

عنوان مقاله:

روشی جدید برای وزن دهی واژه ها در خوشه بندی مستندات

محل انتشار:

اولین همایش ملی فناوری و مدیریت دانش با محوریت اقتصاد مقاومتی (سال: 1393)

تعداد صفحات اصل مقاله: 8

نویسندگان:

سمیه گراوند - دانشجوی ارشد موسسه آموزش عالی پویش

سیداحسان یثربی نایینی - عضو هیات علمی دانشگاه تربت حیدریه

خلاصه مقاله:

بازیابی اطلاعات فرایند یافتن اطلاعات مربوط به جستجوی کاربر در مجموعه مدارک مورد نظر می باشد. با پیاده سازی الگوریتم های متفاوت استراتژی های مختلفی در بازیابی اطلاعات وجود دارد وجه مشترک استراتژی های بازیابی، یافتن مدارک مشابه با موضوع جستجوی کاربر است. یکی از الگوریتم های مهمی که در بازیابی اطلاعات، کاربرد بسیار دارد، الگوریتم های فضای برداری است که می کوشد تمام مدارک را در مجموعه وجستجوهای کاربر به صورت بردارها نشان دهد و ضریب تشابه میان بردارها ی مدارک و بردار جستجو را جهت بازیابی مدارک مربوط محاسبه می کند. در این مقاله یک روش جدید برای وزن دهی واژه ها ارائه شده که بابررسی روش tf و idf و مشکلات آنها و آزمایشاتی که انجام شده روش مناسبی برای وزن دهی واژه ها می باشد.

کلمات کلیدی:

وزن دهی، مدل فضای برداری، $tfidf$ ، tf

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/359045>

