

عنوان مقاله:

تحلیل آماری اخبار جعلی فارسی مربوط به کوید-۱۹

محل انتشار:

فصلنامه زبان شناسی اجتماعی، دوره 5، شماره 4 (سال: 1401)

تعداد صفحات اصل مقاله: 18

نویسنده:

مسعود قیومی - پژوهشگر زبان شناسی، پژوهشگاه علوم انسانی و مطالعات فرهنگی

خلاصه مقاله:

در این پژوهش تلاش می شود با استفاده از تحلیل آماری، ویژگی های اخبار جعلی فارسی مربوط به کوید-۱۹ بررسی گردد. برای این هدف، ابتدا یک پیکره زبانی که حاوی اخبار موثق و جعلی در حوزه کرونا است تهیه می شود. سپس الگوهای زبانی این دو دسته داده و همچنین دو تحلیل آماری مقدار اطلاعات و خوانایی اخبار موثق و جعلی مورد بررسی قرار گرفته و با یکدیگر مقایسه می شود. براساس اطلاعات استخراج شده و نتایج عملی به دست آمده از پیکره خبرهای جعلی، الگوهای زبانی مشترک بین این دو دسته داده وجود دارد. همچنین، مقدار اطلاعات در اخبار موثق براساس دو معیار آنتروپی و شگفتی بیشتر از اخبار جعلی است. همچنین سطح خوانایی خبرهای جعلی با استفاده از تساوی های اندازه گیری خوانایی متن مورد ارزیابی قرار گرفت و این نتیجه به دست آمد که اخبار جعلی در مقایسه با اخبار موثق عمدتاً ساده بوده و دشوار نیست. در فرایند برچسب گذاری خودکار خبرهای موثق و جعلی براساس سطح دشواری حجم زیادی از اخبار جعلی ساده تشخیص داده شد. اگرچه سطح دشواری برای تعدادی خبر جعلی تشخیص داده شد، تعدادشان کمتر از اخبار موثق با سطح زبانی دشوار بود. علاوه بر این دستاورد و بررسی آماری ویژگی های زبانی براساس میزان اطلاعات و خوانایی اخبار جعلی، جنبه کاربردی این اطلاعات آماری جهت تشخیص خبر جعلی با استفاده از روش های یادگیری ماشینی مورد مطالعه قرار گرفت.

کلمات کلیدی:

اخبار جعلی فارسی، کوید-۱۹، نظریه اطلاعات، آنتروپی، شگفتی، خوانایی

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/1639736>

