

روش یادگیری دسته بند فازی Emmff و بکارگیری آن در داده‌های پزشکی

سعید جلیلی، فاطمه فرجی دانشگر

f_daneshgar@modares.ac.ir, sjalili@modares.ac.ir

دانشگاه تربیت مدرس، دانشکده فنی، بخش برق، آزمایشگاه یادگیری نمادین ماشین

چکیده

سیستم‌های خبره مبتنی بر قانون اغلب برای پشتیبانی تصمیم‌گیری در دامنه‌های مختلفی مانند تشخیص خطا^۱، زیست شناسی^۲، و پزشکی به کار گرفته می‌شوند. در بعضی از دامنه‌ها مانند پزشکی، ترجیح داده می‌شود از روش‌های دسته بندی به صورت جعبه سیاه (مانند شبکه عصبی) استفاده نشود تا کاربر قادر به درک دانش دسته‌بند باشد. دسته‌بندهای مبتنی بر قانون فازی برای این منظور بسیار مناسب هستند. زیرا آنها حاوی قوانین تفسیرپذیر زبانی ساده‌ای هستند و بعضی محدودیت‌های دسته‌بندهای قطعی^۳ را ندارند. دسته بندها اغلب باید طی یک فرآیند یادگیری از روی داده‌ها ایجاد شوند، زیرا دانش متخصص در بسیاری از دامنه‌ها محدود است و نمی‌توان تمام پارامترهای سیستم را بوسیله دانش متخصص مشخص کرد.

یکی از کاربردهای داده کاوی در پزشکی، مسئله تشخیص سرطان است. تا کنون از روش‌های مختلفی برای این مسئله استفاده شده است که به طور کلی می‌توان به SVM، دسته‌بندی فازی، شبکه عصبی و الگوریتم ژنتیک اشاره کرد.

MMFF^۴ یکی از روش‌های دسته‌بندی فازی است که طی سه مرحله: (۱) انتخاب خصیصه (۲) تولید توابع عضویت و (۳) تولید جدول تصمیم‌گیری، قوانین فازی را از نمونه‌های آموزشی استخراج می‌کند. این روش از یک سو الگوریتم بسیار ساده‌ای دارد که می‌تواند برای کاربردهای پزشکی مناسب باشد و از سوی دیگر نتایج خوبی از لحاظ دقت دسته‌بندی بین روش‌های مشابه دارد.

مجموعه داده Wisconsin Breast Cancer (WBC) شامل ۶۹۹ نمونه است که هر نمونه با ۹ خصیصه توصیف شده است. پس از آزمایشات مختلف روی این مجموعه داده با روش MMFF، مشاهده شد که نتایج پایین‌تر از حد انتظار است. بنابراین روش را بررسی کرده و دریافتیم که این الگوریتم در قسمت تولید قوانین فازی دچار ضعف است، به گونه‌ای که در مجموعه داده‌های بزرگ منجر به کاهش دقت دسته‌بندی می‌شود. بنابراین ما در این مقاله روش MMFF را توسعه دادیم، به نحوی که مشکل مطرح شده مرتفع گردد. ما این روش را Emmff نامگذاری کردیم.

آزمایش روش Emmff با مجموعه داده WBC نشان داد که دسته‌بند فازی تولید شده به روش Emmff، نسبت به دسته‌بند روش MMFF و همچنین نسبت به سایر روش‌های موجود در تشخیص سرطان، قادر به دسته‌بندی و تشخیص دقیق‌تری است.

کلمات کلیدی - یادگیری قوانین فازی، یادگیری ماشین، تشخیص سرطان، داده‌های پزشکی

¹ Fault Detection

² Biology

³ Crisp

⁴ Merging Membership Function First