

عنوان مقاله:

ارزیابی و بررسی تاثیر و پیشرفت استخراج اتوماتیک عبارات کلیدی در مجموعه داده ها

محل انتشار:

دوازدهمین کنفرانس ملی مهندسی برق، کامپیوتر و مکانیک (سال: 1400)

تعداد صفحات اصل مقاله: 9

نویسنده:

عادلہ غلامی - کارشناسی ارشد مهندسی کامپیوتر-نرم افزار

خلاصه مقاله:

استخراج کلمه کلیدی که گاهی اوقات تشخیص کلمه کلیدی یا آنالیز کلمه کلیدی نیز نامیده می شود یک تکنیک مورد استفاده برای آنالیز متن است. هدف اصلی این تکنیک این است که به طور خودکار پربسامدترین کلمات و عبارات را از بدنه یک متن استخراج کند. اغلب به عنوان اولین قدم برای خلاصه کردن ایده های اصلی یک متن و ارائه ایده های کلیدی ارائه شده در متن استفاده می شود. استخراج عبارات کلیدی به طور گسترده ای، مورد آزمون قرار گرفته است، پیشرفت کارایی این عمل هنوز بسیار پایین تر از اعمال اصلی پردازش زبان طبیعی است. از طرفی رشد فزاینده اسناد منتهی بر روی وب در مجموعه داده ها، ضرورت استفاده از ابزارهای پردازش خودکار متون را بیش از پیش آشکار می سازد. همچنین در اکثر روش های استخراج کلمات کلیدی در مجموعه داده ها، مفهوم و معنای متن نادیده گرفته می شوند. از طرفی دیگر، بدون ساختار بودن متون جدید در اخبار و اسناد الکترونیکی، استخراج این کلمات را مشکل می سازد. در گذشته جستجوی کلمات کلیدی به صورت دستی انجام می شد ولی محققان روش هایی را بکار بردند که جستجوی کلمات کلیدی و خلاصه سازی آن بصورت اتوماتیک انجام شود. در این تحقیق به بررسی و ارزیابی پیشرفت استخراج کلمات کلیدی اتوماتیک و بررسی منابع اصلی خطا در سیستم های موجود، پرداخته شده و در نهایت به رهیافت مبتنی بر گراف برای استخراج عبارات کلیدی به صورت خودکار بیان می شود.

کلمات کلیدی:

کلمات کلیدی، پردازش زبان، استخراج کلمات، مجموعه داده ها

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/1312177>

