بهینه‌سازی سامانه‌های پرسش و پاسخ با استفاده از بهینه‌سازی ازدحام ذرات تسریع شده

نسیم توحیدی و سید محمدحسین هاشمی‌نژاد*
گروه مهندسی کامپیوتر، دانشگاه الزهرا (س)، تهران، ایران

چکیده

سامانه‌های پرسش و پاسخ، موتورهای جستجویی هستند که توانایی ارائهٔ پاسخی کوتاه و دقیق را به یک پرسش دارند. به عبارت دیگر، پرسشی که یک موتور جستجو، با مجموعه‌ای از اسناد پاسخ می‌دهد، یک ساماهٔ پرسش و پاسخ، با یک پاراگراف، جمله یا کلمه پاسخ می‌دهد. در این مقاله، یک راه‌کار برای بهینه‌سازی عملکرد سامانه‌های پرسش و پاسخ تک‌زبانه به زبان انگلیسی و مبتنی بر وب، ارائه شده است. با توجه به اینکه الگوریتم‌های تکاملی برای مسائل با فضای جستجوی بزرگ مناسب هستند، در این مقاله برای بهینه‌سازی عملکرد این سامانه‌ها، APSO مورد استفاده قرار گرفته است. در این پژوهش هدف، ارائه روشی است که دقت و سرعت بالاتری نسبت به سامانه‌های موجود در انتخاب پاسخ از میان اسناد بازیابی شده داشته باشد. روش پیشنهادی بر روی مجموعه‌داده استاندارد به میزان دقتی (Top1 Accuracy) برابر با ۰/۵۲۷ دست یافته است و همچنین شاخص MRR نتایج حاصل از آن برابر ۰/۷۱۱ محاسبه شده است، که این نتایج نسبت به بیش‌تر پژوهش‌های مرتبط پیشرفت داشته‌اند. در عین حال، سرعت آن نسبت به همه کارهای مشابه بهبود یافته است.

واژگان کلیدی: سامانه پرسش و پاسخ، پردازش زبان طبیعی، الگوریتم بهینه‌سازی ازدحام ذرات تسریع‌شده

Optimizing question answering systems by Accelerated Particle Swarm Optimization (APSO)

Nasim Tohidi & Seyed Mohammad Hossein Hasheminejad*
Faculty of Engineering and Technology, Alzahra University, Tehran, Iran

Abstract

One of the most important research areas in natural language processing is Question Answering Systems (QASs). Existing search engines, with Google at the top, have many remarkable capabilities. However, there is a basic limitation; search engines do not have deduction capability which a QAS is expected to have. In this perspective, a search engine may be viewed as a semi-mechanized QAS. Upgrading a search engine such to a QAS is a task whose complexity is hard to exaggerate. To achieve success, new concepts and ideas are needed to address difficult problems which arise when knowledge has to be dealt with in an environment of imprecision, uncertainty and partial truth .

QASs are search engines that have the ability to provide a brief and accurate answer to each question in natural language for instance, the question that a search engine answers with a set of documents, a QAS answers with a paragraph, sentence or etc. In this paper, a solution is proposed to optimize the performance and speed of web-based QASs for answering English questions .

As evolutionary algorithms are suitable for issues with large search space, in this approach we have used an evolutionary algorithm to optimize QASs. In this regard, we have chosen APSO which is a simplified

* نویسندهٔ عهده‌دار مکاتبات
* Corresponding author

• تاریخ ارسال مقاله: ۱۳۹۸/۹/۲۳ • تاریخ پذیرش: ۱۳۹۹/۱۱/۱۵ • تاریخ انتشار: ۱۴۰۱/۷/۷ • نوع مطالعه: پژوهشی

[DOI: 10.52547/jsdp.19.2.161]

version of PSO. The proposed method consists of five main stages: question analysis, pre-process, retrieval, extraction and ranking. We have tried to provide a method that would be more accurate in choosing the most probable answer from the documents that have been retrieved by the standard search engine and at the same time, be faster than similar methods. In ranking process, various attributes can be extracted from the text that are used in APSO. For this purpose, in addition to selecting a sentence from the text and examining its attributes, different cut parts of the sentence are selected each time by changing the beginning and end points of the cut part. The attributes which have been used in this study are: 1. Number of unigrams similar to the question words, 2. Number of bigrams similar to the question words, 3. Number of unigrams similar to the question words in the cut part, 4. Number of bigrams similar to the question words in the cut part, 5. Number of synonyms with the question words and 6. Number of synonyms with the question words in the cut part. The fitness function is the weighted sum of these attributes.

Top-1 accuracy and MRR are the most valid metrics for measuring the performance of QASs. The proposed method has achieved the accuracy (top-1 accuracy) of 0.527 with respect to the standard dataset and the MRR of it, is 0.711. Both of these results are improved compared to most similar systems. In addition, the time taken to answer the input question in the proposed method, has been significantly reduced compared to similar methods. In general, the accuracy and MRR in this paper have progressed and the system needs less time to find the answer, in comparison with existing QASs.

Keywords: question answering system, natural language processing, accelerated particle swarm optimization (APSO)

۱- مقدمه

یک سامانه پرسش و پاسخ، به کاربران این امکان را می‌دهد که به‌طور مستقیم پرسش خود را به زبان طبیعی به سامانه بدهند و پاسخی دقیق براساس دانش نهفته در مستنداتی خارج از سامانه دریافت کنند. بر این اساس، این قبیل سامانه‌ها نسبت به سامانه‌های معمول بازیابی اطلاعات، با روش‌های پیچیده‌تر پردازش زبان طبیعی (NLP)[1] سر و کار دارند و در محافل علمی به‌عنوان نسل آینده موتورهای جستجوی اطلاعات، مطرح هستند [1]. در این راستا نیاز به تحلیل‌های زبانی در سطوح مختلف وجود خواهد داشت. سامانه‌های پرسش و پاسخ، با انواع مختلفی از پرسش‌های کاربران روبه‌رو هستند، برای مثال اطلاعات در مورد حقایق، ارائه تعاریف مفاهیم، چگونگی و چرایی مسائل مختلف، تئوری‌ها، پرسش‌های منطقی (بولین) و نیز پرسش‌های بین زبانی. مجموعه اسناد مورد جستجو نیز می‌تواند در مقیاس‌های مختلفی از جمله مجموعه اسناد یک سازمان تا اطلاعات موجود در محیط وب، متغیر باشد. به‌طورکلی سامانه‌های پرسش و پاسخ را از ابعاد متفاوتی (دامنه، زبان، مأخذ پاسخ‌ها و ماهیت پرسش‌گران) می‌توان بررسی کرد [2].

پرسش‌های کاربران نیز به دو دسته تقسیم می‌شود. گروه نخست پرسش‌هایی است که پاسخ آن‌ها به‌طور صریح و مستقیم در مجموعه اطلاعات قابل دسترس سامانهٔ موجود است و گروه دوم که سامانه برای پاسخ به آن‌ها نیازمند استدلال است. از سوی دیگر، کیفیت عملکرد یک سامانهٔ پرسش و پاسخ به‌شدت وابسته به اسناد و اطلاعات در دسترس آن سامانه است. این مسأله ممکن است موجب تبادر این مطلب در ذهن شود که هر چه‌قدر حجم اطلاعات در دسترس بیشتر باشد، به کیفیت بهتری دست خواهیم یافت؛ درحالی‌که، در حجم‌های زیاد اطلاعات، مسألهٔ دیگری به‌وجود می‌آید که عبارت است از هم‌پوشانی و تکرار اطلاعات. بدین معنی که اطلاعات خاصی، به روش‌های گوناگون، در بخش‌های مختلف اسناد، وجود خواهند داشت [3].

سامانه‌های پرسش و پاسخ مبتنی بر وب یک جستجو با دامنه بسیار وسیع در وب انجام می‌دهند. از آنجایی که الگوریتم‌های تکاملی برای مسائل با فضای جستجوی بزرگ مناسب هستند [4]، می‌توانند نقش بسزایی در ساختار این سامانه‌ها داشته باشند.

۲- طرح مسأله

در دهه‌های اخیر حجم اطلاعات به حدی افزایش یافته که پیداکردن اطلاعات مورد نیاز از انبوه داده‌های موجود کاری دشوار شده است؛ علاوه‌برآن، ظهور وب از طرفی باعث شده تا اطلاعات، راحت‌تر و سریع‌تر در اختیار اشخاص قرار بگیرد و از طرف دیگر به هر کاربر اینترنت اجازه داده است تا به یک تولیدکننده و ناشر اطلاعات تبدیل شود. اگر چه گفته می‌شود که اطلاعات قدرت است ولی اکنون بیش از پیش آشکار می‌شود که اگر نتوان اطلاعات مفید را از انبوه داده‌ها جدا کرد، اطلاعات زیاد باعث سرگردانی می‌شوند؛ به همین منظور، ابزارهایی خودکار برای دسترسی به اطلاعات ایجاد شده‌اند. سامانه‌های متداول فعلی بازیابی اطلاعات (IR)[2]، براساس یک یا چندکلمه کلیدی که کاربر وارد می‌کند، تعدادی

[1] Natural Language Processing

[2] Information Retrival

[DOI: 10.52547/jsdp.19.2.161]

سند را برمی‌گرداند [5]. این سامانه‌ها چند مشکل عمده دارند. نخست این‌که کاربران به‌طورمعمول تمایل دارند به جای تعدادی کلمه کلیدی، پرسش خود را به‌طورمستقیم وارد کنند. از طرفی، به‌طورمعمول کاربران برای تبدیل پرسش خود به کلمات کلیدی مناسب، با مشکل مواجه هستند و این تبدیل، مستلزم مهارتی است که باید در طول زمان کسب شود. علاوه‌براین، به‌طوراصولی چند کلمه کلیدی به‌تنهایی نمی‌توانند منظور کاربر را به شکل واضح بیان کنند. بسیاری اوقات حتی اگر کاربر مطمئن باشد که یک سند در مجموعهٔ متون وجود دارد که نیاز اطلاعاتی او را برآورده می‌کند، ممکن است نتواند پرس‌وجوی[1] مناسب را برای سامانه بیان کند. درمجموع استفاده از کلمات کلیدی روش مطلوبی برای برقراری ارتباط بین سامانه و کاربر نیست. ازطرف دیگر، کاربران به‌دنبال پاسخ صریح هستند؛ درحالی‌که خروجی این سامانه‌ها تعداد زیادی سند است که ممکن است پاسخ را در خود داشته باشند. بدین ترتیب کاربر برای پیداکردن پاسخ مورد نظر خود مجبور است تعدادی سند را بخواند.

در سال‌های اخیر، سامانه‌های پرسش و پاسخ به‌عنوان یکی از ابزارهای مهم برای دسترسی به اطلاعات تکامل پیدا کرده‌اند. برخلاف سامانه‌های بازیابی اطلاعات، هدف سامانه‌های پرسش و پاسخ این است که به‌طورمستقیم از طریق زبان طبیعی با کاربر ارتباط برقرار کنند. به این شکل که آن‌ها پرسش کاربر را به زبان طبیعی در ورودی دریافت کرده و پاسخ را به‌صورت دقیق به کاربر برمی‌گردانند؛ بنابراین لازم نیست پرسش ابتدا به چندین کلمه کلیدی تبدیل شود و بعد از بازیابی لازم باشد تا کاربر تعداد نامعلومی سند را مطالعه کند تا به پاسخ خود برسد.

از طرفی برای انتخاب یک پاسخ از یک مجموعه از نامزدها، آن‌ها باید اولویت‌دهی شوند [6]. مسأله اولویت‌دهی شامل محاسبه یک امتیاز مبتنی بر مجموعه‌ای از ویژگی‌هاست. از آن‌جایی که ویژگی‌های مربوط به متون متعددند و می‌توان آن‌ها را در گروه‌های مختلفی طبقه‌بندی کرد، یافتن راهی برای به‌دست‌آوردن یک امتیاز مبتنی بر این ویژگی‌ها یک مسألهٔ کلیدی است. هدف از انجام این پژوهش، ارائه راه‌کارهایی مناسب برای حل چالش‌های مطرح شده است.

در بخش ۳ به معرفی الگوریتم بهینه‌سازی ازدحام ذرات می‌پردازیم که در این پژوهش مورد استفاده قرار گرفته است و جزئیات مراحل آن را تشریح می‌کنیم. سپس، در بخش ۴ به کارهای مشابه انجام‌شده در گذشته اشاره خواهد شد. بخش ۵، شامل اطلاعات تکمیلی در مورد سامانه‌های پرسش و پاسخ است. صورت کلی روش پیشنهادی و بخش‌های مختلف آن در بخش ۶ به‌تفصیل شرح داده خواهد شد. در بخش‌های ۷ و ۸ به معرفی داده آموزشی و شاخص‌های ارزیابی می‌پردازیم. جزئیات مربوط به چگونگی ارزیابی روش مطرح‌شده در این پژوهش، در بخش ۹ بیان خواهند شد. درنهایت، بخش ۱۰ به نتیجه‌گیری و توسعه‌های آینده ممکن برای بهبود کارایی روش پیشنهادی خواهد پرداخت.

۳- APSO

الگوریتم بهینه‌سازی ازدحام ذرات یک الگوریتم جستجوی مبتنی بر جمعیت است که رفتار اجتماعی دسته پرندگان را شبیه‌سازی می‌کند و نخستین‌بار توسط کندی و ابرهارت[2] در سال ۱۹۹۵ ارائه شده است [4]. کشف الگوهایی که توانایی پرواز هماهنگ، تغییر مسیر ناگهانی و گروه‌بندی مجدد با بهینه‌ترین ترکیب را به پرندگان می‌دهد، هدف اصلی پژوهشی بوده که درنهایت منجر به ایجاد الگوریتم بهینه‌سازی کارا و ساده PSO شده است [7].

در PSO، ذرات در فضای چند بعدی جستجو حرکت می‌کنند. تغییر در موقعیت ذرات در فضای جستجو مبتنی بر رفتار اجتماعی موجودات در تقلید از موفقیت سایر موجودات است. تغییرات یک ذره در ازدحام، تحت تاثیر تجربیات خود و یا دانش همسایگانش است؛ بنابراین رفتار جستجوی یک ذره از ازدحام تحت تأثیر سایر ذرات در ازدحام است.

بهینه‌سازی ازدحام ذرات بیشتر برای مسائل بهینه‌سازی با پارامترهای پیوسته به کار رفته و یکی از نخستین کاربردهای این الگوریتم برای آموزش شبکه عصبی پیشرو بوده است. مطالعات نشان داده‌اند که کارایی بهینه‌سازی ازدحام ذرات برای آموزش شبکه‌های عصبی بسیار بالاست؛ بنابراین، پژوهش‌های زیادی برای بررسی توانایی‌های بهینه‌سازی ازدحام ذرات به‌عنوان یک الگوریتم آموزشی برای معماری‌های مختلف شبکه‌های عصبی انجام شده است. نتایج حاصل از این پژوهش‌ها نشان داده‌اند که استفاده از PSO، منجر به افزایش دقت می‌شود [7]. الگوریتم PSO در مسائل خوشه‌بندی نیز کاربرد فراوان دارد [8].

[1] Query

[2] Kennedy & Eberhart

[DOI: 10.52547/jsdp.19.2.161]

۱-۳- بهینه‌سازی ازدحام ذرات (PSO)

در الگوریتم PSO، هر ذره، نمایش‌دهنده یک راه حل است. در مقایسه با نمونه‌های محاسباتی تکاملی، ازدحام، معادل جمعیت و یک ذره، معادل یک موجود است [7]. $x_i(t)$ مشخص‌کننده موقعیت ذره i در فضای جستجو در گام زمانی t است؛ با توجه به فرمول (۱) موقعیت ذره با اضافه‌کردن سرعت $v_i(t)$، به موقعیت کنونی تغییر می‌کند.

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad (۱)$$

بردار سرعت، علاوه‌بر هدایت فرایند بهینه‌سازی، دانش تجربی ذره و نیز اطلاعات تغییرات ذرات همسایه را بازتاب می‌کند.

۱-۱-۳- پارامترهای بهینه‌سازی ازدحام ذرات

در PSO تمام ذرات یک مقدار شایستگی دارند که به‌وسیلهٔ تابع شایستگی که باید بهینه شود ارزیابی می‌شود. علاوه‌بر این هر ذره i، دارای یک موقعیت در فضای D-بعدی مسأله است که که در تکرار tام، با یک بردار مانند رابطه (۲) نمایش داده می‌شود [7]:

$$X_i^t = (x^t_{i1}, x^t_{i2}, ..., x^t_{iD}) \quad (۲)$$

سرعت ذره در تکرار tام با بردار (۳) نشان داده می‌شود:

$$V_i^t = (v^t_{i1}, v^t_{i2}, ..., v^t_{iD}) \quad (۳)$$

این ذره در هر تکرار یک حافظه از بهترین موقعیت قبلی خودش را دارد که با بردار P مانند رابطه (۴) نشان داده می‌شود:

$$P_i^t = (p^t_{i1}, p^t_{i2}, ..., p^t_{iD}) \quad (۴)$$

در هر تکرار جستجو، هر عضو با در نظر داشتن دو مقدار بهترین به‌روزرسانی می‌شود: Pbest مربوط به بهترین راه حلی است که ذره تاکنون آن را تجربه کرده و Gbest بهترین موقعیتی است که تاکنون در جمعیت به‌دست آمده است [7]. بعد از این‌که به تعداد دو بهترین مقدار پیدا شدند موقعیت و سرعت هر عضو به‌وسیلهٔ فرمول‌های (۵) و (۶) به‌روزرسانی می‌شوند:

$$v_i(t+1) = wv_i(t) + c_1 r_{1.i}(t)\big(P_i(t) - X_i(t)\big) + c_2 r_{2.i}(t)\left(P_g(t) - X_i(t)\right) \quad (۵)$$

$$X_i(t+1) = X_i(t) + V_i(t+1) \quad (۶)$$

در فرمول‌های بالا t بیان‌گر شماره تکرار و متغیرهای c_1 و c_2 فاکتورهای یادگیری هستند. r_1 و r_2 دو عدد تصادفی یکنواخت در بازه [۰،۱] هستند. w وزنی‌ست که در بازه [۰،۱] مقداردهی اولیه می‌شود. یک w بزرگ یک اکتشاف عمومی و وزن کوچک اکتشاف محلی را تسهیل می‌کند.

سمت راست معادله سرعت از سه قسمت تشکیل شده که قسمت نخست، سرعت فعلی ذره است و قسمت‌های دوم و سوم تغییر سرعت ذره و چرخش آن به سمت بهترین تجربه شخصی و بهترین تجربه گروه را به عهده دارند. اگر قسمت نخست را در این معادله در نظر نگیریم، آن‌گاه سرعت ذرات تنها با توجه به موقعیت فعلی و بهترین تجربه ذره و بهترین تجربه جمع تعیین می‌شود. به این ترتیب، بهترین ذره جمعیت، در جای خود ثابت می‌ماند و سایرین به سمت آن ذره حرکت می‌کنند. در واقع حرکت دسته جمعی ذرات بدون قسمت اول معادله سرعت، فرایندی خواهد بود که طی آن فضای جستجو به‌تدریج کوچک می‌شود و جستجویی محلی حول بهترین ذره شکل می‌گیرد. در مقابل اگر فقط قسمت نخست معادله سرعت را درنظر بگیریم، ذرات راه عادی خود را طی می‌کنند تا به انتهای محدوده برسند و به نوعی جستجویی سراسری را انجام می‌دهند [7].

در الگوریتم PSO استاندارد جمعیت به‌صورت تصادفی مقداردهی اولیه می‌شود و تا رسیدن به شرط خاتمه به‌صورت تکراری شایستگی جمعیت محاسبه، مقادیر Pbest و Gbest، سرعت و موقعیت نیز به‌ترتیب به‌روزرسانی می‌شوند. در آخر هم Gbest و مقدار شایستگی‌اش به‌عنوان خروجی بیان می‌شوند. شرط خاتمه می‌تواند رسیدن به بیشینه تعداد نسل‌ها یا رسیدن به یک مقدار خاص شایستگی در Gbest باشد.

۲-۱-۳- مزایا و معایب

الگوریتم بهینه‌سازی ازدحام ذرات یکی از پرکاربردترین الگوریتم‌های تکاملی است. از جمله مزایای این الگوریتم می‌توان به موارد زیر اشاره کرد [7]:

- ✓ پیاده‌سازی آن نسبت ساده است؛
- ✓ جمعیت آن کم و بنابراین مقداردهی اولیه راحت‌تر است؛
- ✓ تعداد پارامترها کم و درنتیجه حجم حافظه کم‌تر است.
- ✓ احتمال کمی دارد که در بهینه محلی به دام بیافتد. (وابسته به تعیین مناسب پارامترهاست.)

[DOI: 10.52547/jsdp.19.2.161]

همچنین یکی از معایب آن این است که سرعت ذرات با افزایش تعداد تکرار کاهش می‌یابد و به بهترین نتیجه‌ای که تاکنون یافته‌اند هم‌گرا می‌شوند.

۲-۳- بهینه‌سازی ازدحام ذرات تسریع‌شده (APSO)

PSO استاندارد از بهترین موقعیت شخصی و سراسری برای به‌روزرسانی سرعت ذره استفاده می‌کند. استفاده از بهترین موقعیت شخصی در واقع برای افزایش تنوع ازدحام است که با استفاده از مقادیر تصادفی هم این هدف تأمین می‌شود؛ بنابراین، برای تسریع هم‌گرایی الگوریتم PSO یک نسخه ساده‌شده از آن به نام APSO معرفی شده است که در آن تنها بهترین موقعیت سراسری مورد توجه قرار می‌گیرد [9]. در این نسخه از فرمول (۷) برای به‌روزرسانی سرعت ذرات و از فرمول (۸) برای جلوگیری از هم‌گرایی زودرس استفاده می‌شود:

$$v_i^{t+1} = v_i^t + \beta(g^* - x_i^t) + \alpha\epsilon_t \qquad (۷)$$

$$\alpha = \alpha_0\gamma^t \qquad (۸)$$

همان‌طور که در فرمول (۷) می‌بینید، بخش نخست نشان‌دهنده موقعیت فعلی ذره iام است. بخش دوم معادل مؤلفه اجتماعی ذره است که آن را به سمت بهترین موقعیت سراسری سوق می‌دهد و بخش سوم معادل یک حرکت تصادفی در فضای جستجوست. در بخش دوم β نشان‌دهنده میزان جذابیت بهترین موقعیت سراسری است و مقداری بین صفر و یک دارد که هرچه به یک نزدیک‌تر باشد، سرعت هم‌گرایی افزایش خواهد یافت. در بخش سوم فرمول، α یک پارامتر تصادفی بین صفر و یک است [9]. برای جلوگیری از هم‌گرایی زودرس، α می‌تواند متغیر باشد و به‌تدریج با استفاده از فرمول سوم مقدار آن کاهش یابد. که در آن γ یک پارامتر کنترلی است [10].

با توجه به این فرمول‌ها در محاسبات هم exploration (با استفاده از مقادیر تصادفی) و هم exploitation (بهترین موقعیت) وجود دارد. ترکیب مناسب این دو به‌طورمعمول به یافتن بهینه سراسری می‌انجامد.

در ابتدا مقادیر اولیه پارامترهای α و β تعیین می‌شوند؛ سپس جمعیت اولیه با N ذره به‌صورت تصادفی تولیدشده، تابع برازش برای هر ذره جمعیت محاسبه و بهترین آن‌ها انتخاب می‌شود. APSO به‌صورت تکرارشونده اجرا شده و در هر تکرار موقعیت ذرات با استفاده از فرمول‌های معرفی‌شده به‌روزرسانی می‌شود. این روند تا رسیدن به شروط خاتمه ادامه می‌یابد و درنهایت بهترین ذره به‌عنوان خروجی معرفی خواهد شد.

۴- کارهای انجام‌شده

نخستین سامانه‌های پرسش و پاسخ مربوط به سال‌های ۱۹۶۱ و ۱۹۷۲ هستند که هر دوی آن‌ها پاسخ پرسش به زبان طبیعی را از پایگاه داده ساخت‌یافته پیدا می‌کردند. نام این سامانه‌ها به‌ترتیب BASEBALL و LUNAR بود. سامانه BASEBALL سؤالاتی را در مورد بازی‌های Baseball یک سال گذشته پاسخ می‌داد. از آن‌جایی که این سامانه‌ها جستجوی خود را در یک پایگاه داده محدود انجام می‌دادند، الگوهایی به‌صورت دستی برای اطلاعات پایگاه داده تعیین می‌شد تا پاسخ پرس‌وجوی ورودی به‌سادگی با استفاده از این الگوها به‌دست آید [11]. گرچه این سامانه‌ها به‌عنوان نخستین گام در عرصه سامانه‌های پرسش و پاسخ نقش مهمی ایفا کردند، اما استفاده از پایگاه داده ساخته‌یافته نقطه ضعف اصلی آن‌ها بود.

تفاسیر زیادی در وب موجود است و این اطلاعات به زبان‌های مختلف هستند؛ بنابراین، از روش‌های آماری نیز برای استخراج پاسخ پرسش‌های ورودی استفاده شده است. برای مثال سامانه‌ای که Echihabi و همکارانش معرفی کردند. در این سامانه دامنه‌محدود، هدف یافتن پاسخ‌ها از میان مجموعه‌ای از اسناد فنی بود و قواعد نحوی متنوعی که خروجی منطقی یکسانی داشتند؛ به‌علاوه جابه‌جایی‌های مختلف کلمات که نتیجه یکسانی تولید می‌کنند، در نظر گرفته می‌شد [12]. این سامانه یکی از نخستین سامانه‌هایی بود که به شکل جدی قواعد پردازش زبان طبیعی (قواعد نحوی متن) را نیز برای استخراج پاسخ در نظر می‌گرفت.

موضوع سامانه‌های پرسش و پاسخ در کتب و مقالات پژوهشی و مروری مختلف بررسی شده است [11-21]. همچنین تلاش‌هایی در راستای مهندسی ویژگی‌ها در [19-21] صورت گرفته است که روش یادگیری خودکاری برای الگوهای پیچیده مانند ارتباطات معنایی موجود در پرسش‌ها و متون معرفی کرده‌اند. آن‌ها روش یادگیری را به‌وسیله درخت‌های حاصل از قواعد نحوی

[DOI: 10.52547/jsdp.19.2.161]

به‌دست می‌آورند. در [21] یک روش جدید نظارتی[1] برای کشف ساختار ارتباطی بین یک پرسش و بخش‌هایی از متن که پاسخ محتمل پرسش بودند، معرفی کرد که هدف آن یادگیری مدل رتبه‌بندی بود. ساختار پیشنهادی آن‌ها شامل SVM[2] و هسته‌هایی در قالب درخت بود که می‌توانست زوج‌های پرسش و پاسخ را در یک فضای بزرگ از ویژگی‌ها نمایش دهد.

مقالات مختلفی در مورد رتبه‌بندی متن وجود دارند [22-24]. Muschitti و همکارانش هسته‌های زبانی برای رتبه‌بندی پاسخ‌ها پیشنهاد کرد که از مدل‌های تشخیص نظارتی استفاده می‌کرد. آن‌ها از چهار ویژگی برای این منظور استفاده می‌کردند. Heie و همکارانش از کلمات کلیدی و ویژگی انواع مختلف پرسش‌ها برای رتبه‌بندی پاسخ‌ها و از یک مدل ریاضی برای استخراج آن‌ها استفاده می‌کردند. استفاده از ویژگی‌های معنایی توسط Moreda و همکارانشان پیشنهاد شد. آن‌ها از دو ویژگی معنایی برای پاسخ به پرسش‌ها با دامنه باز استفاده کردند. از آن‌جایی که این پژوهش‌ها فقط به قواعد نحوی بسنده نکردند و ویژگی‌های معنایی متن را نیز مد نظر قرار دادند، موفق شدند به نتایج مناسبی در آن زمان دست یابند؛ به‌علاوه، دامنه بازبودن این سامانه یک نقطه عطف دیگر در عرصه سامانه‌های پرسش و پاسخ محسوب می‌شود.

در برخی مقالات از ویژگی‌های پرسش برای تشخیص نوع پاسخ (عدد، شخص و تعریف) استفاده کردند. در این روش تعدادی از المان‌های پاسخ مثل موضوع آن توسط کاربر وارد می‌شد. همچنین، نمونه‌ای دیگر از این پژوهش‌ها هر کلمه کلیدی در پرس و جو را به یک عنوان نسبت می‌داد و از یک روش قاعده‌محور برای تشخیص پاسخ استفاده می‌کرد. در این پژوهش‌ها توجه به ساختار و کلمات کلیدی پرسش تأثیر به‌سزایی در سرعت و دقت سامانه‌های پیشنهادی داشت و تحولی مهم در این عرصه به وجود آمده [18, 25].

در ادامه این روند، مقالاتی در مورد تفکیک پرسش‌ها با استفاده از روش‌های یادگیری ماشین ارائه شده‌اند [26-29]. این مقالات از روش‌های نظارتی و نیمه‌نظارتی برای رتبه‌بندی پاسخ‌های محتمل برای پرسش ورودی استفاده کرده‌اند و عملکرد بهتری نسبت به سامانه‌های قبلی دست یافتند.

تا به اینجا، بیش‌تر مقالات هدفشان پاسخ به پرسش‌هایی در مورد حقایق بوده است. پاسخ به پرسش‌های تعریفی در [15, 18, 29] مورد بررسی قرار گرفته است. در [18] از برنامه‌نویسی ژنتیک[3] برای تولید یک فرمول (مبتنی بر ویژگی‌های متن) جهت پاسخ به پرسش‌هایی در مورد حقایق و تعریفی استفاده شد. در [15, 29] از ویژگی‌های زبانی برای پاسخ به پرسش‌های تعریفی استفاده و یک سامانه پرسش و پاسخ به نام MedQA معرفی شد که به‌صورت دامنه‌محدود به پرسش‌های تعریفی در حوزهٔ پزشکی پاسخ می‌داد.

درهمین‌اواخر تمرکز بسیاری از پژوهش‌ها بر بهینه‌سازی سامانه‌های پرسش و پاسخ است. در برخی پژوهش‌ها برای بهینه‌کردن عملکرد این سامانه‌ها از الگوریتم‌های تکاملی به‌خصوص الگوریتم ژنتیک استفاده شده است [13, 14]. این روش‌ها توانسته‌اند تأثیر چشم‌گیری در افزایش دقت این سامانه‌ها داشته باشند [6, 11, 18, 29-31].

یکی از مهم‌ترین پژوهش‌هایی که از الگوریتم‌های تکاملی در این راستا استفاده کرده است، مقاله سال ۲۰۱۹ ارائه‌شده توسط توحیدی و هاشمی‌نژاد است. در این پژوهش یک سامانه پرسش و پاسخ ارائه شد که از الگوریتم تکاملی چندهدفه NSGA-II استفاده می‌کرد. درواقع، این پژوهش از چند ویژگی مختلف برای بررسی نتایج بازیابی‌شده به وسیله موتور جستجو، از ابعاد متفاوت (لغوی، نحوی و معنایی) استفاده کرد و توانست به دقت قابل توجهی در مقایسه با پژوهش‌های دیگر در پاسخ به پرسش‌های ورودی دست یابد [16].

۵- سامانه‌های پرسش و پاسخ

این سامانه‌ها در حالت کلی، گونه‌ای از سامانه‌های بازیابی اطلاعات به‌شمار می‌روند که با دراختیارداشتن مجموعه‌ای از اسناد، می‌کوشند تا برای پرسش‌های مطرح‌شده که اغلب در قالب زبان طبیعی است، پاسخ‌های مناسب استخراج کنند؛ بنابراین، این قبیل سامانه‌ها نسبت به سامانه‌های معمول بازیابی اطلاعات، با روش‌های پیچیده‌تر پردازش زبان طبیعی سر و کار دارند و در محافل علمی به‌عنوان نسل آینده موتورهای جستجوی اطلاعات مطرح هستند [3].

سامانه‌های پرسش و پاسخ، با انواع مختلفی از پرسش‌های کاربران سر و کار دارند، برای مثال اطلاعات در

[1] Supervised
[2] Support Vector Machine
[3] Genetic Programming

[DOI: 10.52547/jsdp.19.2.161]

مورد حقایق موجود، تعاریف مفاهیم، چگونگی و چرایی مسائل مختلف، تئوری‌ها، پرسش‌های منطقی و نیز پرسش‌های بین زبانی. مجموعه‌ی اسناد مورد جستجو نیز می‌تواند در مقیاس‌های مختلفی از جمله مجموعه اسناد یک سازمان تا اطلاعات موجود در محیط وب، متغیر باشد [32].

در زبان انگلیسی پرسش‌های پرسیده‌شده توسط کاربر براساس نوع پاسخ و شیوه استخراج آن می‌توانند به گروه‌های مختلفی تقسیم شوند. مانند: پرسش‌های در مورد حقایق، تعریفی، در مورد ارتباطات، استدلالی و لیست [2]. سه گروه مهم را در ادامه آورده‌ایم:

- پرسش در مورد حقایق (Factoid): پرسش‌هایی هستند که پاسخ‌شان یک واقعیت کوتاه است و اغلب تعداد کلمات محدودی به‌عنوان پاسخ دارند. مانند:

"who is the headmaster of Hogwarts?"

- پرسش تعریفی (Definitional): پرسش‌هایی هستند که یک تعریف را از سیستم می‌خواهند. جواب‌های این‌گونه پرسش‌ها به‌طورمعمول طولانی‌تر از پاسخ پرسش‌های factoid است. مانند:

"What is DNA?"

- پرسش فهرست: این پرسش‌ها فهرستی از پاسخ‌ها دارند. هر عضو لیست شبیه یک پاسخ پرسش factoid است. پاسخ این پرسش‌ها، می‌تواند ترکیبی از بخش‌های مختلف یک سند باشد. مثل:

"In which countries French is an official language?"

تاکنون اغلب پژوهش‌ها در مورد سامانه‌های پرسش و پاسخ بر پرسش‌های factoid تمرکز داشته‌اند.

بعضی پژوهش‌ها به صورت‌های دیگر پرسش‌ها و پاسخ‌ها را طبقه‌بندی کرده‌اند. برای مثال گروه‌بندی پرسش‌ها با توجه به ماهیت پاسخ آن‌ها [15]. نمونه‌ای از این طبقه‌بندی شامل موارد زیر است [2]:

- ✓ Human
- ✓ Location
- ✓ Description
- ✓ Numeric
- ✓ Abbeviation
- ✓ Entity

در این روش‌ها نیاز به تسلط بر قواعد زبان و آشنایی با روش‌های ساخت پرسش و پاسخ‌های محتمل برای آن‌هاست. برای مثال پاسخ پرسشی با فرمول زیر همواره در گروه Human قرار می‌گیرد [2]:

Who {is| was| are| were}

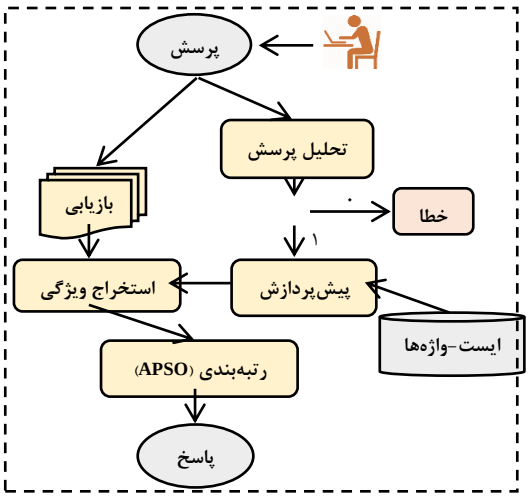
۶- سامانه پرسش و پاسخ پیشنهادی

مسأله پاسخ به پرسش‌های ورودی درواقع معادل بازیابی اطلاعات با توجه به پرسش ورودی است. همان‌طور که اشاره شد برای بازیابی اطلاعات نیاز است تا نامزدها را اولویت‌دهی کنیم. این نامزدها بخش‌هایی از جملات سند بازیابی‌شده به‌وسیلهٔ موتور جستجوی استاندارد هستند.

هدف این است تا با توجه به پرسش، مناسب‌ترین بخش از این سند به‌عنوان پاسخ پرسش انتخاب شود.

مسأله اولویت‌دهی شامل محاسبه امتیاز مبتنی بر مجموعه‌ای از ویژگی‌هاست. در این پژوهش، با توجه به گسترده‌بودن وب، از الگوریتم APSO که نسخه ساده‌شده PSO است و ویژگی‌های متن، برای این منظور استفاده می‌کنیم.

شکل (۱) شمای سامانه پرسش و پاسخ پیشنهادی را نشان می‌دهد.



(شکل-۱): شمای سامانه پرسش و پاسخ پیشنهادی

(Figure-1): The structure of the proposed question answering system

در ادامه به توضیح بخش‌های مختلف شکل (۱) می‌پردازیم.

۶-۱- تحلیل پرسش

همان‌طور که اشاره شد انواع مختلفی از پرسش‌ها وجود دارند. هدف این فاز تعیین نوع پرسش و تشخیص امکان پاسخ‌دهی به آن به‌وسیلهٔ مدل ارائه شده است؛ زیرا این مدل توان پاسخ‌دهی به همه انواع پرسش‌ها را ندارد؛ بنابراین، ورودی این فاز پرسش ورودی است و بررسی می‌کند که پرسش واردشده قابل پاسخ‌دهی به‌وسیلهٔ سامانه پیشنهادی هست یا خیر. در این سامانه پاسخ به پرسش‌های تعریفی و پرسش‌های در مورد حقایق را بررسی می‌کنیم که بیش‌تر پرسش‌های ممکن را تشکیل می‌دهند و رایج‌ترین انواع پرسش‌ها هستند.

[DOI: 10.52547/jsdp.19.2.161]

۲-۶- پیش‌پردازش

در این فاز پرسش به‌عنوان ورودی دریافت می‌شود و کلمات کم اهمیت‌تر و یا ایست‌واژه‌ها[1] از متن مورد پردازش (پرسش ورودی)، حذف می‌گردند. ایست‌واژه‌ها لغاتی پرکاربرد و اغلب کم‌اهمیت هستند که در متن به‌وفور ظاهر می‌شوند. مثل "or"، "if".

در نگاه نخست کلمات ربط و تعریف، ایست‌واژه به‌نظر می‌آیند؛ درحالی‌که برخی افعال کمکی، قیدها و صفات نیز به‌عنوان ایست‌واژه شناخته شده‌اند. این کلمات برخلاف این‌که بسیار استفاده می‌شوند، اما از *لحاظ معنایی دارای اهمیت کمی بوده* و به همین دلیل به‌طورعمومی در فعالیت‌هایی که با حجم انبوهی از داده‌ها روبه‌رو هستند، در فاز پیش‌پردازش حذف می‌شوند. برای حذف این کلمات به‌طورعمومی فهرستی از این کلمات از پیش تهیه می‌شود و سپس در صورت رخداد این کلمات در متن، از سند حذف می‌شوند.

۳-۶- بازیابی

گرچه موتور جستجو مجموعه اسناد را براساس میزان مرتبط‌بودن با پرسش ورودی رتبه‌بندی می‌کند، اما اسناد با اولویت بالاتر در این رتبه‌بندی لزوماً حاوی پاسخ نیستند؛ زیرا این رتبه‌ها متناسب با اهداف سامانه پرسش و پاسخ به دست نیامده‌اند؛ به‌علاوه حتی اگر مرتبط‌ترین سند انتخاب شود، یک متن طولانی که مکان پاسخ در آن مشخص نیست، نتیجه ایده‌آلی نخواهد بود؛ بنابراین، این مرحله شامل استخراج مجموعه‌ای از پاسخ‌های محتمل از متون بازیابی شده است.

برای پرسش و پاسخ از وب، به جای پردازش تمام متن همه اسناد بازیابی‌شده، می‌توان از Snippetهای (تکه‌های متن) تولید شده توسط موتورهای جستجوی استاندارد استفاده کرد و پردازش را روی این متون که خلاصه هستند، انجام داد. درنتیجه سرعت سامانه بیشتر خواهد شد.

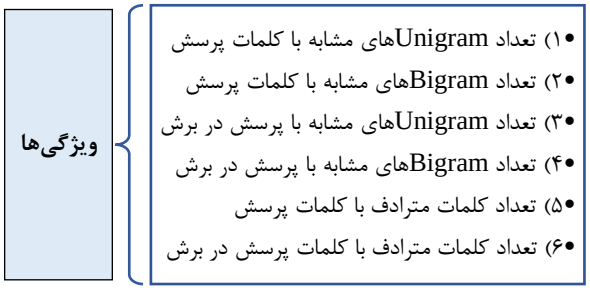
در این مرحله پرسش ورودی به موتور جستجوی استاندارد (گوگل) داده می‌شود و منابع برتر بازیابی شده به‌وسیلهٔ موتور جستجو به‌عنوان خروجی این بخش تولید خواهند شد.

۴-۶- استخراج ویژگی‌ها

کارایی یک سامانه پرسش و پاسخ مبتنی بر وب وابسته به شیوه رتبه‌بندی و تعداد اسناد مرتبط بازیابی‌شده به‌وسیلهٔ موتور جستجو است. اگر سامانه پرسش و پاسخ توانایی پاسخ‌گویی نداشته باشد و یا پاسخ اشتباه تولید کند، دیگر کارایی مناسب نخواهد داشت؛ بنابراین، ویژگی‌هایی برای متون و تشخیص مرتبط‌بودن آن‌ها با پرسش تعریف می‌شود تا سامانه را به سمت بخش‌هایی از متن که حاوی پاسخ محتمل برای پرسش هستند، سوق دهد.

انواع مختلفی از ویژگی‌ها تا به حال برای متن در سامانه‌های پرسش و پاسخ مختلف تعریف شده‌اند. مسأله مهم، انتخاب ویژگی‌های درست با توجه به دامنهٔ سامانه و نوع پرسش‌های محتمل برای آن است.

در شکل (۲) ویژگی‌های مورد استفاده در این مقاله معرفی شده‌اند. منظور از Unigram، تک‌کلمه و منظور از Bigram، دو کلمه متوالی است که در متن ظاهر شده‌اند. منظور از برش، بخشی از یک جمله انتخاب شده است که می‌تواند شامل یک یا بیشتر لغت باشد.



(شکل-۲): ویژگی‌ها
(Figur-2): Features

APSO -۶-۵

این الگوریتم نسخه ساده‌شده و تسریع‌شده از الگوریتم پرکاربرد PSO است، که در آن ابتدا مقادیر اولیه برای پارامترها تنظیم می‌شود. جمعیت اولیه به‌صورت تصادفی تولید و سپس مراحل APSO طی می‌شود و فرآیند تا رسیدن به نتیجه مطلوب ادامه می‌یابد.

در قدم نخست به یک نمایش مناسب برای راه‌حل‌های (ذرات) مسأله نیاز داریم. در اینجا یک ذره معادل یک پاسخ برای پرسش ورودی است که در فضای جستجوی سه‌بعدی پاسخ‌ها حرکت می‌کند. پس در جمعیت اولیه که شامل N ذره است، هر ذره یک بردار سه بعدی است که هر مؤلفه آن به‌صورت تصادفی از مجموعه انتخاب‌شده است. مؤلفه نخست نشان‌دهنده شماره جمله و مؤلفه‌های دوم و سوم به‌ترتیب نشان‌دهنده نقاط برش ابتدا و انتهای پاسخ مربوطه در جمله انتخاب شده هستند.

هدف تابع برازش، ارزیابی کیفیت پاسخ است. در اینجا تابع برازش معادل رابطه (۹)، شامل مجموع وزن‌دار ویژگی‌های استخراج‌شده از متن است. در این رابطه حرف f مخفف کلمه feature به معنای ویژگی است و عدد

[1] Stopwords

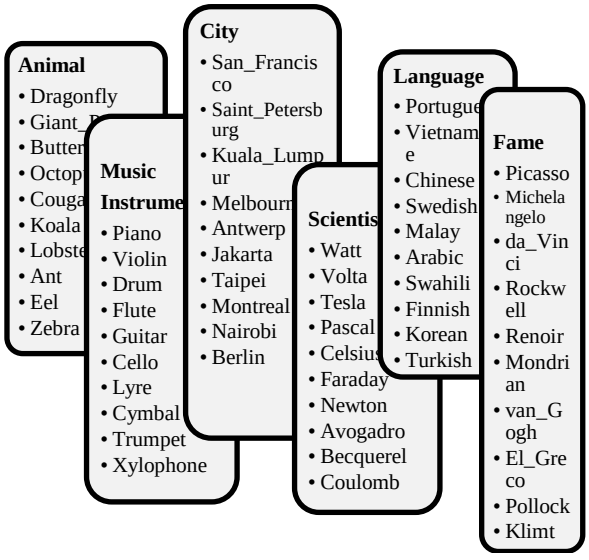
[DOI: 10.52547/jsdp.19.2.161]

مقابل آن شماره ویژگی مربوطه را در شکل (۲) نشان می‌دهد. در این مسأله هدف ما بیشینه‌کردن تابع برازش است. سه مؤلفه ذره به ما کمک می‌کنند تا بخشی از متن را که مربوط به آن ذره است، پیدا کنیم.

$$w_1 f(1) + w_2 f(2) + w_3 f(3) + w_4 f(4) + w_5 f(5) + w_6 f(6) \quad (۹)$$

۷- مجموعه‌داده

برای ارزیابی روش پیشنهادی از مجموعه‌داده دانشگاه Pittsburg استفاده شده است، زیرا متناسب با کاربرد این پژوهش پاسخ‌ها را در میان متونی شامل پاراگراف‌ها مشخص کرده است. به‌علاوه، متون آن تنوع مناسبی دارند و در شش گروه اصلی تقسیم‌بندی شده‌اند. این مجموعه‌داده یکی از مجموعه‌داده‌های استاندارد در این حوزه است که شمای کلی آن در شکل (۳) مشاهده می‌شود [17]:



(شکل-۳): شمای کلی مجموعه داده دانشگاه Pittsburgh
(Figure-3): The big picture of Pittsburgh university's dataset

۸- شاخص ارزیابی

در پژوهش‌های مشابه برای ارزیابی سامانه پرسش و پاسخ از شاخص‌هایی استفاده می‌شود که در ادامه معرفی شده‌اند.

✓ MRR[1]: این معیار در بخش QA در TREC8 استفاده شده است. این معیار از مجموعه آزمایشی برچسب‌خوردهتوسط انسان با پاسخ‌های صحیح، استفاده می‌کند. فرض این معیار، این است که خروجی سامانه جواب‌های کوتاه رتبه‌بندی‌شده است. بدین معنی که هر پرسش با معکوس رتبه نخستین پاسخ صحیح امتیازدهی می‌شود [11].

$$MRR = \frac{1}{Q_n} \times \sum_{\forall q \in Q_u} \frac{1}{rank\_answer_q} \quad (۱۰)$$

در این فرمول Q_n تعداد پرسش‌ها و Q_u مجموعه پرسش‌هاست.

✓ Top-n accuracy: این معیار رایج برای اندازه‌گیری دقت روش به کار می‌رود. درصورتی‌که n را برابر یک در نظر بگیریم به این معنی است که تنها پاسخ تولیدشده توسط روش پیشنهادی باید پاسخ دقیق و درست باشد. اما اگر مقدار n را بزرگ‌تر از یک در نظر بگیریم امکان حدس‌های نادرست را هم به سامانه می‌دهیم. فرمول محاسبه این شاخص صورت زیر است [18]:

$$Top1\ accuracy = \frac{number\ of\ correct\ answers}{number\ of\ questions} \quad (۱۱)$$

۹- آزمون روش پیشنهادی

در این بخش به بررسی نتایج آزمون روش پیشنهادی خواهیم پرداخت؛ سپس این روش با روش‌های مرتبط مقایسه می‌شوند. به‌منظور ارزیابی روش مطرح‌شده، روش پیشنهادی بر روی مجموعه‌دادهٔ دانشگاه Pittsburgh که پیش‌تر توضیح داده شد، مورد آزمایش قرار گرفته است.

کارایی این روش همان‌طور که در بخش شاخص ارزیابی نیز به آن اشاره شد بر طبق معیارهای Accuracy و MRR مورد ارزیابی قرار خواهد گرفت. از آن‌جایی که سرعت سامانه‌های پرسش و پاسخ هم حائز اهمیت است، زمان اجرای الگوریتم نیز مورد بررسی قرار گرفته است.

روش آزمون

در این پژوهش برای تمامی پیاده‌سازی‌ها از زبان‌های برنامه‌نویسی Java و Matlab استفاده شده است؛ به‌علاوه، این پیاده‌سازی‌ها بر روی سیستمی با مشخصات زیر تحت سیستم عامل ویندوز صورت گرفته‌اند:

Intel(R) Core(TM) i7-3520M CPU @ 2.90GHz, 2901 Mhz, 2 Cores, 4 Logical Processors, 8 GB RAM

قبل از ارزیابی و مقایسه روش پیشنهادی با روش‌های مشابه، لازم است تا مقادیر پارامترهای APSO تنظیم شوند. پارامترهای α_0، γ، N و T به‌ترتیب معادل پارامتر تصادفی اولیه، پارامتر کنترلی، اندازه جمعیت و بیشینه تعداد تکرارها هستند، که در آن دو پارامتر نخست مقداری متغیر و دو پارامتر دوم مقادیر ثابتی دارند. مقادیر نهایی محاسبه‌شده برای این پارامترها در پیاده‌سازی روش پیشنهادی در جدول (۱) قابل مشاهده است.

[1] Mean Reciprocal Rank

[DOI: 10.52547/jsdp.19.2.161]

(جدول-۱): مقدار پارامترهای APSO
(Table-1): Parameter values of APSO

α_0	۰/۱
γ	۰/۹۱
N	۳۰
T	۲۰

با توجه به این‌که نتایج حاصل از اجرای الگوریتم‌های تکاملی بر روی یک مجموعه‌داده در هر بار اجرا می‌تواند متفاوت باشد، در پژوهش‌ها، الگوریتم انتخابی به‌صورت مکرر اجرا می‌شود تا اطمینان حاصل شود که نتایج به‌دست‌آمده تصادفی نبوده‌اند. به‌طورمعمول سی بار تکرار برای گزارش نتیجه مناسب است [33]؛ بنابراین، در اینجا اجرای APSO به‌ازای هر پرسش، سی بار تکرار شده و نتیجه نهایی میانگین این نتایج در نظر گرفته شده است.

برای یافتن ضرایب ویژگی‌ها در فرمول تابع هدف، مقادیر مختلفی آزمایش شده‌اند که نتایج آن در جدول (۲) مشاهده می‌شود.

(جدول-۲): ضرایب مختلف برای ویژگی‌ها
(Table-2): Weights of each feature for the fitness function.

$W = \{w_1, w_2, \ldots, w_6\}$	MRR	Top1 Accuracy
{۱, ۱, ۱, ۱, ۱, ۱}	۰/۵۲	۰/۲۹۱
{۲, ۱, ۱, ۱, ۱, ۱}	۰/۵۸۶	۰/۲۵۲
{۱, ۲, ۱, ۱, ۱, ۱}	۰/۵۹۷	۰/۴۶۵
{۱, ۱, ۲, ۱, ۱, ۱}	۰/۵۳۴	۰/۲۸۶
{۱, ۱, ۱, ۲, ۱, ۱}	۰/۶۴۳	۰/۴۳۵
{۱, ۱, ۱, ۱, ۲, ۱}	۰/۵۷۷	۰/۳۸۲
{۱, ۱, ۱, ۱, ۱, ۲}	۰/۵۶۹	۰/۳۸۹
{۱, ۲, ۱, ۲, ۱, ۱}	**۰/۷۱۱**	**۰/۵۲۷**
{۱, ۱/۵, ۱, ۱/۵, ۱, ۱}	۰/۶۷	۰/۵۰۳
{۱, ۱/۷۵, ۱, ۱/۷۵, ۱, ۱}	۰/۶۸۴	۰/۵۱۵
{۱, ۲/۲۵, ۱, ۲/۲۵, ۱, ۱}	۰/۶۸۴	۰/۵۱۹

در نتیجه، تابع هدف از رابطه (۱۲) به‌دست می‌آید:

$$[f(1) + f(3)] + 2 * [f(2) + f(4)] + [f(5) + f(6)] \quad (۱۲)$$

به‌عنوان مثال برای این پرسش:

"Who shares credit with Isaac Newton for developing calculus?"

و متن زیر:

"… Newton also formulated an empirical law of cooling and studied the speed of sound.
In mathematics, Newton shares the credit with Gottfried Leibniz for the development of the differential and integral calculus. He also demonstrated the generalised binomial theorem, developed the so-called "Newton's method" for approximating the zeroes of a function, and contributed to the study of power series. ..."

الگوریتم APSO پاسخ زیر (موقعیت ذره در فضای سه‌بعدی) را به‌عنوان پاسخ این پرسش از متن بالا انتخاب می‌کند:

(۳،۲،۱۸)

گفتنی است، پاسخ بالا پرتکرارترین پاسخ تولیدشده در میان تکرارهای مختلف برای این مثال است.

حال بررسی می‌کنیم که این پاسخ، نشان‌دهنده کدام قسمت از متن است و توابع هدف برای این پاسخ چه مقداری دارند. به این منظور مقادیر ویژگی‌های مربوط به این مثال را محاسبه خواهیم کرد. نخستین عدد (سمت چپ) نشان‌دهنده شماره جمله و دو عدد دوم و سوم نشان‌دهنده برش انتخاب‌شده از این جمله هستند؛ بنابراین، جمله سوم از متن و کلمه دوم تا هجدهم آن انتخاب شده‌اند:

"mathematics, Newton shares the credit with Gottfried Leibniz for the development of the differential and integral calculus."

در تابع هدف، ویژگی نخست تعداد کلمات مشابه پرسش و جمله و ویژگی دوم همین مفهوم را بین برش انتخاب شده از جمله و پرسش محاسبه می‌کند. ویژگی سوم نشان‌دهنده تعداد عبارات دو کلمه‌ای متوالی مشابه بین پرسش و جمله است و ویژگی چهارم همین مفهوم را بین برش انتخاب‌شده از جمله و پرسش محاسبه می‌کند. ویژگی پنجم نشان‌دهنده تعداد کلمات هم‌ریشه بین پرسش و جمله است و ویژگی ششم همین مقدار را بین برش انتخاب‌شده از جمله و پرسش محاسبه می‌کند. کلمه‌های هم‌ریشه در اینجا developing و development هستند؛ بنابراین تابع هدف به صورت زیر محاسبه می‌شود:

(۴+۴) + ۲* (۰+۰) + (۱+۱) = ۱۰

در مجموعه‌داده، برش شامل لغات سوم تا هجدهم، دارای رتبه یک است و APSO در این مثال بهترین پاسخ را انتخاب کرد. APSO در کاربردهای مختلف با تعداد تکرارهای متفاوتی به نتیجه نهایی هم‌گرا می‌شود. برای یافتن تعداد تکرارهای مناسب، این الگوریتم را با تعداد تکرارهای مختلف اجرا کردیم. دقت و MRR اندازه‌گیری‌شده با هر دو شاخص ارزیابی بعد از حدود بیست تکرار به نتیجه نهایی هم‌گرا می‌شود؛ بنابراین، پارامتر تعداد تکرارهای APSO برابر بیست در نظر گرفته شده است.

حال به مقایسه روش پیشنهادی با نتایج برخی پژوهش‌های مرتبط انجام‌شده می‌پردازیم. از آنجایی که هر

[DOI: 10.52547/jsdp.19.2.161]

یک از کارهای پیشین از معیارهای ارزیابی مختلفی برای ارائه نتایج خود استفاده کرده‌اند [32]، در ادامه روش پیشنهادی را با پژوهش‌هایی که از دو معیار ارزیابی استاندارد معرفی شده استفاده کرده‌اند (از جمله پژوهش‌های [11]، [16] و [18]) مقایسه خواهیم کرد. این پژوهش‌ها در سامانه‌های پیشنهادی خود توانسته‌اند به نتایج مناسبی در مقایسه با سایر پژوهش‌ها دست یابند. پژوهش نخست از شاخص ارزیابی MRR، دیگری از شاخص ارزیابی Top1-Accuracy و پژوهش سوم از هر دو شاخص برای تعیین کیفیت کار خود استفاده کرده است.

در جدول (۳) دقت پژوهش‌های انجام‌شده در حوزه سامانه‌های پرسش و پاسخ با دقت روش پیشنهادی مقایسه شده است.

(جدول-۳): مقایسه دقت روش پیشنهادی با پژوهش‌های مرتبط

(Table-3): Comparing accuracy of the proposed method with others

سامانه	Top1 Accuracy
GP-based feature learning QAS [18]	۰/۴۶۰
sv2007c [18]	۰/۲۸۹
FMIX [18]	۰/۲۹۰
LCCFerret [18]	۰/۴۹۴
MOQAS [16]	۰/۵۲۷
مدل پیشنهادی	**۰/۵۲۷**

همان‌طور که از نتایج موجود در جدول مشخص است این روش با درنظرگرفتن ویژگی‌های مناسب و با به‌کارگیری APSO توانسته است، دقت سامانه (محاسبه‌شده با استفاده از شاخص Top1 Accuracy) را در مقایسه با پژوهش‌های مشابه و سامانه‌های پرسش و پاسخ موجود افزایش دهد و یا به دقتی برابر با آن‌ها دست پیدا کند. در اینجا شاید این پرسش پیش بیاید که دلیل کوچک‌بودن اعداد جدول (۳) چیست. برای پاسخ به این پرسش، توجه به یک نکته ضروری است که در اینجا معیار Top1 Accuracy است؛ یعنی در صورتی عملکرد مورد تأیید است که تنها پاسخ تولیدشده به‌وسیلهٔ سامانه، پاسخ صحیح باشد و امکان ارائه یک فهرست از پاسخ‌های محتمل در خروجی سامانه وجود ندارد.

در ادامه، در جدول (۴) به مقایسه روش پیشنهادی با پژوهش‌های مرتبط به‌وسیله شاخص MRR پرداخته‌ایم.

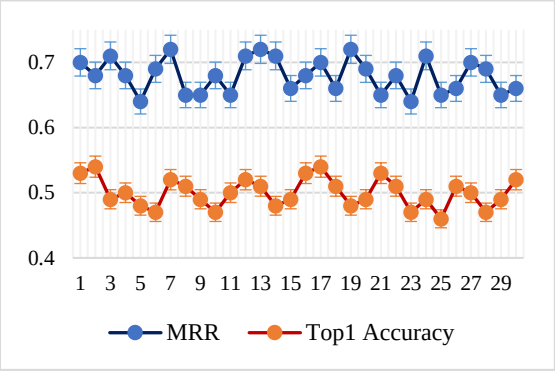
(جدول-۴): مقایسه MRR روش پیشنهادی با پژوهش‌های مرتبط

(Table-4): Comparing MRR values of the proposed method with others

سامانه	MRR
GA for Data-Driven web QAS (PreGA) [11]	۰/۳۸۷
GA for Data-Driven web QAS (GAQA+GASCA) [11]	۰/۵۶۹
FMIX (Perceptron) [18]	۰/۶۴۱
FMIX (SVM-rank) [18]	۰/۶۳۸
MOQAS [16]	۰/۷۱۱
مدل پیشنهادی	**۰/۷۱۱**

با توجه به نتایج جدول بالا، روش پیشنهادی توانسته معیار MRR را نسبت به بیش‌تر پژوهش‌های مشابه تا حد قابل قبولی افزایش دهد.

همان‌طور که در شکل (۴) مشاهده می‌شود، واریانس معیارهای MRR و Top1 Accuracy به‌ترتیب برابر ۰/۰۰۴ و ۰/۰۰۶ هستند.



(شکل-۴): واریانس معیارهای ارزیابی معرفی‌شده در هر تکرار

(Figure-4): The variance of the evaluation metrics in each iteration

کوچک‌بودن مقدار واریانس نشان می‌دهد که معیارهای ارزیابی استفاده‌شده در حوزه سامانه‌های پرسش و پاسخ عملکرد مناسبی داشته‌اند.

تا اینجا به نظر می‌رسد الگوریتم پیشنهادی توانسته راهکاری برای افزایش دقت و بهبود عملکرد سامانه‌های پرسش و پاسخ موجود ارائه دهد و نتایج برابر ویا بهینه‌تری نسبت به آن‌ها تولید کند. حال از آن‌جایی که زمان پاسخ‌گویی به پرسش، مسأله مهمی است، به بررسی سرعت سامانه پیشنهادشده نسبت به سایر پژوهش‌ها در جدول (۵) می‌پردازیم.

[DOI: 10.52547/jsdp.19.2.161]

(جدول-۵): مقایسه سرعت روش پیشنهادی با پژوهش‌های مرتبط

(Table-5): Comparing the speed of the proposed method with others

سامانه	میانگین زمان (میلی ثانیه)
GA for Data-Driven web QAS (PreGA) [11]	۳۶۶۷۱/۰۰۶
GA for Data-Driven web QAS (GAQA+GASCA) [11]	۱۸۵۷۴/۶۳۴
MOQAS [16]	۵۶۱۲۰/۷۶۹
مدل پیشنهادی	**۱۱۹۳/۲۰۰**

اگرچه درمجموع سرعت الگوریتم‌های تکاملی در مقایسه با سایر روش‌ها محدودیت محسوب می‌شود، اما همان‌طورکه در جدول (۵) مشاهده می‌شود، از آنجایی که در این پژوهش از APSO استفاده شده است، زمان پاسخ‌گویی نسبت به همه سامانه‌هایی که تاکنون از الگوریتم تکاملی استفاده کرده‌اند، بهبود یافته است و نسبت به کار پژوهشی قبلی نویسندگان این مقاله ([15]) توانسته بیش از پنج برابر سرعت را افزایش دهد.

۱۰- جمع‌بندی

در این مقاله از APSO برای بهبود کارایی سامانه‌های پرسش و پاسخ استفاده کردیم. دلیل استفاده از الگوریتم‌های تکاملی مناسب‌بودن آن‌ها برای مسائل با فضای جستجوی بزرگ و ذات آن‌ها برای یافتن پاسخ بهینه مسأله است. همچنین، دلیل استفاده از APSO، سرعت بالا و پیاده‌سازی ساده آن است که در ارزیابی‌های انجام‌شده، نشان داده شد که سرعت آن می‌تواند تا پنج برابر سریع‌تر از سامانه‌های پرسش و پاسخ کنونی باشد.

در بخش ارزیابی دیدیم که نسبت به پژوهش‌های موجود، روش پیشنهادی با حفظ دقت و کیفیت، توانسته است معیار حیاتی سرعت را نسبت به سایر روش‌های مشابه بهبود بخشد؛ به‌علاوه، با پیشنهاد کار بر روی Snippetهای بازیابی‌شده به‌وسیلهٔ موتور جستجوی استاندارد که طول کم‌تری نسبت به کل سند دارند، سرعت بهبود بیشتری خواهد یافت.

کارهای آینده در این زمینه می‌تواند پاسخ به سایر انواع پرسش‌ها باشد. در این راستا برای هر یک از انواع پرسش‌ها باید ویژگی‌های جدیدی معرفی و محاسبه شوند و سپس مراحل انجام‌شده در این پژوهش روی آن‌ها اعمال شود؛ به‌علاوه، پویاساختن این راه‌حل به‌طوری که بتواند خود را با توجه به بازخورد کاربر به‌روز کند، می‌تواند گام مثبتی در جهت بهبود عملکرد سامانه پرسش و پاسخ باشد.

همچنین در آینده می‌توان با اضافه‌کردن بخشی که پرسش‌های متداول و یا متون پرتکرار حاوی پاسخ را در خود ذخیره می‌کند، بار پردازشی سامانه را کاهش داد و به افزایش سرعت آن کمک کرد.

11- References ۱۱- مراجع

[1] F. Sebastiani, "Machine Learning in in Automated Text Categorization," *ACM Computing Surveys (CSUR)*, vol. 34, pp. 1-47, 2002.

[2] H. Sundblad, "Question Classification in Question Answering Systems," 2007.

[۳] بذرافشان، مرجان؛ حیدری، سمانه؛ درودی، احسان؛ رهنما، علی؛ سارابی، زهرا؛ سعیدی، پریسا؛ شرکت، احسان؛ فرهودی، مژگان؛ لآلی، مجید؛ مهیار، هومن؛ هاشمی نژاد، محمدحسین, "ایجاد و راهاندازی سامانهی پرسش و پاسخ خودکار در حوزهی قرآنی," پژوهشکده فناوری اطلاعات, ۱۳۹۲.

[3] M. Bazrafshan, S. Heidari, E. Dorodi, A. Rahnama, Z. Sarabi, E. Sherkat, M.Fahordi, M. Lali, M. Homan, SMH. Hashemienjad, "Question Answering for Quran," ITRC, 2012.

[4] Tohidi, Nasim; Rustamov, Rustam B., "Short Overview of Advanced Metaheuristic Methods," *International Journal on Technical and Physical Problems of Engineering (IJTPE)*, vol. 14, pp. 84-97, 2022.

[5] D. Graff, "The AQUAINT Corpus of English News Text LDC2002T31," *Web Download. Philadelphia: Linguistic Data Consortium*, 2002.

[6] I.Khodadi, M. Saniee Abadeh, "A Memetic-Based Approach for Web-Based Question Answering," *Information Technology and Computer Science*, vol. 9, pp. 39-45, 2014.

[7] J. Kennedy and R. C. Eberhart, "Particle swarm optimization," in in Proceedings of the IEEE International Conference on Neural Networks, 1995.

[8] A. L. Ballardini, "A tutorial on Particle Swarm Optimization Clustering," *ArXiv*, vol. abs/1809.01942, 2018.

[9] Y. XS, Nature-Inspired Metaheuristic Algorithms., Luniver Press, 2008.

[DOI: 10.52547/jsdp.19.2.161]

relationships for large-scale learning of answer re-ranking," in *ACM*, In Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval, 2012.

[22] A. Moschitti, S. Quarteroni, "Linguistic kernels for answer re-ranking in question answering systems", *Elsevier, Information Processing and Management*, vol. 47, p. 825–842, 2011.

[23] M. H. Heie, E. W. D.Whittaker, S. Furui, "Question answering using statistical language modelling," *Computer Speech and Language*, vol. 26, no. 3, pp. 193-209, 2012.

[24] P. Moreda, H. Llorens, E. Saquete, M. Palomar, "Combining semantic information in question answering systems," *Information Processing & Management*, vol. 47, no. 6, pp. 870-885, 2011.

[25] S. Yoon, A. Jatowt, K. Tanaka, "Detecting Intent of Web Queries Using Questions and Answers in CQA Corpus," in *2011* IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology, Lyon, 2011.

[26] S. Kandasamy, A.Cherukuri, "Information retrieval for Question Answering System using Knowledge based Query Reconstruction by adapted LESK and LATENT Semantic analysis," *International Journal of Computer Science & Applications*, vol. 14, no. 2, 2017.

[27] D. Croce, A. Moschitti, R. Basili, "Structured lexical similarity via convolution kernels on dependency trees," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 2011.

[28] H.Toba, Z. Ming, Y. Adriani, M. Chua, "Discovering high quality answers in community question answering archives using a hierarchy of classifiers," *Elzevier, Information Sciences*, vol. 261, pp. 101-115, 2014.

[29] Z. Yu, L. Su, L. Zhao, Q. Mao, C. Guo, "Question classification based on co-training style semi-supervised learning," *Elzevier, Pattern Recognition Letters*, vol. 31, no. 13, pp. 1975-1980, 2010.

[30] A. Ulysse Côté, Kh. Richard; Lamontagne, Luc; Bergeron, Jonathan; Laviolette, François; Bergeron-Guyard, Alexandre;, "Optimizing Question-Answering Systems Using Genetic Algorithms," in *Proceedings of the Twenty-Eighth International Florida Artificial Intelligence Research Society Conference*, 2015.

[10] N. Tohidi, Ch. Dadkhah, "Improving the performance of video Collaborative Filtering Recommender Systems using Optimization Algorithm," *International Journal of Nonlinear Analysis and Applications*, vol. 11, no. 1, pp. 283-295, 2020.

[11] A. Figueroa, G. Neumann, "Genetic Algorithms for Data-DrivenWeb Question Answering, Massachusetts Institute of Technology," *Evolutionary Computation*, 2008.

[12] A. Mishra and S. Kumar Jain, "A survey on question answering systems with classification," *Elsevier, Computer and Information Sciences*, vol. 28, pp. 345-361, 2015.

[13] N. Tohidi, Ch. Dadkhah, B. Rustamov, "Optimizing the performance of Persian multi-objective question answering system," in 16th International Conference on Technical and Physical Problems of Electrical Engineering, Istanbul, 2020.

[14] N. Tohidi, Ch. Dadkhah, B. Rustamov, "Optimizing Persian Multi-objective Question Answering System," *International Journal on Technical and Physical Problems of Engineering (IJTPE)*, vol. 13, pp. 62-69, 2021.

[15] H.Yu, D. Kaufman, "A cognitive evaluation of four online Search Engines for Answering Definitional Questioned posed by Physicians," in *Pacific Symposium on Biocomputing*, 2007.

[16] N. Tohidi, S.M.H. Hasheminejad, "MOQAS: Multi-objective question answering system," *Journal of Intelligent & Fuzzy Systems*, vol. 36, no. 4, pp. 3495-3512, 2019.

[17] N. Smith, A. Heilman, M. Hwa, "Question Generation as a Competitive Undergraduate Course Project," in In Proceedings of the NSF Workshop on the Question Generation Shared Task and Evaluation Challenge, Arlington, VA, 2008.

[18] I. Khodadi, M. Saniee Abadeh, "Genetic programming-based feature learning for question answering," *Elsevier, Information Processing and Management*, vol. 40, 2015.

[19] A.Severyn, A.Moschitti, "Automatic Feature Engineering for Answer Selection and Extraction," in *EMNLP Conference*, 2013.

[20] A.Severyn, M.Nicosia, A. Moschitti, "Learning adaptable patterns for passage reranking," in *CoNLL Conference*, 2013.

[21] A.Severyn, A.Moschitti, "Structural

[DOI: 10.52547/jsdp.19.2.161]

[31] K. Karpagam, A. Saradha, "A Hybrid Optimization Technique for Effective Document Clustering in Question Answering System." *ICTACT Journal on Soft Computing*, vol. 7, no. 3, pp. 1447-1451, 2017.

[32] Ojokoh, Bolanle; Adebisi, Emmanue, "A Review of Question Answering Systems," *Journal of Web Engineering*, vol. 17-8, pp. 717–758, 2019.

[33] S.M.H. Hasheminejad, "An Evolutionary Approach to Identify Logical Components," *Journal of Systems and Software*, vol. 96, pp. 24-50, 2014.

نسیم توحیدی درجه کارشناسی ارشد خود را در رشته مهندسی کامپیوتر در سال ۱۳۹۶ از دانشگاه الزهرا(س) دریافت کرد و هم‌اکنون دانشجوی سال آخر دکترای رشته مهندسی کامپیوتر گرایش هوش‌مصنوعی در دانشگاه صنعتی خواجه نصیرالدین طوسی است.

زمینه‌های پژوهشی ایشان عبارت از پردازش زبان طبیعی، الگوریتم‌های تکاملی است.

نشانی رایانامه ایشان عبارت است از:

n.tohidi@student.alzahra.ac.ir

سید محمدحسین هاشمی‌نژاد درجه دکترای خود را در رشته مهندسی کامپیوتر گرایش نرم‌افزار در سال ۱۳۹۳ از دانشگاه تربیت مدرس دریافت کرده و هم‌اکنون عضو هیأت علمی گروه مهندسی کامپیوتر در دانشگاه الزهرا(س) است.

زمینه‌های پژوهشی ایشان، الگوریتم‌های تکاملی، یادگیری ماشین و مهندسی نرم‌افزار است.

نشانی رایانامه ایشان عبارت است از:

smh.hasheminejad@alzahra.ac.ir

[DOI: 10.52547/jsdp.19.2.161]