Journal of Applied And Basic Machine Intelligence Research (ABMIR)

پژوهش‌های نظری و کاربردی هوش ماشینی

سال اول، شماره اول، بهار و تابستان ۱۴۰۲، صفحات ۱۳-۱

ارائه یک روش انتخاب ویژگی مبتنی بر انتگرال فازی در یادگیری چند برچسبی

امین هاشمی[۱]، محمدباقر دولتشاهی[۲*]

[۱] دانشجوی کارشناسی ارشد گروه مهندسی کامپیوتر، دانشکده فنی و مهندسی، دانشگاه لرستان، خرم‌آباد

[۲] استادیار گروه مهندسی کامپیوتر، دانشکده فنی و مهندسی، دانشگاه لرستان، خرم‌آباد

چکیده

مقاله پژوهشی

الگوریتم‌های یادگیری چند برچسبی به دلیل حجم و ابعاد بالای داده‌های چند برچسبی و همچنین وجود نویز در آن‌ها، با چالش‌های فراوانی مواجه هستند. انتخاب ویژگی یک تکنیک مؤثر برای برطرف کردن این چالش‌ها است. در این مقاله یک روش انتخاب ویژگی مبتنی بر یک رویکرد شورایی برای داده‌های چند برچسبی ارائه‌شده است. درروش پیشنهادی، سه ماتریس تصمیم مختلف بر اساس معیارهای ارزیابی ویژگی مختلف با درنظرگرفتن همگرایی ویژگی‌ها با برچسب‌های کلاس و همچنین افزونگی ویژگی‌ها نسبت به هم در فرایند انتخاب ویژگی مؤثر هستند. این سه ماتریس تصمیم درنهایت بر اساس یک رویکرد شورایی مبتنی بر مفهوم انتگرال فازی باهم ترکیب می‌شوند تا ارزیابی ویژگی‌ها بر اساس مقدار تجمیع شده صورت گیرد. برای ارزیابی عملکرد الگوریتم پیشنهادی، مقایساتی با چندین الگوریتم مشابه بر روی چند مجموعه داده مختلف صورت گرفته است. نتایج به‌دست‌آمده از آزمایش‌های انجام‌شده، نشان‌دهنده عملکرد مناسب الگوریتم پیشنهادی در مقایسه با سایر الگوریتم‌ها است.

تاریخ دریافت:

۱۳۹۹/۱۲/۱۵

تاریخ پذیرش:

۱۴۰۰/۰۲/۱۸

کلیدواژه‌ها:

انتخاب ویژگی، یادگیری چند برچسبی، انتگرال فازی، رویکرد شورایی

نویسنده مسئول:

dowlatshahi.mb@lu.ac.ir

: 10.22034/ABMIR.2022.2714

۱- مقدمه

با پیشرفت فناوری‌های جدید، اینترنت اشیا[1] و شبکه‌های اجتماعی، شاهد تولید حجم وسیعی از داده‌ها و با سرعت‌بالا در هرروز هستیم. این داده‌ها عموماً شامل نمونه‌های آموزشی و ویژگی‌های بسیاری هستند که اغلب ویژگی‌ها غیرمرتبط[2] و افزونه[3] بسیاری نیز در مجموعه ویژگی‌ها وجود دارد. ویژگی‌های غیرمرتبط در پیش‌بینی برچسب‌های کلاس مؤثر نیستند و حتی در این فرایند تأثیر منفی نیز می‌گذارند، درحالی‌که ویژگی‌های افزونه ویژگی‌هایی هستند که تقریباً ترکیبی از ویژگی‌های دیگر هستند و اطلاعات جدیدی را ارائه نمی‌دهند. وجود چنین ویژگی‌هایی در مجموعه داده‌ها، الگوریتم‌های یادگیری را زمان‌بر می‌کند و درعین‌حال بر عملکرد آن‌ها نیز تأثیر می‌گذارد [۱,۲].

انتخاب ویژگی[4] به‌عنوان یک روش مؤثر برای رفع این مشکلات استفاده می‌شود. این تکنیک کاهش ابعاد[5]، ویژگی‌های مفید را انتخاب می‌کند و درنهایت ویژگی‌های غیرمرتبط و افزونه را از مجموعه داده حذف می‌کند. این رویکرد هزینه محاسباتی و ذخیره‌سازی را کاهش می‌دهد، درحالی‌که دقت دسته‌بندی[6] را نیز افزایش می‌دهد [۲,۳]. در برخی مسائل دنیای واقعی، هر نمونه با مجموعه‌ای از برچسب‌های کلاس مطابقت دارد. این نوع مسائل به‌عنوان دسته‌بندی چند برچسبی[7] شناخته می‌شوند. همانند داده‌های تک برچسبی[8]، انتخاب ویژگی به دلیل وجود ویژگی‌های بسیار در اغلب مجموعه‌های داده، یک فرایند ضروری در دسته‌بندی چند برچسبی است. روش‌های انتخاب ویژگی چند برچسبی را می‌توان به دو گروه تبدیل دودویی[9] و انطباق الگوریتم[10] دسته‌بندی کرد. در تبدیل دودویی، مسئله انتخاب ویژگی چند برچسبی به چندین مسئله تک‌برچسبی مستقل نگاشت می‌شود. استراتژی انطباق الگوریتم روشی است که در آن الگوریتم مستقیماً به داده‌های چند برچسبی اعمال می‌شود[۴-۶]. همچنین می‌توان روش‌های انتخاب ویژگی را بر اساس نحوه تعامل با الگوریتم‌های یادگیری را در سه دسته کلی در نظر گرفت: فیلتر[11]، بسته‌بند[12] و تعبیه‌شده[13]. روش‌های فیلتر، تعامل با الگوریتم‌های یادگیری را نادیده می‌گیرند و بر اساس برخی معیارهای از پیش تعریف‌شده مانند ضریب همبستگی پیرسون[14]، بهره اطلاعاتی[15] و شباهت کسینوسی به هر ویژگی امتیازی اختصاص می‌دهند و فرایند انتخاب ویژگی صرفاً بر اساس این امتیازات صورت می‌گیرد. روش‌های بسته‌بند اهمیت یک ویژگی را در طول فرایند یادگیری با استفاده از یک دسته‌بند شناسایی می‌کنند. در این استراتژی زیرمجموعه‌های مختلف از ویژگی‌ها بر اساس یک دسته‌بند ارزیابی می‌شوند تا بهترین زیرمجموعه انتخاب شود. درنهایت، استراتژی‌های تعبیه‌شده دارای روش‌های داخلی انتخاب ویژگی هستند. در این روش‌ها، مشابه روش‌های بسته‌بند، ویژگی‌ها توسط یک الگوریتم یادگیری ارزیابی می‌شوند. بااین‌حال، در این روش‌ها الگوریتم یادگیری تنها یک‌بار اعمال می‌شود و زیرمجموعه ویژگی‌ها را به‌طور مکرر ارزیابی نمی‌کند. پس‌ازآن، ویژگی‌ها بر اساس یک رویکرد فیلتر ارزیابی می‌شوند؛ بنابراین، روش‌های تعبیه‌شده به دنبال رفع نقاط ضعف دو روش قبلی و استفاده از نقاط قوت هر دو هستند[۳,۷].

در این مقاله یک روش انتخاب ویژگی مبتنی بر رویکرد فیلتر برای یادگیری چند برچسبی و با استفاده از مفهوم انتگرال فازی ارائه‌شده است. عملگر سوگنو[16] که یکی از عملگرهای انتگرال فازی است به‌عنوان معیار اصلی برای انتخاب ویژگی در نظر گرفته‌شده است. درواقع هدف این مقاله، استفاده از چندین معیار برای انتخاب ویژگی برخلاف روش‌های کنونی انتخاب ویژگی چند برچسبی است. زیرا تنوع روش‌های مختلف و بررسی یک ویژگی بر اساس

[1] Internet of things
[2] Irrelevant
[3] Redundant
[4] Feature selection
[5] Dimensionality reduction
[6] Classification
[7] Multi-label
[8] Single label
[9] Binary transformation
[10] Algorithm adaptation
[11] Filter
[12] Wrapper
[13] Embedded
[14] Pearson correlation coefficient
[15] Information gain
[16] Sugeno

چندین معیار متنوع می‌تواند منجر به نتیجه بهتری نسبت به یک معیار شود.

انتخاب ویژگی شورایی[1] مبتنی بر ترکیب چندین معیار انتخاب ویژگی به‌جای یک روش واحد برای حل مسئله انتخاب ویژگی است و بر پایه ضرب‌المثل قدیمی "دو فکر بهتر از یک فکر است" استوار است. منطق مبتنی بر این ایده، ساخت مجموعه‌ای از فرضیه‌ها با استفاده از روش‌های مختلف است، و سپس آن‌ها باهم ترکیب می‌شوند تا نتایج بهتری نسبت به تنها یک روش انتخاب ویژگی به دست آید. این تنوع رویکردها و کنترل واریانس است که باعث موفقیت این رویکرد می‌شود [۸].

بنابراین در این رویکرد از عملگر انتگرال فازی سوگنو برای ترکیب معیارهای در نظر گرفته‌شده استفاده‌شده است و مقدار به‌دست‌آمده از این ترکیب به‌عنوان شاخص انتخاب ویژگی در نظر گرفته‌شده است. به این صورت که اگر این مقدار بالاتر از یک حد آستانه از پیش تعریف‌شده قرار گرفت، ویژگی انتخاب و در غیر این صورت صرف‌نظر می‌شود.

برای نشان‌دادن بهینه بودن و کارایی روش پیشنهادی، این رویکرد با برخی از تکنیک‌های مشابه در مجموعه داده‌های مختلف دنیای واقعی مقایسه شده است. نتایج نشان‌دهنده برتری روش پیشنهادی ازنظر معیارهای دسته‌بندی و زمان اجرا در برابر استراتژی‌های رقابتی است.

این مقاله به شرح زیر سازمان‌دهی شده است: در بخش ۲، برخی از روش‌های مرتبط درزمینه انتخاب ویژگی چند برچسبی بررسی خواهند شد. در بخش ۳، مفاهیم اساسی الگوریتم پیشنهادی معرفی می‌شوند، و سپس در بخش ٤، الگوریتم به‌تفصیل بحث می‌شود. نتایج تجربی در بخش ۵ ارائه‌شده است. درنهایت، در بخش ٦، مسیر آینده خود را مشخص می‌کنیم و به جمع‌بندی مقاله می‌پردازیم.

۲- کارهای مرتبط

اخیراً، تلاش‌های زیادی برای بهبود انتخاب ویژگی چند برچسبی انجام‌شده است. پنیری و همکاران [۹] یک روش انتخاب ویژگی چند برچسبی را بر اساس الگوریتم بهینه‌سازی کلونی مورچگان[2] پیشنهاد کردند. رویکرد آن‌ها دو تابع اکتشافی[3] با نظارت[4] و بدون نظارت[5] را معرفی کرد که حداکثر همبستگی بین ویژگی‌ها و برچسب‌های کلاس برای اندازه‌گیری ارتباط ویژگی استفاده می‌شود. همچنین حداقل همبستگی بین ویژگی‌ها برای اندازه‌گیری افزونگی بین ویژگی‌ها استفاده‌شده است.MGFS[6] [۱۰] یک رویکرد مبتنی بر فیلتر است که مسئله انتخاب ویژگی چند برچسبی را به یک گراف کامل مدل می‌کند. در این مدل، هر رأس گراف نشان‌دهنده یک ویژگی است. روش MGFS از الگوریتم معروف پیج رنک[7] برای ساختن یک سیستم رتبه‌بندی ویژگی‌ها بر اساس یک مدل مبتنی بر گراف استفاده کرد.

الگوریتم ارائه‌شده در مرجع [٤] رویکرد جدیدی است که برای کاهش ابعاد داده‌های چند برچسبی پیشنهادشده است. در این روش، ویژگی‌ها بر اساس رابطه آن‌ها با برچسب‌های کلاس در یک فرایند تصمیم‌گیری چند شاخصه[8]ارزیابی می‌شوند. BMFS[9] [۱۱] یک روش انتخاب ویژگی چند برچسبی است که رابطه ویژگی‌ها و برچسب‌های کلاس را به یک مدل بر اساس گراف دوبخشی[10] تبدیل می‌کند. برای ارزیابی ویژگی‌ها در این فضا، یک روش تطبیق گراف[11] با استفاده از الگوریتم مجارستانی[12] بر اساس همبستگی بین ویژگی‌ها و برچسب‌ها انجام‌شده است. در مرجع[۱۲] یک الگوریتم انتخاب ویژگی چند برچسبی جدید بر اساس همبستگی‌های محلی و سراسری[13] ارائه‌شده است. در این روش برچسب‌ها بر اساس روابطشان دسته‌بندی‌شده‌اند و برای هر گروه زیرمجموعه‌ای از ویژگی‌ها انتخاب‌شده است.

[1] Ensemble feature selection
[2] Ant colony optimization
[3] Heuristic
[4] Supervised
[5] Unsupervised
[6] Multi-label Graph-based Feature Selection
[7] PageRank
[8] Multi-Criteria Decision-Making
[9] Bipartite Multi-label Feature Selection
[10] Bipartite
[11] Graph matching
[12] Hungarian algorithm
[13] Global

پنیری و همکاران [۱۳]یک الگوریتم انتخاب ویژگی مبتنی بر الگوریتم بهینه‌سازی کلونی مورچگان برای یادگیری چند برچسبی پیشنهاد کردند. این الگوریتم برخلاف همه الگوریتم‌های انتخاب ویژگی مبتنی بر الگوریتم بهینه‌سازی کلونی مورچگان، یک تابع ابتکاری پویا را اعمال کرده است. در این روش، یک روش یادگیری اکتشافی[1] بر اساس الگوریتم یادگیری تقویتی[2] تفاوت زمانی[3]انجام‌شده است. ژانگ و همکاران[۱۴]یک روش انتخاب ویژگی جدید ارائه کردند که تأثیر برچسب‌ها را به‌طور یکسان در نظر نمی‌گیرد. آن‌ها مجموعه برچسب را در یک ساختار معنایی مدل‌سازی کردند که در آن برچسب‌ها به دسته‌های اصلی و فرعی طبقه‌بندی‌شده‌اند. در این مقاله با استفاده از یک رویکرد خوشه‌بندی، ویژگی‌ها ابتدا با توجه به موضوعات اصلی انتخاب و سپس با موضوعات فرعی مطابقت داده می‌شوند. یک استراتژی انتخاب ویژگی برخط برای داده‌های چند برچسبی توسط پل و همکاران[۱۵] ارائه‌شده است که مبتنی بر سه فازی اصلی است. در دو مرحله اول، ویژگی‌ها بر اساس الگوریتم بهینه‌سازی ازدحام ذرات[4] چند هدفی و افزونگی ویژگی‌ها فیلتر می‌شوند. درنهایت در فاز سوم، ویژگی‌هایی که در انتخاب ویژگی‌های جدید کارایی ندارند حذف می‌شوند. فن و همکاران[۱۶] یک تکنیک مبتنی بر رگرسیون غیرهمبسته[5] و گراف طیفی تطبیقی[6] را برای در نظر گرفتن جنبه‌های مرتبط بودن و افزونگی ویژگی‌ها در داده‌های چند برچسبی ارائه کرده است.

۳- مفاهیم پایه

۳-۱ دسته‌بندی چند برچسبی

یک داده چند برچسبی شامل یک بردار ویژگی $X_i=(x_{i1},x_{i2},\ldots\ldots,x_{iM})$ و یک بردار برچسب دودویی $Y_i=(y_{i1},y_{i2},\ldots\ldots,y_{iL})$ است که M و L به ترتیب تعداد ویژگی‌ها و برچسب‌ها را برای هر نمونه نشان می‌دهند. یادگیری چند برچسبی فرآیندی است که یک مدل یادگیری متشکل از N نمونه آموزشی یک مجموعه را داده می‌سازد که برای نمونه‌های جدید می‌تواند مجموعه برچسب‌ها را پیش‌بینی کند. شکل (۱) ساختار مجموعه داده چند برچسبی را نشان می‌دهد.

۳-۲ انتگرال فازی سوگنو

در مسائل شورایی و یا ترکیب که در آن به دنبال ترکیب چندین مقدار برای رسیدن به یک مقدار واحد به‌منظور تصمیم‌گیری هستیم، انتگرال‌های فازی می‌توانند استفاده می‌شوند. در انتگرال‌های فازی تصور می‌شود که مقادیر ورودی در یک بازه مشخص را قرار دارند. درواقع انتگرال‌های فازی بر اساس شاخص‌های فازی عمل می‌کنند.

X				**Y**			
X_1	X_2	⋯	X_M	Y_1	Y_2	⋯	Y_L
X_{11}	X_{12}	⋯	X_{1M}	Y_{11}	Y_{12}	⋯	Y_{1L}
X_{21}	X_{22}	⋯	X_{2M}	Y_{21}	Y_{22}	⋯	Y_{2L}
⋮	⋮	⋱	⋮	⋮	⋮	⋱	⋮
X_{N1}	X_{N2}	⋯	X_{NM}	Y_{N1}	Y_{N2}	⋯	Y_{NL}

شکل (۱): ساختار مجموعه داده چند برچسبی

تعریف ۱. فرض کنیم $N=\{1,2,\ldots,n\}$ باشد. یک شاخص فازی را می‌توان به‌عنوان یک مجموعه تابع $\mu:2^N\rightarrow[0,1]$ تعریف کرد که شرطهای زیر را ارضا کند [۱۷]:

1) اگر $A\subset B$ آنگاه $\mu(A)\leq\mu(B)$ یکنوا[7] باشد:
2) $\mu(N)=1$ و $\mu(\emptyset)=0$ نرمال شده باشد[8]:

تعریف ۲. اگر فرض کنیم " X=" {"x" _"1" "," "x" _"2" "," " {"x" _"n" "," "x" _"3" ",…," یک مجموعه متناهی باشد و "∃λ" (" −1,∞")، آنگاه شاخص فازی "λ" می‌تواند به‌صورت یک تابع "x"^ "2" ":" "λ" _"μ" تعریف شود که شرایط زیر را ارضا کند[۱۷]:

1) $\mu_\lambda(X)=1$
2) اگر $B\in 2^X$ ،
آنگاه $\mu_\lambda(A\cup B)=\mu_\lambda(A)+\mu_\lambda(B)+\lambda\mu_\lambda(A)\mu_\lambda(B)$
که $A\cap B=\emptyset$

[1] Heuristic learning
[2] Reinforcement learning
[3] Temporal difference
[4] Particle swarm optimization
[5] Uncorrelated regression
[6] Adaptive spectral graph
[7] Monotic
[8] Normalized

در این شرایط ما باید λ را در ابتدا محاسبه کنیم تا درنهایت به μ_λ برای ترکیبات مختلف دست پیدا کنیم. شاخص تراکم فازی[1] $\mu(X)=\mu_\lambda(\{x_1,x_2,x_3,\ldots,x_n\})$ به‌صورت زیر تعریف می‌شود:

$$\mu_\lambda(\{x_1,x_2,x_3,\ldots,x_n\})=\sum_{i=1}^{n}\mu_i+\lambda\sum_{i_1=1}^{n-1}\sum_{i_2=i_1+1}^{n}\mu_{i_1}.\mu_{i_2}+\ldots \quad (۱)$$

که $\lambda\in(-1,\infty)$ است. همان‌طور که در قبل اشاره شد، $\mu_\lambda(X)=1$ است، بنابراین می‌توانیم رابطه (۱) را برای محاسبه λ به‌صورت زیر ساده‌سازی کنیم.

$$\lambda=\prod_{i=1}^{n}(\lambda\mu_i+1)-1 \quad (۲)$$

اثبات رابطه (۲) در مرجع [۱۸]ارائه‌شده است.

تعریف ۳. فرض کنیم μ یک شاخص فازی در X باشد، آنگاه انتگرال فازی یک تابع $f:X\longrightarrow[0,\infty]$ با توجه به شاخص فازی μ به‌صورت زیر تعریف می‌شود:

$$\int f(x)d\,\mu=\max_{1\le i\le n}\left(\min\left(f(x_i),\mu(A_i)\right)\right) \quad (۳)$$

که $A_i\subset X$ برای $\{f(x_1),f(x_2),f(x_3),\ldots,f(x_n)\}$, $i=1,2,\ldots,n$ بازه‌ها[2] هستند که به‌صورتی تعریف می‌شوند که $f(x_1)\le f(x_2)\le f(x_3)\le\ldots\le f(x_n)$ و $f(x_0)=0$ است.

برای درک بهتر مسئله از یک مثال عددی استفاده می‌کنیم. مجموعه $X=\{x_1,x_2,x_3\}$ را در نظر بگیرید که بازه‌ها به‌صورت $f(x_1)=$ ۰/۶، $f(x_2)=$ ۰/۸ و $f(x_3)=$ ۰/٤ هستند. همچنین مقادیر تراکم فازی در آن به‌صورت $\mu_\lambda(\{x_1\})=$ ۰/۳، $\mu_\lambda(\{x_2\})=$ ۰/٤ و $\mu_\lambda(\{x_3\})=$ ۰/۲ وجود دارند. حال برای به دست آوردن μ_λ برای ترکیبات مختلف عناصر X باید ابتدا λ را با استفاده از رابطه (۲) محاسبه کنیم:

$$\lambda=(0.4\lambda+1)(0.3\lambda+1)(0.2\lambda+1)-1$$

اگر این معادله را حل کنیم به مقادیر [$\lambda=$ ۰ و ۱۱/۸۷- و ۰/۳۷۱۹] می‌رسیم. از قبل می‌دانیم که $\lambda\in(-1,\infty)$ پس مقادیر قابل‌قبول برای λ برابر عدد صفر و ۰/۳۷۱۹ هستند. در این مثال، رابطه را برای $\lambda=$ ۰/۳۷۱۹ محاسبه می‌کنیم. مقادیر μ_λ برای محاسبه تمام حالت‌ها در X بر اساس تعریف ۲ به دست می‌آیند:

$$\mu_\lambda(\{x_1,x_2\})=\mu_\lambda(\{x_1\})+\mu_\lambda(\{x_2\})+\lambda\mu_\lambda(\{x_1\})\mu_\lambda(\{x_2\})=0.7446$$

$$\mu_\lambda(\{x_2,x_3\})=\mu_\lambda(\{x_2\})+\mu_\lambda(\{x_3\})+\lambda\mu_\lambda(\{x_2\})\mu_\lambda(\{x_3\})=0.5223$$

$$\mu_\lambda(\{x_1,x_3\})=\mu_\lambda(\{x_1\})+\mu_\lambda(\{x_3\})+\lambda\mu_\lambda(\{x_1\})\mu_\lambda(\{x_3\})=0.7323$$

$$\mu_\lambda(\{x_1,x_2,x_3\})=1$$

درنهایت بر اساس رابطه (۳) انتگرال فازی سوگنو به‌صورت زیر محاسبه می‌شود:

$$\int f(x)d\,\mu=\max\begin{pmatrix}\min\left(f(x_1),g_\lambda(\{x_1,x_2,x_3\})\right),\\ \min\left(f(x_2),g_\lambda(\{x_2,x_3\})\right),\min\left(f(x_3),g_\lambda(\{x_3\})\right)\end{pmatrix}=$$

$$\max\begin{pmatrix}\min(0.4,1),\\ \min(0.6,0.5223),\min(0.8,0.4)\end{pmatrix}=$$

$$\max(0.4,0.5223,0.4)=0.5223$$

۳-۳ ضریب همبستگی پیرسون

ضریب همبستگی پیرسون [۱۰]بین دو متغیر تصادفی از طریق نرمال‌سازی کوواریانس بین آن‌ها به دست می‌آید. به این صورت که مقدار کوواریانس بر حاصل‌ضرب واریانس دو متغیر تقسیم می‌شود. این معیار برای اندازه‌گیری میزان مرتبط بودن دو متغیر تصادفی استفاده می‌شود که مقدار آن در بازه[۱-,۱] قرار دارد. رابطه زیر برای محاسبه ضریب همبستگی پیرسون برای دو متغیر تصادفی X و Y استفاده می‌شود.

$$r(X,Y)=\frac{Cov(X,Y)}{S_X.S_Y} \quad (٤)$$

$$d(X,Y)=1-r(X,Y) \quad (۵)$$

[1] Fuzzy density

[2] Ranges

٤- روش پیشنهادی

در این بخش به معرفی و تشریح روش پیشنهادی می‌پردازیم. الگوریتم پیشنهادی برای مسئله انتخاب ویژگی چند برچسبی ارائه‌شده است که از یک رویکرد شورایی برای تصمیم‌گیری برای انتخاب ویژگی استفاده‌شده است. در پژوهش‌های مختلف به‌دفعات نشان داده‌شده است که تنوع روش‌های مختلف در رویکردهای شورایی از به دام افتادن در بهینه محلی جلوگیری می‌کند و همچنین باعث بهبود عملکرد الگوریتم‌ها می‌شود [۸,۱۹]. دلیل این بهبود عملکرد به این مسئله برمی‌گردد که فرایند تصمیم‌گیری در مورد هر ویژگی بر اساس چندین شاخص مختلف صورت می‌گیرد و هر شاخص ویژگی‌ها را از یک جنبه مشخص ارزیابی می‌کند؛ بنابراین ویژگی‌های متمایز و بهینه آن دسته از ویژگی‌ها هستند که در برآیند کلی شاخص‌های موردنظر امتیاز بالاتری را کسب کنند. درروش پیشنهادی، از سه معیار فیلتر برای ارزیابی ویژگی‌ها به‌عنوان شاخص‌های تصمیم‌گیری استفاده‌شده است. دو معیار برای اندازه‌گیری میزان ارتباط ویژگی‌ها با برچسب کلاس و یک مورد برای محاسبه افزونگی ویژگی ورودی با ویژگی‌هایی که از قبل انتخاب‌شده‌اند، در نظر گرفته‌شده است. ترکیب این سه شاخص از طریق روش انتگرال فازی سوگنو انجام‌گرفته است. مراحل گام‌به‌گام روش پیشنهادی در الگوریتم (۱) نمایش داده‌شده‌اند و در ادامه نیز این مراحل را بر اساس الگوریتم (۱) شرح خواهیم داد.

در مرحله اول الگوریتم، یک بردار تهی را مقداردهی اولیه می‌کنیم تا ویژگی‌های مرتب‌سازی را اضافه کنیم. یک داده چند برچسبی به‌عنوان دو ماتریس ویژگی و برچسب کلاس در دسترس قرار دارد. سطرها به نمونه‌ها در ماتریس ویژگی‌ها و برچسب‌ها اشاره دارند، درحالی‌که ستون‌های این دو ماتریس به ترتیب نشان‌دهنده ویژگی‌ها و برچسب‌ها هستند که این ساختار تحت عنوان دو ماتریس X و Y در شکل (۱) نمایش داده‌شده‌اند.

در انتخاب ویژگی چند برچسبی، ویژگی‌هایی که همبستگی بیشتری با برچسب‌ها دارند بهتر هستند. بنابراین، ما همبستگی بین ویژگی‌ها و برچسب‌ها را به‌عنوان معیار مهمی در ارزیابی ویژگی‌ها در نظر گرفته‌ایم.

الگوریتم ۱. ارائه یک روش انتخاب ویژگی مبتنی بر مفهوم انتگرال فازی در یادگیری چند برچسبی

ورودی: ماتریس برچسب کلاس با ابعاد $N \times L$، ماتریس ویژگی با ابعاد $N \times M$

خروجی: زیرمجموعه ویژگی انتخاب‌شده S

۱. ایجاد بردار تهی S

۲. محاسبه ماتریس ضریب همبستگی پیرسون و در نظر گرفتن حداکثر مقدار همبستگی هر ویژگی با برچسب‌های کلاس بر اساس رابطه (٤)

۳. درنظرگرفتن حداقل مقدار همبستگی هر ویژگی با برچسب‌های کلاس بر اساس رابطه (٤)

٤.. محاسبه ماتریس فاصله همبستگی و درنظرگرفتن ماکسیمم مقدار فاصله همبستگی هر ویژگی با سایر ویژگی بر اساس رابطه (٥)

٥. تعیین مقدار تراکم فازی برای سه روش فیلتر: محاسبه فاصله همبستگی بین روش‌های فیلتر و محاسبه میانگین فاصله‌ها برای هر روش

٦. نرمال‌سازی فاصله بر اساس رابطه (۷) و نسبت دادن مقادیر به‌دست‌آمده به‌عنوان تراکم فازی هر روش

۷. محاسبه λ بر اساس رابطه (۲) و مقادیر تراکم فازی

۸. دو گام زیر را برای هر ویژگی انجام شود:

۹. برای هر ویژگی مرتب‌سازی سه مقدار ضریب همبستگی پیرسون، شباهت کسینوسی و فاصله همبستگی به‌صورت صعودی و محاسبه مقادیرμ_λ با توجه به ترتیب روش‌ها

۱۰. محاسبه انتگرال فازی سوگنو بر اساس مقادیر بازه‌ها و تراکم فازی بر اساس رابطه (۳)

۱۱. مرتب‌سازی ویژگی‌ها بر اساس مقدار انتگرال فازی به‌صورت نزولی

۱۲. مجموعه S به‌عنوان مجموعه ویژگی‌های مرتب مشخص می‌شود و کاربر می‌تواند به‌اندازه دلخواه انتخاب کند.

در این مقاله از ترکیبی از روش‌های ارزیابی ویژگی‌ها با یک رویکرد شورایی به‌منظور انتخاب ویژگی‌ها استفاده کرده‌ایم. در این

رویکرد، دو روش سنجش همبستگی ویژگی‌ها با برچسب‌های کلاس و یک روش برای سنجش افزونگی ویژگی‌ها در نظر گرفته‌شده است.

معیارهایی همبستگی می‌توانند قدرت پیش‌بینی ویژگی‌ها را منعکس کنند. درجه ارتباط ویژگی‌های مختلف با برچسب‌های کلاس متفاوت است و ویژگی‌های باهمبستگی پایین ممکن است الگوریتم یادگیری را گمراه کنند. همچنین، اندازه‌گیری افزونگی می‌تواند ویژگی‌های منحصربه‌فردی درداده را مشخص کند که بدان معناست که برخی از ویژگی‌ها فقط ترکیبی از سایر ویژگی‌ها هستند و حاوی اطلاعات مفیدی نیستند؛ بنابراین، بهترین زیرمجموعه ویژگی، زیرمجموعه‌ای است که مرتبط‌ترین و حداقل ویژگی‌های افزونه را داشته باشد. بر اساس این منطق، در این مقاله ترکیبی از روش‌های سنجش همبستگی و افزونگی استفاده می‌شود. میزان حداکثر و حداقل همبستگی هر ویژگی با مجموعه برچسب‌های کلاس بر اساس ضریب همبستگی پیرسون به‌عنوان شاخص‌های همگرایی و معیار فاصله همبستگی به‌منظور محاسبه میزان شباهت هر ویژگی با سایر ویژگی‌ها در مجموعه داده استفاده شده‌اند.

لذا در گام‌های دوم و سوم الگوریتم این معیارها بر روی مجموعه داده اعمال می‌شوند و ماتریس متناظر حاصل می‌شود. به این صورت که در گام دوم با استفاده از شاخص ضریب همبستگی پیرسون و بر اساس رابطه (٤)، همبستگی هر ویژگی با برچسب‌های کلاس محاسبه می‌شود. درنهایت یک ماتریس حاصل می‌شود که سطرهای آن برابر ویژگی‌ها و ستون‌های آن نشان‌دهنده برچسب‌ها هستند. بر اساس این ماتریس، مقدار متناظر در هر درایه ماتریس برابر مقدار ضریب همبستگی پیرسون بین ویژگی و برچسب متناظر آن است. درنهایت مقدار حداکثر و حداقل هر سطر از این دو ماتریس به‌عنوان مقدار امتیاز هر ویژگی بر اساس روش یادشده، در نظر گرفته می‌شود. درواقع با این رویکرد، حداکثر همبستگی هر ویژگی با برچسب‌های کلاس به‌عنوان شاخص ارزیابی آن ویژگی تعیین می‌شود. در گام چهارم الگوریتم، فاصله همبستگی هر ویژگی با سایر ویژگی‌ها محاسبه و ماتریس فاصله همبستگی حاصل می‌شود. درنهایت ماکسیمم مقدار هر سطر از این ماتریس به‌عنوان شاخص ارزیابی افزونگی آن ویژگی در نظر گرفته می‌شود؛ بنابراین خروجی سه‌گام یادشده سه بردار به‌اندازه ویژگی‌ها است که هرکدام یک نوع ارزیابی از ویژگی‌ها را در بردارند.

روش پیشنهادی در این مقاله مبتنی بر یک رویکرد شورایی است که از ترکیب چندین روش ارزیابی ویژگی‌ها برای تصمیم‌گیری استفاده می‌کند. با توجه به این‌که فرایند ترکیب شاخص‌های ارزیابی توسط عملگر انتگرال فازی سوگنو انجام می‌شود، بنابراین در ابتدا باید مقادیر تراکم فازی برای هر یک از شاخص‌ها تنظیم شود. این روند در گام پنجم تا هفتم الگوریتم قابل‌مشاهده است. به این صورت که با استفاده از رابطه (٥) فاصله همبستگی بین بردارهای به‌دست‌آمده در مراحل دوم تا چهارم الگوریتم، محاسبه می‌شود و سپس میانگین آن‌ها به‌عنوان مقدار تراکم فازی در نظر گرفته می‌شود. اما بر اساس تعریف ۱ در بخش سوم مقاله، باید مقادیر تراکم فازی در بازه [۰,۱]قرار داشته باشند. بنابراین بر اساس رابطه (٦)، مقادیر به‌دست آمده نرمال می‌شوند.

$$\mu_i = \frac{a_i}{\left(\sum_{i=1}^{F} a_i\right)+1} \qquad (٦)$$

که a یک بردار حاوی مقادیر فاصله همبستگی است که درایه‌های آن به ترتیب حداکثر و حداقل همبستگی پیرسون و فاصله همبستگی هستند و μ نیز نشان‌دهنده بردار تراکم فازی است.

همان‌طور که در بخش ۳.۲ اشاره شد، برای محاسبه مقدار انتگرال فازی سوگنو باید مقادیر μ_χ را برای تمام حالت‌های ممکن روش‌های دخیل در فرآیند ترکیب محاسبه کنیم. بنابراین در گام هفتم الگوریتم، ابتدا مقدار λ بر اساس مقادیر تراکم فازی و بر اساس رابطه (۲) محاسبه می‌شود و بر اساس مقدار λ همه حالت‌های μ_χ حاصل می‌شوند. حال تمام موارد بر محاسبه انتگرال فازی سوگنو و عمل ترکیب مقادیر شاخص‌های ارزیابی ویژگی در دسترس است. همچنین با توجه به در نظر گرفتن شرط $f(x_1) \leq f(x_2) \leq f(x_3) \leq \ldots \leq f(x_n)$ برای بازه‌ها در انتگرال سوگنو، برای هر ویژگی مقادیر شاخص‌ها به‌صورت صعودی مرتب می‌شوند و تحت عنوان $f(x_1)$ $f(x_2)$ و $f(x_3)$در نظر گرفته می‌شوند. بنابراین در ارزیابی هر ویژگی جدید، هر بار ممکن است ترتیب متفاوتی از روش‌های ارزیابی را داشته باشیم. شاخص‌های تراکم فازی متناظر نیز بر همین اساس در هر محاسبه انتگرال فازی

چوکت متناظر با تغییرات در مقادیر شاخص‌ها به شاخص متناظر خود نسبت داده می‌شوند. در پایان برای هر ویژگی بر اساس مقادیر بازه‌ها و تراکم فازی، مقدار انتگرال فازی سوگنو بر اساس رابطه (۳) محاسبه می‌شود. این مراحل در گام‌های هشتم تا یازدهم الگوریتم قرار دارند. در پایان ویژگی‌ها بر اساس مقدار به دست توسط انتگرال فازی سوگنو به‌صورت نزولی مرتب می‌شوند و در بردار S قرار می‌گیرند. بردار S درواقع دربردارنده بهترین ویژگی‌ها به ترتیب ارزش است. کاربر می‌تواند به‌اندازه دلخواه از بردار S ویژگی انتخاب کند نتایج آزمایش‌های تجربی.

۵-۱ مجموعه‌های داده

شش مجموعه داده دنیای واقعی برای اندازه‌گیری عملکرد روش پیشنهادی در برابر روش‌های مشابه استفاده‌شده‌اند. جدول (۱) شامل مشخصات مجموعه داده‌های موردنظر است.

جدول (۱): مشخصات مجموعه داده‌ها

نام مجموعه داده	تعداد نمونه‌های آموزشی	تعداد ویژگی‌ها	تعداد برچسب‌ها	مرجع
Arts	۵۰۰۰	٤٨٨	٢٦	[۲۰]
Chemsitry	٦٩٦١	٧١٥	١٧٥	[۲۱]
Cooking	١٠٤٩١	٩٧٧	٤٠	[۲۱]
Science	٥٠٠٠	٤٨٣	٤٠	[۲۰]
Philosophy	٣٩٧١	١٠٧٥	٢٣٣	[۲۱]
CS	٩٢٧٠	٩٠٩	٢٧٤	[۲۱]

۵-۲ معیارهای ارزیابی

برای اندازه‌گیری عملکرد روش پیشنهادی و مقایسه روش‌ها، از زیان همینگ[1] و دقت[2] به‌عنوان معیارهای ارزیابی چند برچسبی و زمان اجرا الگوریتم‌ها استفاده کردیم. فرض کنیم $T=\{(x_i, Y_i),i=1,\ldots\ldots,N\}$ مجموعه آزمون، $Y_i \subseteq L$ مجموعه برچسب‌های حقیقی و $Z_i \subseteq L$ مجموعه برچسب‌های پیش‌بینی‌شده نسبت به نمونه آموزشی x_i باشند. همچنین $f(x,y)$ برابر امتیاز نسبت داده‌شده به برچسب y برای نمونه آموزشی x باشد. بنابراین معیارهای زیان همینگ و دقت به‌صورت زیر تعریف می‌شوند [۱]:

زیان همینگ: این معیار اندازه‌گیری می‌کند که چند بار یک جفت نمونه–برچسب به درسته دسته‌بندی نشده است:

$$Hamming\ loss(h,T)=\frac{1}{N}\sum_{i=1}^{N}\frac{|Y_i \Delta Z_i|}{|L|} \qquad (۷)$$

که Δ تفاضل متقارن بین دو مجموعه است.

دقت: تعداد برچسب‌های پیش‌بینی‌شده درست را در بین همه برچسب‌های واقعی و پیش‌بینی‌شده محاسبه می‌کند.

$$Accuracy(h,T)=\frac{1}{N}\sum_{i=1}^{N}\frac{|Y_i \cap Z_i|}{|Y_i \cup Z_i|} \qquad (۸)$$

۵-۳ نتایج

روش پیشنهادی در این مقاله، با شش الگوریتم انتخاب ویژگی چند برچسبی مقایسه شده است. این روش‌ها عبارت‌اند از: PPT-ReleifF[3] [۲۲]، MCLS[4] [۲۳]، LRFS[5] [۱٤]، MLACO[6] [۹]، PMFS[7] [۱] و MFS-MCDM[8] [٤].

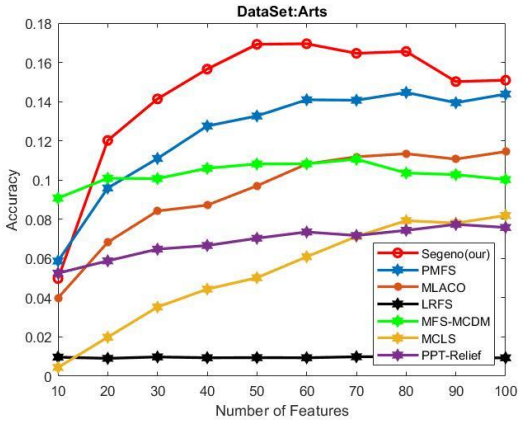
دسته‌بند ML-KNN[9] [٢٤] به‌عنوان یک دسته‌بند معروف که نسخه چند برچسبی الگوریتم KNN[10] است برای ارزیابی عملکرد روش‌های انتخاب ویژگی در نظر گرفته‌شده است. تعداد همسایگان در این دسته‌بند برابر با ۱۰ است. ٦٠ درصد از نمونه‌های آموزشی به‌طور تصادفی به‌عنوان داده‌های آموزشی برای هر آزمون انتخاب می‌شوند، درحالی‌که ٤٠ درصد نمونه باقی‌مانده به‌عنوان داده‌های آزمون در نظر گرفته می‌شود. نتایج گزارش‌شده میانگین ۳۰ اجرای جداگانه از هر روش را نشان می‌دهد. برای هر روش ارزیابی به تعداد ویژگی‌های انتخابی با تعداد [۱۰,۲۰,۳۰,٤٠,٥٠,٦٠,۷۰, و

[1] Hamming loss
[2] Accuracy
[3] Pruned Problem Transformation ReliefF multi-label feature selection
[4] Manifold-based Constraint Laplacian Score for Label Redundancy Feature Selection
[5] Label Redundancy Feature Selection
[6] Multi-Label Ant Colony Optimization
[7] Multi-label Feature Selection based on Multi-Criteria Decision-Making
[8] Multi-label Feature Selection based on Multi-Criteria Decision-Making
[9] Multi-Label K Nearest Neighbor
[10] K Nearest Neighbor

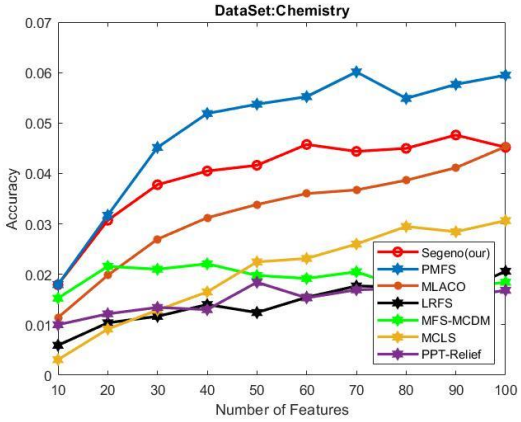
۱۰۰،۹۰،۸۰] انجام‌شده است. در این رویکرد، کاربر می‌تواند تعداد ویژگی‌ها را تعیین کند.

در شکل‌های (۲) تا (۷) نتایج به‌دست‌آمده روی مجموعه داده‌های مختلف ازنظر معیار دقت دسته‌بندی و در شکل‌های (۸) تا (۱۳) نیز نتایج برحسب معیار زیان همینگ نمایش داده‌شده‌اند. همچنین در جدول (۲) میانگین زمان اجرای الگوریتم در مجموعه داده‌های مختلف گزارش‌شده است. در این جدول بهترین روش با رنگ قرمز و روش دومین بهترین با رنگ آبی نشان داده‌شده است.

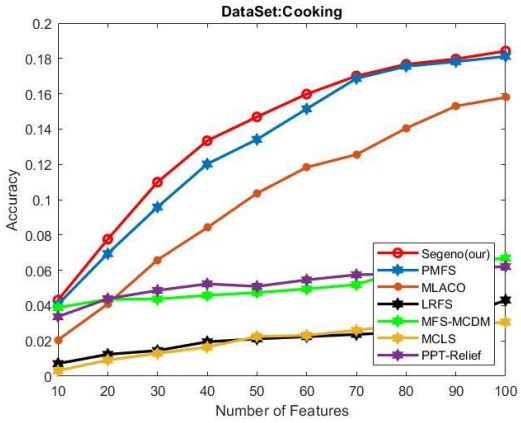
با توجه به نتایج به‌دست‌آمده بر اساس دو معیار دسته‌بندی چند برچسبی، می‌توان مشاهده کرد که روش پیشنهادی بر سایر روش‌ها برتری دارد. همچنین روش پیشنهادی دارای زمان اجرای بسیاری کوتاهی است و فقط در مقایسه با روش‌های MFS-MCDM و PMFS کندتر است که با مشاهده میانگین زمان اجرا در جدول (۲) می‌توانی مشاهده کنیم که اختلاف‌زمان اجرا با این دو روش بسیار کم است؛ بنابراین می‌توان گفت که روش پیشنهادی ازلحاظ عملکرد و زمان اجرا در شرایط مناسبی قرار دارد.



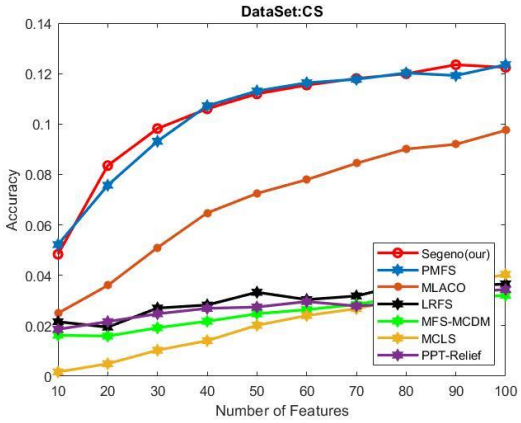
شکل (۲): نتایج بر اساس معیار دقت در مجموعه داده Arts



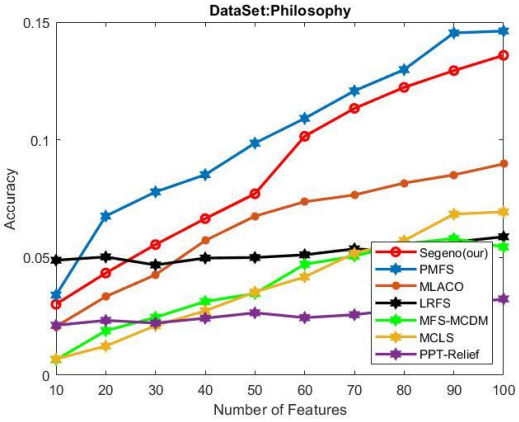
شکل (۳): نتایج بر اساس معیار دقت در مجموعه داده Chemistry



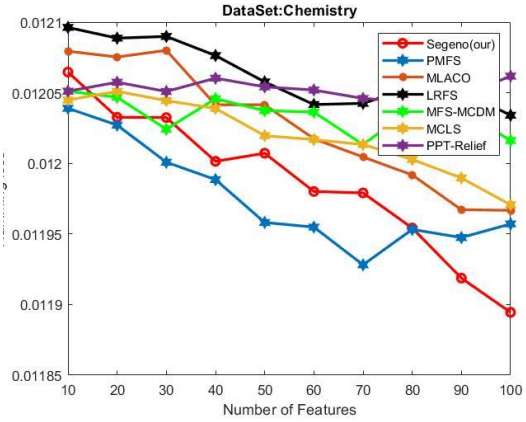
شکل (٤): نتایج بر اساس معیار دقت در مجموعه داده Cooking



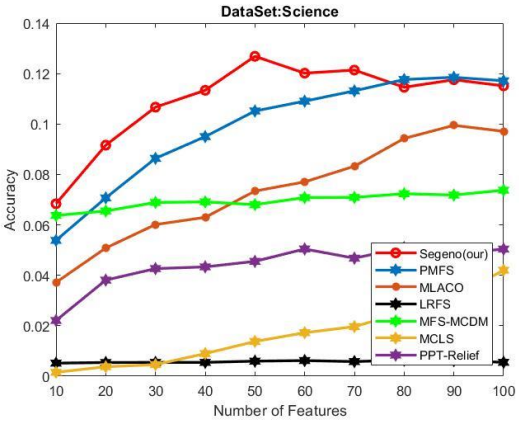
شکل (٥): نتایج بر اساس معیار دقت در مجموعه داده CS



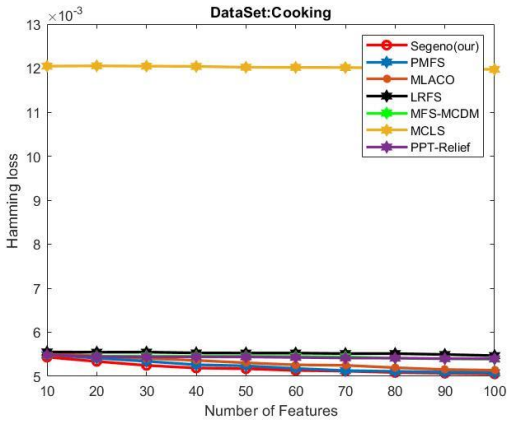
شکل (٦): نتایج بر اساس معیار دقت در مجموعه داده **Philosphy**



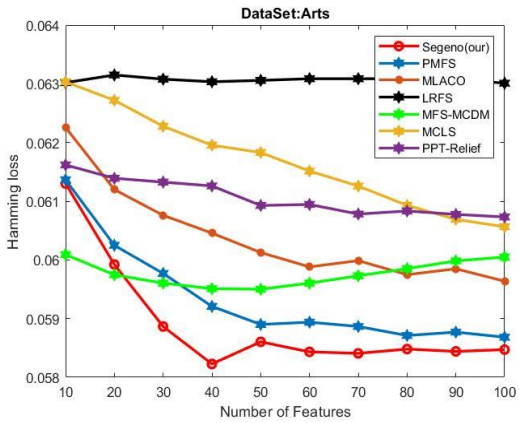
شکل (۹): نتایج بر اساس معیار زیان همینگ در مجموعه داده**Chemistry**



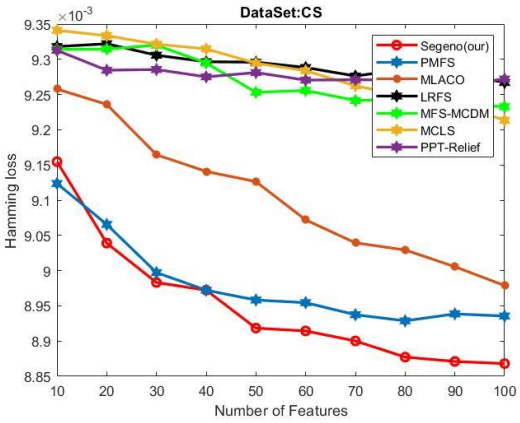
شکل (۷): نتایج بر اساس معیار دقت در مجموعه داده **Science**



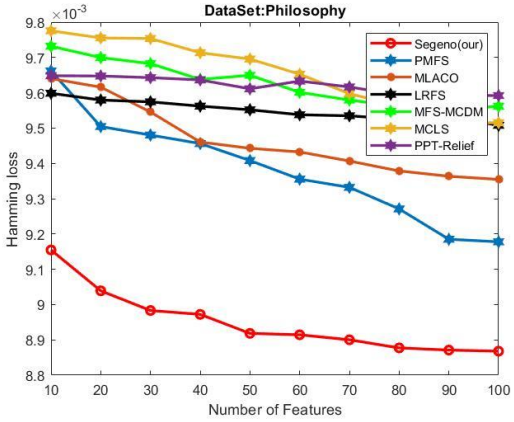
شکل (۱۰): نتایج بر اساس معیار زیان همینگ در مجموعه داده **Cooking**



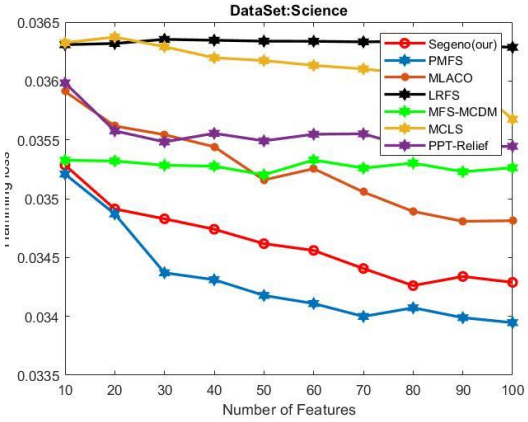
شکل (۸): نتایج بر اساس معیار زیان همینگ در مجموعه داده **Arts**



شکل (۱۱): نتایج بر اساس معیار زیان همینگ در مجموعه داده **CS**



شکل (۱۲): نتایج بر اساس معیار زیان همینگ در مجموعه داده Philosphy



شکل (۱۳): نتایج بر اساس معیار زیان همینگ در مجموعه داده Science

نام مجموعه داده	روش پیشنهادی	PMFS	MALCO	MCLS	LRFS	MFS-MCDM	PPT-ReliefF
Science	۰/۱۸	۰/۱۶	۵/۴۴	۳۸/۰۹	۵/۹۹	۰/۱۷	۳۶/۲۴
Philosophy	۰/۲۰	۰/۱۹	۷/۶۱	۵۰۴/۱۲	۱۷/۷۳	۰/۲۳	۸/۰۱
CS	۰/۳۷	۰/۱۴	۴/۲۶	۶۳۸/۵۲	۳۷/۱۰	۰/۱۷	۲۸/۹۰
Cooking	۰/۴۰	۰/۱۴	۳/۷۲	۲۵۰/۵۰	۵۲/۳۹	۰/۱۷	۳۳/۸۹
Chemsitry	۰/۲۰	۰/۰۸	۳/۱۷	۲۵۰/۷۴	۱۶/۴۱	۰/۱۲	۱۵/۱۱
Arts	۰/۱۱	۰/۰۵	۱/۵۲	۱۱/۱۰	۲/۸۸	۰/۰۷	۱۸/۸۱

جدول (۲): میانگین زمان اجرای الگوریتم‌ها برحسب ثانیه

بر اساس نتایج به‌دست‌آمده می‌توان مشاهده کرد که روش پیشنهادی در مقایسه با روش‌های جدید و قدرتمند در حوزه انتخاب ویژگی چند برچسبی عملکرد مناسبی داشته است. دلیل این عملکرد به عقیده ما استفاده از یک رویکرد تجمعی با درنظرگرفتن تأثیر و تنوع روش‌های مختلف است. در هیچ‌یک از روش‌های مقایسه شده در این مقاله از رویکرد شورایی در سطح ارزیابی ویژگی استفاده‌نشده است. همچنین این روش از ترکیبی از معیارهای همبستگی و افزونگی استفاده می‌کند. در دو الگوریتم MFS-MCDM و PPT-ReliefF تنها همبستگی ویژگی‌ها با برچسب‌های کلاس به‌عنوان شاخص ارزیابی در نظر گرفته‌شده است؛ بنابراین می‌توان دلیل برتری روش پیشنهادی بر این دو الگوریتم را عدم تأثیر افزونگی ویژگی‌ها دانست.

در سایر روش‌ها ترکیب همبستگی و افزونگی ویژگی‌ها به‌عنوان معیار ارزیابی در نظر گرفته‌شده است. اما درروش پیشنهادی، ما همبستگی ویژگی‌ها را از دو دیدگاه مدنظر قرار داده‌ایم تا بتواند نسبت به سایر الگوریتم عملکرد بهتری داشته باشد. به عقیده ما، در یک داده چند برچسبی یک ویژگی که همبستگی بالایی با یک برچسب کلاس داشته باشد، می‌تواند ویژگی مفیدی برای پیش‌بینی آن برچسب باشد؛ بنابراین درروش پیشنهادی، حداکثر همبستگی هر ویژگی با مجموعه برچسب‌ها را به‌عنوان یک شاخص در نظر گرفته‌ایم. در بسیاری از الگوریتم‌های انتخاب ویژگی چند برچسبی ازجمله روش‌های مقایسه شده از مجموع یا حداکثر همبستگی استفاده‌شده است. مجموع همبستگی با افزایش تعداد برچسب‌ها می‌تواند منجر ازبین‌رفتن تأثیر هر ویژگی برای برچسب‌های خاصی

از کلاس شود؛ بنابراین حداکثر همبستگی معیار مناسب‌تری است، زیرا ویژگی مؤثر در پیش‌بینی یک برچسب تأثیر بالاتری می‌گیرد. اما برای بهبود این موضوع، ما حداقل همبستگی هر ویژگی با برچسب کلاس‌ها را نیز یک عامل در ارزیابی ویژگی‌ها در نظر گرفته‌ایم که در هیچ‌یک از روش‌های مقایسه شده دیده نمی‌شود. منطق این موضوع بر این اساس است که ویژگی که کمترین همبستگی آن با برچسب کلاس از سایر ویژگی‌ها بیشتر باشد، ویژگی مهم‌تری است؛ بنابراین ترکیب حداکثر و حداقل همبستگی با دیدگاه حداکثری می‌تواند منجر به ارزیابی مناسب‌تری از ویژگی‌ها شود.

همچنین به‌منظور تأثیر دادن افزونگی نیز از معیار فاصله همبستگی بین ویژگی‌ها استفاده‌شده است. همچنین وزن دهی به این مقادیر نیز بر اساس فاصله همبستگی صورت گرفته است که معیار با تنوع بیشتر، وزن بالاتری را در فرایند تجمیع به خود اختصاص دهد. ما از عملگر فازی سوگنو به‌عنوان یک روش تجمیع خطی و سریع استفاده کرده‌ایم. نتایج حاصل نشان‌دهنده عملکرد مناسب روش پیشنهادی ازنظر معیارهای دسته‌بندی و زمان اجرا دارد. ازنظر زمان اجرا روش پیشنهادی سومین عملکرد برتر را دارد که دلیل اجرای عملگر سوگنو به‌اندازه تعداد ویژگی‌ها و محاسبه سه معیار برای هر ویژگی است. بااین‌حال، الگوریتم ازنظر زمانی بسیار نزدیک به دو الگوریتم PMFS و MFS-MCDM عمل کرده است.

۵- نتیجه‌گیری

این مقاله یک روش انتخاب ویژگی چند برچسبی را با استفاده از انتگرال فازی سوگنو پیشنهاد می‌کند. این مقاله سه معیار ارزیابی ویژگی را برای تصمیم‌گیری در خصوص انتخاب ویژگی‌ها در نظر می‌گیرد. این مقادیر به عملگر انتگرال فازی سوگنو تحویل داده می‌شود تا فرایند انتخاب ویژگی بر اساس ترکیب این مقادیر و با استفاده از یک رویکرد شورایی انجام پذیرد. نتایج اجرای الگوریتم بر روی مجموعه داده‌های مختلف نشان‌دهنده کارایی و بهینه بودن روش پیشنهادی است. زیرا همگرایی و تنوع الگوریتم با انواع روش‌های انتخاب ویژگی متعادل می‌شود که در نتایج به‌دست‌آمده عملکرد الگوریتم پیشنهادی بر روش‌های رقیب برتری دارد. همچنین به دلیل محاسبات ساده در فرایند الگوریتم، روش در زمان اجرای بسیار کوتاهی اجرا می‌شود. به‌عنوان کارهای آینده قصد داریم از تکنیک‌های شورایی در انتخاب ویژگی برای سایر کاربردهای یادگیری ماشین استفاده کنیم و رویکرد شورایی را در مسائل مختلف بهینه‌سازی در یادگیری ماشین آزمایش می‌کنیم. همچنین، قصد داریم کار خود را در مسائل انتخاب ویژگی برخط و غیر برخط توسعه دهیم و مدل‌های پیشنهادی خود را در کاربردهای مختلف آزمایش کنیم.

References

[1] Hashemi, A., Bagher Dowlatshahi, M., and Nezamabadi-pour, H. (2021) An efficient Pareto-based feature selection algorithm for multi-label classification. Information Sciences. 581 428–447.

[2] Dhal, P. and Azad, C. (2022) A comprehensive survey on feature selection in the various fields of machine learning. Applied Intelligence. 52 (4), 4543–4581.

[3] Deng, X., Li, Y., Weng, J., and Zhang, J. (2019) Feature selection for text classification: A review. Multimedia Tools and Applications. 78 (3), 3797–3816.

[4] Hashemi, A., Dowlatshahi, M.B., and Nezamabadi-pour, H. (2020) MFS-MCDM: Multi-label feature selection using multi-criteria decision making. Knowledge-Based Systems. 206 106365.

[5] Kashef, S., Nezamabadi-pour, H., and Nikpour, B. (2018) Multilabel feature selection: A comprehensive review and guiding experiments. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. 8 e1240.

[6] Hashemi, A., Dowlatshahi, M.B., and Nezamabadi-Pour, H. (2020) A bipartite matching-based feature selection for multi-label learning. International Journal of Machine Learning and Cybernetics.

[7] Cai, J., Luo, J., Wang, S., and Yang, S. (2018) Feature selection in machine learning: A new perspective. Neurocomputing. 300 70–79.

[8] Bolón-Canedo, V. and Alonso-Betanzos, A. (2019) Ensembles for feature selection: A review and future trends. Information Fusion. 52 1–12.

[9] Paniri, M., Dowlatshahi, M.B., and Nezamabadi-pour, H. (2020) MLACO: A multi-label feature selection algorithm based on ant colony optimization. Knowledge-Based Systems. 192 105285.

[10] Hashemi, A., Dowlatshahi, M.B., and Nezamabadi-pour, H. (2020) MGFS: A multi-label graph-based feature selection algorithm via PageRank centrality. Expert Systems with Applications. 142 113024.

[11] Hashemi, A., Dowlatshahi, M.B., and Nezamabadi-Pour, H. (2021) A bipartite matching-based feature selection for multi-label learning. International Journal of Machine Learning and Cybernetics. 12 (2), 459–475.

[12] Che, X., Chen, D., and Mi, J. (2020) A novel approach for learning label correlation with application to feature selection of multi-label data. Information Sciences. 512 795–812.

[13] Paniri, M., Dowlatshahi, M.B., and Nezamabadi-pour, H. (2021) Ant-TD: Ant colony optimization plus temporal difference reinforcement learning for multi-label feature selection. Swarm and Evolutionary Computation. 64 100892.

[14] Zhang, P., Liu, G., and Gao, W. (2019) Distinguishing two types of labels for multi-label feature selection. Pattern Recognition. 95 72–82.

[15] Paul, D., Jain, A., Saha, S., and Mathew, J. (2021) Multi-objective PSO based online feature selection for multi-label classification. Knowledge-Based Systems. 222 106966.

[16] Fan, Y., Liu, J., Weng, W., Chen, B., Chen, Y., and Wu, S. (2021) Multi-label feature selection with constraint regression and adaptive spectral graph. Knowledge-Based Systems. 212 106621.

[17] Beliakov, G. and Divakov, D. (2020) On representation of fuzzy measures for learning Choquet and Sugeno integrals. Knowledge-Based Systems. 189 105134.

[18] Ayub, M. (2009) Choquet and Sugeno Integrals, 2009.

[19] Hashemi, A., Dowlatshahi, M.B., and Nezamabadi-pour, H. (2022) Ensemble of feature selection algorithms: a multi-criteria decision-making approach. International Journal of Machine Learning and Cybernetics. 13 (1), 49–69.

[20] Ueda, N. and Saito, K. (2003) Parametric mixture models for multi-labeled text. in: Adv. Neural Inf. Process. Syst., pp. 737–744.

[21] Charte, F. and Charte, D. (2015) Working with multilabel datasets in R: The mldr package. R Journal. 7 (2), 149–162.

[22] Reyes, O., Morell, C., and Ventura, S. (2015) Scalable extensions of the ReliefF algorithm for weighting and selecting features on the multi-label learning context. Neurocomputing. 161.

[23] Huang, R., Jiang, W., and Sun, G. (2018) Manifold-based constraint Laplacian score for multi-label feature selection. Pattern Recognition Letters. 112 346–352.

[24] Cherman, E.A., Spolaôr, N., Valverde-Rebaza, J., and Monard, M.C. (2015) Lazy Multi-label Learning Algorithms Based on Mutuality Strategies. Journal of Intelligent and Robotic Systems: Theory and Applications. 80 261–276.

Journal of Applied And Basic Machine Intelligence Research (ABMIR)

Applied and Basic Machine Intelligence Research

Vol. 1, No. 1, Spring & Summer 2023, 1-13

A feature selection algorithm based on fuzzy integral in multi-label learning

Amin Hashemi [1], Mohammad Bagher Dowlatshahi [2*]

[1] Department of Computer Engineering, Faculty of Engineering, Lorestan University, Khorramabad, Iran

Article Information

Original Research Paper

Received:
17 February 2022

Accepted:
14 June 2022

Keywords:
Feature selection, Multi-label learning, Fuzzy integral, Ensemble approach

Corresponding Author*:
dowlatshahi.mb@lu.ac.ir

Abstract

Multi-label learning algorithms face many challenges due to the high volume and dimensions of multi-label data and the existence of noise. Feature selection methods are an effective technique for addressing these challenges. This paper presents a feature selection method based on an ensemble approach for multi-label data. In this approach, three different decision matrices based on various feature evaluation criteria, taking into account the relevancy of features with class labels and their redundancy relative to each other, are effective in the feature selection process. These three decision matrices are finally combined based on an ensemble approach using the concept of fuzzy integral to evaluate the features according to the aggregate value. Comparisons have been made with several similar algorithms to illustrate the performance of the proposed method.

doi : 10.22034/ABMIR.2022.2714