2026, Volume 13, Number 1 (pp. 83--83)
http://www.jcomsec.org

Journal of Computing and Security

تحلیل احساسات مبتنی بر جنبه در زبان فارسی با استفاده از مدل تنظیم‌دقیق‌شده LLaMA3

Persian Abstract

نیلوفر رنجبر [آ،*] حامد باغبانی [آ]

[آ] گروه مهندسی کامپیوتر، دانشکده مهندسی جم، دانشگاه خلیج فارس، بوشهر، ایران.

اطلاعات مقاله

تاریخچه مقاله:
دریافت: 12 July 2025
اصلاح: 28 June 2026
پذیرش: 22 July 2026
انتشار آنلاین: 28 July 2026

کلمات کلیدی:
تحلیل احساسات مبتنی بر جنبه، زبان فارسی، تحلیل احساسات، مدل LLaMA3، مدل‌های زبانی بزرگ، پردازش زبان طبیعی.

چکیده

تحلیل احساسات مبتنی بر جنبه در زبان فارسی، در مقایسه با زبان انگلیسی، کمتر توسعه یافته است؛ به‌ویژه هنگامی که هدف، استخراج ساختارهای کاملِ جنبه، دسته، عبارت نظر و احساس باشد، نه صرفاً طبقه‌بندی احساس برای جنبه‌های از پیش تعیین‌شده. این مقاله یک چارچوب مولد برای تحلیل احساسات مبتنی بر جنبه فارسی بر پایه مدل‌های دستورمحورِ تنظیم‌دقیق‌شده خانواده LLaMA3 ارائه می‌کند. مدل با دریافت یک نقد فارسی، آرایه‌ای JSON از چهارتایی‌های شامل عبارت جنبه، دسته، عبارت نظر و برچسب احساس تولید می‌کند. داده‌های آموزشی از یک پیکره تحلیل احساسات مبتنی بر جنبه فارسی با پیش‌برچسب‌گذاری GPT-3.5 گرفته شده‌اند که برچسب‌های آن توسط ارزیابان انسانی بررسی و پالایش شده‌اند. ازاین‌رو، در این مقاله از آن با عنوان پیکره‌ای کمک‌گرفته از مدل زبانی و اعتبارسنجی‌شده توسط انسان یاد می‌کنیم، نه یک مجموعه کاملاً انسان‌نویس. برای کاهش خطر ارزیابی دوری، یک مجموعه آزمون مستقل با برچسب‌گذاری صرفاً انسانی نیز از نقدهای خام فارسی ساخته شد. سه ارزیاب، ۲۴۰ نقد را از ابتدا برچسب‌گذاری کردند و توافق دست‌کم دو نفر از سه نفر، ۵۵۵ برچسب طلایی در سطح جنبه از ۲۳۹ نقد ایجاد کرد. ما LLaMA3 در حالت صفرنمونه، مدل‌های LLaMA3 تنظیم‌دقیق‌شده، و Dorna-LLaMA3-8B-Instruct را به‌عنوان یک خط‌مبنای مولد فارسی‌محور، با طرح خروجی، تجزیه‌گر و معیارهای یکسان ارزیابی می‌کنیم. در مجموعه آزمون صرفاً انسانی، LLaMA3 صفرنمونه به امتیاز F1 استخراج جنبه برابر با 0.1426 و امتیاز F1 جنبه و احساس برابر با 0.1185 می‌رسد. میانگین پنج اجرای LLaMA3-CG این مقادیر را به‌ترتیب به 0.5477±0.0229 و 0.4394±0.0212 افزایش می‌دهد، در حالی که Dorna-CG به 0.5809±0.0123 و 0.4897±0.0194 می‌رسد. افزودن ParsiNLU عملکرد درون‌دامنه‌ای ParsiNLU را به‌طور چشمگیری بهبود می‌دهد، اما عملکرد روی مجموعه صرفاً انسانی را کاهش می‌دهد و نوعی موازنه میان دامنه و سبک برچسب‌گذاری را آشکار می‌کند.

* نویسنده مسئول.
آدرس‌های رایانامه: nranjbar@pgu.ac.ir (ن. رنجبر)،
baghbani.hamed@gmail.com (ح. باغبانی)
 ISSN: 2322-4460