

## عنوان مقاله:

روش جدید انتخاب ویژگی برای طبقه بندی اسناد فارسی با استفاده از الگوریتم های فراابتکاری

## محل انتشار:

پنجمین کنفرانس ملی محاسبات توزیعی و پردازش داده های بزرگ (سال: 1398)

تعداد صفحات اصل مقاله: 7

## نویسندگان:

پیمان فضلی - کارشناسی ارشد نرم افزار، دانشگاه آزاد اسلامی، میاندوآب، ایران

عادل رضایی - هیات علمی گروه کامپیوتر، دانشگاه آزاد اسلامی، میاندوآب، ایران

## خلاصه مقاله:

امروزه با افزایش روزافزون اطلاعات و گستردگی موضوعات داده‌کاوی یکی از موضوعات مهم درزمینه‌ی پردازش انبوه اطلاعات است. متن کاوی شاخه‌ای از داده‌کاوی است که در آنها متون در دسته‌ها یا گروه‌هایی قرار می‌گیرند. دسته‌بندی متون یکی از مواردی است که اخیراً موردتوجه جوامع علمی قرار گرفته است و کاربردهای بسیاری دارد از جمله مهمترین آنها، فیلترکردن پیام‌های ورودی نرم افزارهای پیام رسان مانند سروش و آی‌گپ میباشد. دسته‌بندی متون فرایندی است که در آن متون براساس وزن کلمات، که یکی از نکات مهم در افزایش دقت و کارایی الگوریتم‌های دسته‌بندی است، در دسته‌هایی قرار می‌گیرند. در این پژوهش، روشی برای دسته‌بندی متون فارسی با استفاده از ضریب همبستگی پیرسون به همراه روش جدید وزندهی Var-TFIDF و الگوریتم فراابتکاری بهینه‌سازی انبوه ذرات (PSO) جهت انتخاب ویژگی بهتر، ارائه شده است. برای شبیه‌سازی از نرم افزار متلب به همراه داده‌های روزنامه همشهری استفاده شده است، نتایج آزمایشها نشان داد که روش جدید وزندهی به همراه استفاده از الگوریتم فراابتکاری نسبت به سایر روشها دارای دقت بالاتری میباشد.

## کلمات کلیدی:

دسته بندی، متن کاوی، وزن دهی، الگوریتم فراابتکاری PSO، ضریب همبستگی پیرسون

## لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/961949>

