a quantitative study", Artificial Intelligence Review, vol. 22, no. 3, pp. 177–210, 2004.

زیرنویس‌ها

[1] Fuzzy Associative Classifier (FAC)
[2] fuzzy association rule
[3] lateral tuning
[4] validity
[5] Extended Fuzzy Logic (FLe)
[6] Wisdom of crowds
[7] *f*-transformation
[8] noise
[9] Interval type2 fuzzy set (IT2 FS)
[10] Footprint Of Uncertainty
[11] Codebook
[12] interval end point data
[13] linguistic weighted average (LWA)
[14] support
[15] confidence
[16] the maximum association degree
[17] the weighted vote
[18] matching degree
[19] *f*-frequent patterns
[20] individual
[21] Blend crossover
[22] 10 fold cross validation
[23] recall
[24] trade off

www.SID.ir