

عنوان مقاله:

بهبود تصحیح غلط املائی در زبان کردی با تغییر وزندهی در کمترین فاصله ویرایشی

محل انتشار:

دومین کنفرانس ملی محاسبات نرم (سال: 1396)

تعداد صفحات اصل مقاله: 6

نویسندگان:

آرش امانی - دانشجوی کارشناسی ارشد، دانشگاه آزاد واحد علوم و تحقیقات تهران، تهران

عباس کوچاری - استادیار، دانشگاه آزاد واحد علوم و تحقیقات تهران، تهران

خلاصه مقاله:

در این مقاله روشی برای تصحیح خودکار غلط های املائی زبان کردی ارائه شده است در روش پیشنهادی از Soundex برای کدگذاری کلمات و پیشنهاد کلمات کاندید برای تصحیح غلط املائی استفاده شده است برای اولویت دادن به کلمات پیشنهادی نخست فاصله ویرایشی لونشتاین به کارگرفته شد سپس با تغییر در هزینه ی جایگزینی حروف ویژه ی زبان کردی نرخ تصحیح خطا و کارایی سیستم افزایش داده شد برای آزمون الگوریتم ارائه شده یک مجموعه آزمون با تعداد 399 کلمه دارای خطای املائی جمع اوری گردیده است برای ایجاد فرهنگ لغات کردی گزینش کلمات براساس تعداد تکرار وقواعد زبانی انجام گرفته تا از قرارگیری واژه نامعتبر در فرهنگ لغات جلوگیری شود در روش پیشنهادی نسبت به روشهای عمومی برای دیگر زبانها مانند لونشتاین در نرخ تصحیح خطا 3.76 درصد بهبود داشته و وجود کلمه ی درست در اولین پیشنهاد 10.03 درصد بهبود یافته است

کلمات کلیدی:

غلط گیر املائی، زبان کردی، تصحیح خودکار، کمترین فاصله ویرایشی، کدگذاری Soundex

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/696641>

