

عنوان مقاله:

ارائه کاربست مقاوم سازی با تندمتن روی مجموعه داده عربی

محل انتشار:

نخستین همایش ملی هوش مصنوعی و علوم اسلامی (سال: 1399)

تعداد صفحات اصل مقاله: 1

نویسندگان:

امیر هاونگی - کارشناسی ارشد، مهندسی کامپیوتر، نرم افزار دانشگاه علم و صنعت ایران

علی اصغر تقی زاده - دکتر، مهندسی کامپیوتر، هوش مصنوعی، دانشگاه علم و صنعت ایران

یوسف علیزاده - دکتر، مهندسی کامپیوتر، هوش مصنوعی، دانشگاه علم و صنعت ایران

بهروز مینایی بیدگلی - دانشیار، (دانشگاه علم و صنعت ایران)، دکتری هوش مصنوعی

خلاصه مقاله:

یکی از حوزه های علم هوش مصنوعی، کار هوشمند بر روی متن می باشد که به پردازش زبان طبیعی معروف است. در این حوزه با پردازش و روش های مختلف روی متون هر چند بدون ساختار و دارای زبان طبیعی، نتایجی هوشمندانه، معتبر و بسیار مفید ارائه می گردد. قریب به اتفاق آزمایش های پردازش زبان طبیعی صورت گرفته، بر روی زبان انگلیسی می باشد. مقاوم سازی مجموعه داده های زبان غیر از زبان انگلیسی، یکی از چالش های این حوزه است. زبان عربی که اکثر کتب و متون علوم اسلامی به این زبان ارائه می گردد از این قاعده مستثنی نیست. پردازش زبان طبیعی روی این زبان و روی کتب و متون اسلامی می تواند نتایج جالب و شگف تانگیز را برای دانشمندان این زمینه به ارمغان آورد. در این مقاله روی مجموعه دادگان عربی، مقاوم سازی صورت گرفته است که به نتایج درخور و مفیدی منجر شده است. تکنولوژی مورد استفاده در این مقاله، کتابخانه تندمتن زبان پایتون می باشد. اقدامات صورت گرفته بر روی این مجموع هدادگان، تولید کلمات مترادف و اعتبارا رسنجی آن و همچنین اصلاح املائی کلمات م بیاشد. امید است که این اقدامات در پیشبرد فعالیت های علمی در حوزه علوم اسلامی موثر واقع گردد.

کلمات کلیدی:

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/1347096>

