

عنوان مقاله:

ارایه الگوریتم جدید optimized-RAKE برای استخراج عبارات کلیدی

محل انتشار:

کنفرانس ملی کامپیوتر، فناوری اطلاعات و کاربردهای هوش مصنوعی (سال: 1396)

تعداد صفحات اصل مقاله: 20

نویسندگان:

علی محمدی حسن کیاده - دانشجوی کارشناسی ارشد، دانشگاه علم و صنعت ایران

سعید پارسا - دانشیار دانشکده مهندسی کامپیوتر، دانشگاه علم و صنعت ایران

قربان شیخکانلوی میلان - دانشجوی دکترا، دانشگاه آزاد اسلامی واحد شبستر

خلاصه مقاله:

استخراج عبارات کلیدی به عنوان هسته پردازش های خودکاری است که روی داده های متنی انجام می شود. بنابراین در الگوریتم های متنکاوی استخراج عبارات کلیدی از اهمیت ویژه ای برخوردار است. در واقع استخراج عبارات کلیدی یک مرحله بنیادی از اکثر پروژه های متن کاوی است، که تحقیق بر روی آن و پیدا کردن بهترین روش استخراج عبارات کلیدی که دقت و سرعت مناسبی داشته باشد، اهمیت ویژه ای دارد. در این مقاله، الگوریتم جدیدی ارایه میگردد که علاوه بر سرعت بالای استخراج عبارات کلیدی، دارای دقت مناسبی نسبت به سایر الگوریتم های موجود در این زمینه میباشد. در این مقاله یک لیست بهینه و مناسب برای کلمات زاید ارایه میشود که دقت و سرعت حذف کلمات زاید از متن ورودی را افزایش میدهد. همچنین برای استخراج عبارات کلیدی مناسب از رویکرد پردازش زبانی استفاده شده است، که در این رویکرد با دو روش نحوی و لغوی عبارات کاندید مناسب را جهت پردازش و استخراج عبارات کلیدی، مشخص مینماید. علاوه بر بهینه سازی هایی که در بالا ذکر شد، در این مقاله، بهینه سازی هایی روی الگوریتم RAKE انجام داده و الگوریتم جدیدی به نام optimized-RAKE ارایه میگردد. نتایج مقایسه الگوریتم مورد نظر با تعدادی از الگوریتم های مطرح در این حوزه، حاکی از عملکرد بهتر این الگوریتم میباشد

کلمات کلیدی:

عبارت کلیدی، پردازش زبان طبیعی، برچسب بخش کلامی، فراوانی عبارت، متن کاوی، پردازش زبانی، کلمات زاید

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/762492>

