

عنوان مقاله:

دسته بندی متون با تکنیک های انتخاب ویژگی و یادگیری ماشین

محل انتشار:

همایش ملی رویکرد عملی به پیاده سازی مفاهیم علمی، مباحث تئوری و پژوهش های کاربردی علوم فنی (سال: 1393)

تعداد صفحات اصل مقاله: 5

نویسندگان:

مژگان رحیمی راد - گروه کامپیوتر، واحد علوم و تحقیقات خوزستان، دانشگاه آزاد اسلامی واحد اهواز، ایران - گروه کامپیوتر، واحد اهواز، دانشگاه آزاد اسلامی اهواز،

محمدعلی مصلح - گروه مهندسی کامپیوتر، واحد دزفول، دانشگاه آزاد اسلامی دزفول، ایران

امیرمسعود رحمانی - گروه کامپیوتر، واحد علوم و تحقیقات تهران، دانشگاه آزاد اسلامی تهران، ایران

خلاصه مقاله:

اسناد الکترونیکی حجم انبوهی از اطلاعات را در خود جای داده اند، اما برای متن کاوی بسیاری از آنها زائد و اضافی می باشند. از این رو لازم است بعنوان پیش پردازش، ویژگی های مناسب از بین این حجم اطلاعات استخراج شود. برای انتخاب ویژگی روش های بسیاری ارائه شده است. در اینجا ما دو الگوریتم تکاملی بهینه سازی کلونی مورچه ها و ژنتیک را برای انتخاب ویژگی متن مورد مقایسه قرار داده ایم. این الگوریتم ها بر روی مجموعه داده رویترز -21578 شبیه سازی شده اند. نتایج به دست آمده نشان دهنده برتری الگوریتم ژنتیک با استفاده از الگوریتم یادگیری Decision Tree در میانگین دقت برابر 72.35%، میانگین فراخوانی برابر 48.53%، میانگین f1 برابر با 49.54% و با استفاده از الگوریتم یادگیری KNN در میانگین دقت برابر 80.41%، میانگین فراخوانی برابر 59.74% و میانگین f1 برابر با 60.33% می باشد. همچنین نمودارهای Macro-f1 و Micro-f1 برای مقایسه الگوریتم های ACO و GA رسم شده است که برتری الگوریتم GA را نشان می دهد.

کلمات کلیدی:

متن کاوی، انتخاب ویژگی، الگوریتم ژنتیک، الگوریتم بهینه سازی کلونی مورچه ها، دسته بندی متون

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/387645>

