

عنوان مقاله:

انتخاب عنوان برای متون فارسی با استفاده از پردازش زبان طبیعی

محل انتشار:

پنجمین کنفرانس بین المللی پژوهش های کاربردی در علوم و مهندسی (سال: 1399)

تعداد صفحات اصل مقاله: 23

نویسندگان:

کژال سبزهزاده - فارغ التحصیلی کارشناسی ارشد دانشگاه علوم و تحقیقات واحد ایلام، مهندسی کامپیوتر

طیبه سلیمانیان - فارغ التحصیل موسسه آموزش عالی باختر، مهندسی کامپیوتر

خلاصه مقاله:

در این جادو روش جدید به منظور تعیین خودکار عنوان برای متون زبان فارسی ارائه شده است. در روش اول، ابتدا الگوهای رایج برای عناوین در متون فارسی با استفاده از یک پایگاه داده جمع آوری شده و سپس عناوین براساس موضوع متن دسته بندی می شوند. به منظور تعیین عنوان برای متون جدید، در این روش الگوهای پایگاه داده در متن ورودی جستجو می شود. سپس هر عبارت بصورت یک عنوان احتمالی در نظر گرفته می شود و با استفاده از یک معیار امتیازبندی بهترین عنوان انتخاب می شود. در روش دوم نیز تمامی عبارات اسمی با طول بیشتر از چهار کلمه از متن استخراج شده و به صورت عناوین احتمالی در نظر گرفته می شوند. در هر دو روش از یک معیار امتیاز بندی مبتنی برای تعیین میزان مناسب بودن عناوین برای متن استفاده شده است. این معیار، از دو روش TF-IDF و امتیازدهی براساس موقعیت استفاده می کند. نتایج پیاده سازی نشان می دهد که روش های پیشنهاد شده می توانند عناوین مناسبی را برای متون زبان فارسی تولید کنند.

کلمات کلیدی:

پردازش زبان طبیعی، پردازش متن، استخراج عنوان از متن، عبارت اسمی.

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/1128561>

