

عنوان مقاله:

روش جدید وزن دهی ویژگی مبتنی بر اطلاعات کلاس: TFCRF در حوزه طبقه بندی مستندات

محل انتشار:

دوازدهمین کنفرانس سالانه انجمن کامپیوتر ایران (سال: 1385)

تعداد صفحات اصل مقاله: 6

نویسندگان:

مینا ملکی - دانشجوی کارشناسی ارشد دانشگاه صنعتی امیرکبیر، دانشکده مهندسی کامپی

احمد عبدالله زاده بارفروش - دانشیار دانشگاه صنعتی امیرکبیر، دانشکده مهندسی کامپیوتر و فناوری ار

خلاصه مقاله:

وزن دهی ویژگی به عنوان یکی از تکنیک های پیش پردازش در حوزه طبقه بندی مستندات ، نقش بسیار مهمی در دستیابی به شاخص بندی با کیفیت بالا و در نتیجه دستیابی به طبقه بندی کننده خوب مستندات ایفا می کند. در این مقاله یک روش جدید برای وزن دهی ویژگی به نام TFCRF خاص حوزه طبقه بندی مستندات ارائه می شود که در آن برای وزن دهی ویژگی ها علاوه بر توجه به چگونگی توزیع آنها در مستندات مختلف و مستند د ات کل م جموعه به چگونگی توزیع آنها در طبقات مختلف نیز توجه شده است . نتایج شبیه سازی نشان دهنده بهبود قابل توجهی در کارایی الگوریتم طبقه بندی کننده SVM با بکارگیری روش وزن دهی ویژگی ارائه شده جدید TFCRF در مقایسه با سایر روش های متداول وزن دهی ویژگی پیاده سازی شده نظیر روش های مبتنی بر TF، روش های مبتنی بر IDF، روش های ترکیبی TFIDF و روش های خاص طبقه بندی بر روی مجموعه مستندات inex می باشد.

کلمات کلیدی:

وزن دهی ویژگی، طبقه بندی مستندات، الگوریتم SVM ، انتخاب ویژگی، بازیابی مستندات، متن کاوی

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/44651>

