



SimCat: انتخاب پاسخ آگاه به دسته و مبتنی بر شباهت در سیستم‌های پرسش پاسخ فارسی

نگین آبادانی آ^{*}، افسانه فاطمی آ^{*}، محمدعلی نعمت بخش آ

آ دانشکده مهندسی کامپیوتر، دانشگاه اصفهان، اصفهان، ایران.

Persian
Abstract

چکیده

اطلاعات مقاله

تاریخچه مقاله:

دریافت: 24 July 2024

اصلاح: 31 October 2024

پذیرش: 10 December 2024

انتشار آنلاین: 22 June 2025

کلمات کلیدی:

انتخاب پاسخ، پرسش پاسخ فارسی، پرسش
واقع‌نما، پردازش زبان طبیعی، یادگیری عمیق.

انتخاب پاسخ یکی از وظایف اصلی سیستم‌های پرسش پاسخ (QA) است که هدف آن یافتن مرتبط‌ترین جمله با توجه به پرسش در میان مجموعه‌ای از جملات است؛ هدف آن رتبه‌بند پاسخ‌های کاندید بر اساس ارتباط و شباهت آن‌ها با پرسش و یافتن پاسخی نهایی است. تمام پژوهش‌های انجام شده در این زمینه تا به امروز عمدتاً بر روی زبان انگلیسی متمرکز بوده و هیچ تحقیقی در مورد انتخاب پاسخ دامنه باز به زبان فارسی وجود ندارد؛ یکی از دلایل اصلی آن، عدم وجود مجموعه داده‌های انتخاب پاسخ دامنه باز فارسی است. در این مقاله، یک رویکرد دسته‌بندی مبتنی بر شباهت و آگاه به دسته برای انتخاب پاسخ معرفی نموده، اثربخشی محاسبه شباهت جملات از جنبه‌های مختلف (لغوی، نحوی و معنایی)، به جای یک جنبه را تحلیل کرده و سپس این رویکرد را بر روی سه مجموعه داده‌ی انگلیسی محک متفاوت و چهار مجموعه داده‌ی جدید که برای انتخاب پاسخ واقع‌نما دامنه باز در زبان فارسی ایجاد شده‌اند، ارزیابی می‌کنیم. علاوه بر بهبود دقت انتخاب پاسخ، با حذف جملات کاندید نامرتبط بر اساس دسته‌ی پرسش و دسته‌ی پاسخ کاندید، زمان مورد نیاز برای فرایند انتخاب پاسخ را نیز کاهش داده‌ایم. پس از اجرا و ارزیابی این رویکرد برای دو زبان انگلیسی و فارسی، مشخص شد که رویکرد پیشنهادی انتخاب پاسخ را از نظر MAP و MRR به طور متوسط به ترتیب ۱٪ و ۸٪ برای انگلیسی، ۳٪ و ۲٪ برای فارسی بهبود داده است. علاوه بر این، زمان مورد نیاز را به طور متوسط ۷۹٪ برای انگلیسی و ۶۹٪ برای فارسی کاهش داده است.

© 2025. تمامی حقوق محفوظ است.

* نویسنده مسئول.

آدرس رایانامه: a_fatemi@eng.ui.ac.ir (آ. فاطمی)

تمامی حقوق محفوظ است. ISSN: 2322-4460 © 2025

