

عنوان مقاله:

خوشه بندی اسناد با استفاده از روش کار مبتنی بر افزار داده ها

محل انتشار:

اولین همایش ملی فناوری اطلاعات، ارتباطات و محاسبات نرم (سال: 1395)

تعداد صفحات اصل مقاله: 8

نویسندگان:

سیده نوشین اخوندی - کارشناسی ارشد دانشکده فنی مهندسی دانشگاه آزاد اسلامی علوم و تحقیقات تهران البرز

مجید خلیلیان - استادیار دانشکده فنی و مهندسی دانشگاه آزاد اسلامی کرج

خلاصه مقاله:

بازیابی متون به فن آوری جستجو و استخراج اطلاعات از مجموعه اسناد متنی گفته می شود. طبقه بندی داده های متنی به عنوان یکی از روشها در این راستا می باشد که همراه با چالشهای از قبیل پیچیدگی داده ها، بزرگی چه از نظر تعداد چه از نظر ابعاد و همچنین وجود ساختار سلسله مراتبی برای اکثر داده ای متنی می باشد. برای غلبه بر این مشکلات نیاز به روشهایی است که ضمن برطرف نمودن مشکلات مذکور دقت طبقه بندی را نیز افزایش دهد. در این تحقیق ضمن پیش پردازش داده های متنی و تبدیل آن به ماتریسهای عددی از یک روش دو مرحله ای مبتنی بر افزار بندی برای غلبه بر مشکلات مذکور در داده های متنی استفاده شده است نتایج بدست آمده از اجرای این روش بر روی مجموعه داده متنی در مقایسه با روشهای افرازبندی و سلسله مراتبی اجرا شده و بر روی داده های متنی حای از آن است که دقت طبقه بندی متون نسبت به الگوریتمهای مورد مقایسه افزایش یافته است.

کلمات کلیدی:

الگوریتم k-means خوشه بندی، روش تقسیم و غلبه، روش افزار بندی، مدل فضای برداری

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/517862>

