

عنوان مقاله:

شناسایی کنش گفتار با استفاده از ویژگی های آماری و ساختاری در زبان فارسی

محل انتشار:

یازدهمین کنفرانس سراسری سیستم های هوشمند (سال: 1391)

تعداد صفحات اصل مقاله: 8

نویسندگان:

آرش خوشنود - دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی امیرکبیر

محمد مهدی همایونپور - هیات علمی دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی امیرکبیر

خلاصه مقاله:

هدف از این پژوهش یافتن کنش گفتاری برای استفاده در سامانه های هوشمند پردازش متن فارسی مانند تبدیل متن به گفتار، ترجمه ماشینی، بازیابی اطلاعات، خلاصه سازی متون حساس به کنش های گفتاری می باشد. در این مقاله شش روش مختلف برای انتخاب ویژگی ها در ماشین شناسایی کنش گفتاری مورد آزمایش قرار گرفت. پنج عدد از این الگوریتم ها براساس روش یادگیری ماشین و استخراج ویژگی ها بر مبنای TF-IDF و مقوله واژگانی، آموزش و تست شدند. یک روش نیز براساس ویژگی های ساختاری جملات فارسی، مورد بررسی قرار گرفت. درستی هر شش الگوریتم براساس ارزیابی متقاطع 5 باره 1 ارزیابی شد، با توجه به آزمون های تجربی که بروی یک مجموعه داده واحد انجام گردید، الگوریتمی که ویژگی های آن براساس TF-IDF - Unigram و TF-IDF - bigram و ساختار جملات فارسی پیاده سازی شده است، بیشترین درستی (معادل 93%) را در برداشت.

کلمات کلیدی:

پردازش متن، ترجمه ماشینی، تبدیل متن به گفتار، کنش گفتار، ماشین بردار پشتیبان

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/214584>

