چکیده مبسوط
(چکیده توسعه یافته Extended Abstract)

عنوان: پرسش و پاسخ محاوره‌ای برای زبان‌های کم‌منبع: یک معماری جدیدِ تقویت‌شده با مدل‌های زبانی بزرگ

نام نویسندگان: صفورا آقاداود جلفایی، پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)،
aghadavood@students.irandoc.ac.ir

آزاده محبی، پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)،
mohebi@irandoc.ac.ir

مقدمه :در عصر کنونی، سیستم‌های پرسش و پاسخ محاوره‌ای به مؤلفه‌ای کلیدی در هوش مصنوعی تبدیل شده‌اند، اما شکافی عمیق میان پیشرفت‌ها در زبان‌های پرمنبع (مانند انگلیسی) و نیازهای اساسی زبان‌های کم‌منبع (مانند فارسی) وجود دارد .مسئله اصلی این پژوهش، ناکارآمدی معماری‌های موجود در مواجهه با چالش‌هایی نظیر کمبود داده‌های آموزشی، پیچیدگی‌های ساختاری و ظرافت‌های فرهنگی در زبان‌های کم‌منبع است .هدف اصلی این پژوهش، طراحی و اعتبارسنجی یک چارچوب جامع و ماژولار است که با یکپارچه‌سازی پردازش‌های زبانی و فرهنگی، امکان تولید پاسخ‌های دقیق و متناسب با زمینه را فراهم آورد .این پژوهش به دنبال پاسخ به این پرسش است که مولفه‌های ضروری برای چنین سیستمی کدامند و چگونه می‌توان بر محدودیت‌های محاسباتی و داده‌ای در این زبان‌ها غلبه کرد.

روش‌شناسی :پژوهش حاضر با رویکردی کیفی و توسعه‌ای و در دو مرحله اصلی انجام شد .جامعه پژوهش شامل متون علمی مرتبط و متخصصان حوزه پردازش زبان طبیعی بود .در مرحله اول (طراحی چارچوب)، با استفاده از روش تحلیل شکاف عملکردی و مرور ساختاریافته بیش از ۶۰ مقاله کلیدی، معماری‌های موجود بررسی و چالش‌های زبان‌های کم‌منبع استخراج گردید .در مرحله دوم (اعتبارسنجی)، از روش کیفی «گروه متمرکز» بهره گرفته شد .شرکت‌کنندگان شامل پنج پژوهشگر متخصص با مدرک دکتری در هوش مصنوعی و سابقه پژوهشی در زبان‌های کم‌منبع بودند .ابزار گردآوری داده‌ها، مصاحبه‌های نیمه‌ساختاریافته بر اساس پرسشنامه‌ای دقیق بود و داده‌های حاصل از بحث‌های عمیق پیرامون مولفه‌های چارچوب، مورد تجزیه و تحلیل کیفی قرار گرفت تا کارایی آن‌ها در مواجهه با چالش‌های شناسایی‌شده سنجیده شود.

یافته‌های اصلی :یافته‌های پژوهش منجر به ارائه یک چارچوب نوین شش‌مؤلفه‌ای شد که شامل «پردازش ورودی»، «هسته مدل زبانی بزرگ»، «تقویت دانش»، «مدیریت زمینه»، «تولید پاسخ» و «بازخورد انسانی» بود .نتایج نشان داد که مؤلفه پردازش ورودی با تشخیص زبان و درک پرس‌وجوی میان‌زبانی، بر چالش کمبود منابع فائق آمد. همچنین مشخص شد که استفاده از روش‌های انطباق با پارامتر بهینه در هسته مدل، نیازهای محاسباتی را به طور چشمگیری کاهش داد .یافته‌ها حاکی از آن بود که بر خلاف سیستم‌های پیشین، ادغام «انطباق فرهنگی» به عنوان یک مرحله میانی در تولید پاسخ، منجر به خروجی‌هایی می شود که از نظر سطوح رسمی بودن و هنجارهای اجتماعی دقیق‌تر هستند. ارزیابی متخصصان نیز تأیید کرد که معماری ماژولار و استفاده از مدیریت زمینه فشرده،

گلوگاه‌های حافظه و پردازش را در محیط‌های محدود برطرف می کند و سیستم می تواند با حداقل داده‌های آموزشی، عملکرد قابل قبولی ارائه دهد.

بحث و نتیجه‌گیری: این پژوهش نشان می‌دهد که گذار از رویکرد پس‌پردازش به «پردازش فرهنگی یکپارچه» در تمام مراحل، ارزش افزوده‌ای حیاتی برای پذیرش سیستم‌ها در زبان‌های کم‌منبع ایجاد می‌کند .چارچوب پیشنهادی با ترکیب نوآورانه دانش میان‌زبانی و بازخورد انسانی، سیستمی تطبیق‌پذیر را ارائه می‌دهد که به مرور زمان دقت فرهنگی خود را بهبود می‌بخشد .استنباط می‌شود که این معماری می‌تواند الگویی برای سایر زبان‌های با منابع محدود باشد و توسعه هوش مصنوعی را دموکراتیک‌تر کند . با این حال، محدودیت‌هایی در ارزیابی کمیِ دقت فرهنگی وجود دارد .پیشنهاد می‌شود پژوهش‌های آتی بر توسعه معیارهای استاندارد و کمی برای سنجش «تناسب فرهنگی» تمرکز کنند و این چارچوب را بر روی زبان‌های دیگر نیز پیاده‌سازی و آزمون نمایند.

کلیدواژه‌ها :پرسش و پاسخ محاوره‌ای، زبان‌های کم‌منبع، مدل‌های زبانی بزرگ، انطباق فرهنگی، پردازش زبان طبیعی.

Suggested format/template for extended abstract

Title: Conversational Question Answering for Low-Resource Languages: A Novel Architecture Enhanced by Large Language Models

Names of Authors: First Author; Safoura Aghadavoud Jolfaei, PHD student in ranian Research Institute for Information Science and Technology (IranDoc), aghadavood@students.irandoc.ac.ir

Introduction: Conversational Question Answering (CQA) systems have become a pivotal component of Artificial Intelligence; however, a significant gap remains between advancements in high-resource languages and the fundamental needs of low-resource languages. This research addresses the inefficiency of existing architectures in handling unique challenges such as data scarcity, morphological complexity, and cultural nuances in languages like Persian. The primary objective of this study is to design and validate a comprehensive, modular framework that integrates linguistic and cultural processing to generate accurate, context-aware responses. This research investigates the essential components of such a system and explores how to effectively overcome computational and data limitations to enhance CQA capabilities for under-represented languages.

Methodology: This research employed a qualitative and developmental approach, conducted in two main phases. The study population encompassed relevant scientific literature and experts in Natural Language Processing (NLP). In the first phase (Framework Design), a functional gap analysis and a structured review of over 60 key papers were performed to examine existing architectures and extract specific challenges related to low-resource languages. In the second phase (Validation), the qualitative "Focus Group" method was utilized. The participants included five expert researchers holding PhDs in Artificial Intelligence with specific experience in low-resource languages. Data were collected through semi-structured interviews based on a detailed questionnaire and deep discussions regarding the framework's components. The gathered data were qualitatively analyzed to assess the efficacy of the proposed components in addressing the identified linguistic and cultural challenges.

Main Findings: The study resulted in the development of a novel six-component framework comprising Input Processing, LLM Core, Knowledge Enhancement, Context Management, Response Generation, and Human Feedback. The findings indicated that the Input Processing component successfully addressed resource scarcity through language detection and cross-lingual query understanding. Furthermore, the implementation of Parameter-Efficient Adaptation methods (such as LoRA) in the LLM Core significantly reduced computational requirements. A key finding was that, unlike previous systems, the integration of "Cultural Adaptation" as an intermediate step in response generation led to outputs that were more accurate regarding formality levels and social norms. Expert evaluation confirmed that the modular architecture and compressed context management resolved memory and processing bottlenecks, allowing the system to perform effectively with minimal training data.

Discussion and Conclusions: This research demonstrates that shifting from post-processing to "integrated cultural processing" throughout the pipeline provides critical value for system adoption in low-resource languages. The proposed framework, by innovatively combining

cross-lingual knowledge and human feedback, offers an adaptive system that improves its cultural accuracy over time. It is concluded that this architecture can serve as a model for other resource-constrained languages, effectively democratizing AI development. However, limitations exist regarding the quantitative evaluation of cultural accuracy. Future research should focus on developing standardized metrics for "cultural fit" and testing this framework on extremely low-resource languages to further validate its scalability and robustness.

Keywords: Conversational Question Answering, Low-Resource Languages, Large Language Models (LLMs), Cultural Adaptation, Natural Language Processing.