

عنوان مقاله:

بهبودی بر خوشه بندی مقید با استفاده از انتخاب ویژگی

محل انتشار:

اولین کنفرانس بین المللی راهکارهای نوین در مهندسی، علوم اطلاعات و فناوری در قرن پیش رو (سال: 1398)

تعداد صفحات اصل مقاله: 9

نویسنده:

ماندانا غلامی گزافروزی - کارشناس ارشد مهندسی کامپیوتر در گرایش هوش مصنوعی

خلاصه مقاله:

الگوریتم های خوشه بندی کاربردهای بسیاری در علوم پزشکی، بانک داری، داده کاوی و... دارند. از جمله مشهورترین الگوریتم های خوشه بندی، الگوریتم K-means است که پیاده سازی ساده و سرعت بالایی دارد. اما یکی از جمله ایرادات بزرگ این الگوریتم، یکسان در نظر گرفتن ارزش تمامی ویژگی های پایگاه داده و عدم توانایی آن در استفاده از اطلاعات جانبی (قیدهای باید پیوند و نفی پیوند) است که به میزان قابل توجهی از کارایی آن می کاهد. الگوریتم CVQE با هدف بهبود الگوریتم K-means و ایجاد امکان استفاده از قیدهای مجموعه داده، طراحی شده است و توانایی استفاده از اطلاعات جانبی، یعنی قیدهای باید پیوند و نفی پیوند را دارا می باشد. ایرادی که بر این الگوریتم وارد است، یکسان در نظر گرفتن اهمیت ویژگی ها می باشد که گاه به جای بهبود نتیجه خوشه بندی، موجب اختلال در عملکرد آن خواهد شد. در این مقاله مکانیسمی پیشنهاد شده است که به منظور بهبود کیفیت نتایج خوشه بندی مقید، با بهره گیری از قابلیت های الگوریتم خوشه بندی CVQE و الگوریتم رقابت استعماری، قابلیت انتخاب ویژگی های یک مجموعه داده، به خوشه بندی مقید اضافه شود. معیار بهبود ما در این روش، کاهش فواصل درون خوشه ای و افزایش فاصله میان خوشه ای می باشد. به همین منظور از شاخص دویس بلدین برای مقایسه نتایج الگوریتم پیشنهادی و الگوریتم های خوشه بندی CVQE و Cop-K-means استفاده شده است. پس از خوشه بندی سه مجموعه داده توسط الگوریتم پیشنهادی و الگوریتم های خوشه بندی، نتایج نشان می دهد که الگوریتم پیشنهادی در بهبود خوشه بندی مقید موفق عمل نموده است. اما زمان اجرای الگوریتم پیشنهادی بیشتر از سایر الگوریتم های یاد شده می باشد.

کلمات کلیدی:

خوشه بندی، خوشه بندی مقید، انتخاب ویژگی

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/965942>

