Neural Network Modelling of the Information Behavior of Database Users Based on their Previous Interactions with the Search Results

Fataneh Vahabi*
PhD in Library and information science; Iranian Research Institute for Information Science and Technology (IranDoc); Tehran, Iran;
Email: vahabi@students.irandoc.ac.ir

Saeid Asadi
PhD in Information Technology; Associate Professor; Department of Information Science and Knowledge Studies;
Shahed University; Tehran, Iran Email: S.asadi@shahed.ac.ir

Soheil Ghanjefar
PhD in Electrical Engineering; Professor; School of Electrical Engineering; Iran University of Science and Technology;
Tehran, Iran; Email: s_ganjefar@basu.ac.ir

Iranian Journal of Information Processing and Management

Iranian Research Institute for Information Science and Technology (IranDoc)
ISSN 2251-8223
eISSN 2251-8231
Indexed by SCOPUS, ISC, & LISTA
Vol. 37 | No. 1 | pp. 255-276
Autumn 2021

Received: 24, Nov. 2020 **Accepted: 14, Mar. 2021**

Abstract: In designing search engines, it is important to check the database you are studying and make a connection with it.
This study is an applied one conducted by using observation method. What is used in this study is a case study.
The pattern of most searches is generally partial and specific in most cases. Participants begin their search with general information such as introducing and reviewing facts, and then focusing on specific aspects. In some cases, users come up with new ideas while searching. According to the results, general background, subject knowledge, time range and tools available influence desired response.
According to the analysis of findings and results, it is suggested that information behavior using neural network is more accurate in identifying information skills, barriers, goals and motivation and determining and predicting resources and services. Information and ways of accessing information should be compared with the results of present study. Based on the results, it seems that specialized information retrieval training for all classes of users is needed to increase information skills of current and future users.
Current search engines retrieve only part of relevant documents in a collection. Better models are needed to overcome the huge volume

* Corresponding Author

of documents. The proposed model in this way enables improved data retrieval in a short period of time. It is also possible that search model is also updated each time with user searches and results in more accurate results.

Keywords: Information Retrieval, User, Behavior Feedback, Neural Network, Search Engine

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

مدل‌سازی رفتار اطلاعاتی کاربران پایگاه‌های اطلاعاتی با روش شبکهٔ عصبی با تأکید بر تعاملات پیشین آن‌ها با نتایج جست‌وجو

فتانه وهابی
دانشجوی دکتری کتابداری و اطلاع‌رسانی؛ پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)؛ ایران، تهران؛
پدیدآور رابط vahabi@students.irandoc.ac.ir

سعید اسدی
دکتری فناوری اطلاعات؛ دانشیار؛ دانشگاه شاهد؛ ایران، تهران asadi.s@gmail.com

سهیل گنجه‌فر
دکتری برق؛ استاد؛ دانشگاه علم و صنعت؛ ایران، تهران؛
s_ganjefar@basu.ac.ir

پژوهش‌نامهٔ پردازش و مدیریت اطلاعات

دریافت: ۱۳۹۹/۰۹/۰٤ **پذیرش:** ۱۳۹۹/۱۲/۲٤ مقاله برای اصلاح به مدت ۴ ماه نزد پدیدآوران بوده است.

نشریهٔ علمی | رتبهٔ بین‌المللی
پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)
شاپا (چاپی) ۸۲۲۳-۲۲۵۱
شاپا (الکترونیکی) ۸۲۳۱-۲۲۵۱
نمایه در LISTA، ISC، و SCOPUS
jipm.irandoc.ac.ir
دورهٔ ۳۷ | شمارهٔ ۱ | ص‌ص ۲۵۵-۲۷۶
پاییز ۱٤۰۰

چکیده: در طراحی موتورهای جست‌وجو بررسی پایگاه دادهٔ مورد مطالعه و ایجاد ارتباط با آن دارای اهمیت است. در پژوهش حاضر با بررسی تعاملات کاربران پایگاه‌های اطلاعاتی با سیستم در حین جست‌وجو، به مدل‌سازی رفتار اطلاع‌یابی آن‌ها پرداخته می‌شود.

این پژوهش یک مطالعهٔ کاربردی است که با استفاده از شبکهٔ عصبی اجرا شده است. داده‌ها از طریق مشاهدهٔ رفتار کاربران در استفاده از پایگاه دادهٔ «ایرانداک» و مطالعهٔ لاگ کاربران گردآوری شده است.

الگوی اغلب جست‌وجوهای انجام‌شده در بیشتر موارد به‌صورت «کلی به جزیی» و «اختصاصی» است. شرکت‌کنندگان جست‌وجوی خود را با اطلاعات کلی در مورد موضوع، مانند معرفی و بررسی حقایق، شروع نموده و سپس، هر یک بر روی جنبه‌های خاصی از موضوع تمرکز کردند. در بعضی موارد، کاربران در حین جست‌وجو ایده‌های جدیدی به‌دست آوردند.

با توجه به نتایج به‌دست‌آمده می‌توان چنین استنتاج کرد که بر اساس مدل استخراجی، عواملی چون پیشینهٔ کلی فرد در مورد موضوع، دانش موضوعی، محدودهٔ زمان و ابزارهای در اختیار، پاسخ مورد نظر در

جست‌وجو را تحت تأثیر قرار می‌دهد.
موتورهای جست‌وجوی کنونی تنها بخشی از مدارک مرتبط با موضوع را در یک مجموعهٔ داده بازیابی می‌کنند. برای دسترسی به مطالب بیشتر و مناسب‌تر در حجم عظیم داده‌ها لازم است مدل‌های بهتری مورد بررسی قرار گیرد. روش شبکهٔ عصبی پیشنهادی در این مطالعه امکان بهبود بازیابی اطلاعات را در مدت زمان کم فراهم می‌آورد. همچنین، با این روش می‌توان این امکان را به‌وجود آورد که مدل جست‌وجو هر مرتبه با جست‌وجوی کاربران به‌روزرسانی شده و نتایج کامل‌تر و دقیق‌تری به‌دست آید.

کلیدواژه‌ها: بازخورد رفتاری، بازیابی اطلاعات، شبکهٔ عصبی، کاربر، موتور جست‌وجو

۱. مقدمه

در عصر الکترونیک و دیجیتال دسترسی به اطلاعات صحیح به کمک موتورهای جست‌وجو، بررسی پایگاه دادهٔ مورد مطالعه، و ایجاد ارتباط با آن از اهمیت بالایی برخوردار است. در طراحی موتورهای جست‌وجو دسترسی به اطلاعات صحیح در صفحات وب و بررسی مشابهت داده‌های مختلف در پایگاه داده به‌منظور بازیابی اطلاعات مورد توجه قرار می‌گیرد. در بازیابی اطلاعات توسط کاربر، کلمهٔ مورد جست‌وجو به‌صورت یک کلیدواژه در نظر گرفته می‌شود. در این روند چنانچه اطلاعات به‌صورت دقیق و با سرعت بالا و متناسب با کلمهٔ مورد جست‌وجو بازیابی شود، بسیار ثمربخش خواهد بود.

با توجه به رشد روزافزون اطلاعات دیجیتال، وابستگی محققان به سیستم‌های مختلف برای دریافت اطلاعات مورد نظر به‌طور چشمگیری افزایش یافته است. فناوری‌های مرتبط با بازیابی اطلاعات بیشتر بر مدل‌های ریاضی متن‌مدار تأکید داشته و بر چگونگی تعامل کاربر با اطلاعات در محتوا کمتر توجه داشته است. از آنجا که نیاز به بازیابی اطلاعات صحیح و سریع در کاربردهای روزانه تا کاربردهای تخصصی روبه‌افزایش است، فعالیت‌های انسانی زیادی لازم است تا اطلاعات مناسب بازیابی شود. مفید بودن اطلاعات و موجود بودن و دسترس‌پذیری آن مشکل اصلی در بازیابی اطلاعات است (Jarvelin &Ingwersen 2004; Jarvelin 1986).

اطلاع‌یابی می‌تواند به کمک کلمات تایپ‌شده بر روی صفحه‌کلید، صوت و یا تصویر مورد نظر فرد آغاز شود. هدف اصلی در بسیاری از جست‌وجوهای علمی، دستیابی به اطلاعاتی است که علاوه بر هماهنگی با عبارت مشخص‌شدهٔ کاربر، متناسب با تخصص و خواستهٔ وی ارائه شود. به کمک این روند می‌توان پایگاه‌ها و اطلاعات مناسب را با

دقت بالا در کمترین زمان ممکن در اختیار پژوهشگر قرار داد. دریافت مشخصات کاربر از نظر علمی در کنار کلمات کلیدیِ مد نظر وی به‌عنوان یک پیش‌فرض مناسب در بازیابی اطلاعات صحیح قابل تأمل است. باید توجه داشت که افراد دارای علایق علمی مختلف هستند و ممکن است بخش‌های خاص مرتبط با اطلاعات خود را مورد بررسی قرار دهند. بنابراین، چنانچه بتوان مسیر علاقهٔ علمی فرد را تعیین نمود و هماهنگ با آن عمل کرد، امکان بازیابی اطلاعات با کارایی بیشتر و در زمان کمتر فراهم می‌گردد. به این منظور، لازم است روند اطلاع‌یابی پژوهشگر به‌گونه‌ای بررسی شده و انتخاب‌های بعدی برای وی با توجه به این امر پیشنهاد گردد. از آنجا که بهترین روش در یافتن مسیر مطالعهٔ یک فرد بررسی مطالعات قبلی اوست، می‌توان این فرایند را به کمک شبکه‌های عصبی مصنوعی انجام داد.

بررسی مطالعات انجام‌گرفته در داخل کشور نشان می‌دهد که پژوهش‌هایی که با استفاده از تکنیک شبکهٔ عصبی در داده‌کاوی به دسته‌بندی کاربران کتابخانه‌ها بر مبنای نیاز این افراد در دستیابی به اطلاعات و تعیین رفتارهای اطلاعاتی آن‌ها پرداخته باشند، بسیار اندک است.

پژوهش‌های انجام‌شده در خارج از کشور که مبتنی بر این رویکرد هستند، به‌طور عمده به بخش‌بندی بازار در علم اقتصاد، به‌ویژه مدل‌های ایجاد وفاداری فرد به خرید یک محصول از طریق دسته‌بندی پرداخته و با استفاده از این کار توانسته‌اند کارایی خدمات مشتریان را افزایش و هزینه‌های عملیاتی و بازاریابی را کاهش داده و در نهایت، به افزایش سود کل و رضایت مشتریان دست یابند. بنابراین، انجام پژوهش حاضر از آن جهت ضرورت و اهمیت می‌یابد که در پژوهش‌های صورت‌گرفته، استفاده از رویکرد شبکهٔ عصبی در داده‌کاوی که دارای قابلیت بالایی در کشف الگوهای ناشناخته در رفتار و نیازهای کاربران سازمان‌ها و نهادهای مختلف دارد، در کتابخانه‌ها کمتر مورد توجه قرار گرفته است. افزون بر این، مطالعه بر روی شناخت نیازهای اطلاعاتی جامعهٔ مورد نظر در ایران تاکنون به‌طور خاص صورت نگرفته است.

کتابخانه‌ها نیز مانند سایر سازمان‌ها خواهان آن هستند که اطلاعات گردآوری‌شده از تراکنش‌های مختلف کاربران خود را در بخش‌بندی داده‌ها به‌کار گیرند. اگرچه جست‌وجو در بین حجم عظیمی از داده‌ها امکان‌پذیر است، لکن کتابخانه را در گردآوری، آماده‌سازی، تحلیل و تفسیر داده‌هایی که از انواع متعدد تراکنش‌های محیط‌های فیزیکی

و مجازی ایجاد می‌شود، با مشکلاتی روبه‌رو می‌سازد. به‌عنوان برخی از ابزارهای مورد استفاده برای کشف الگوهای مربوط به جست‌وجو می‌توان به روش‌های دسته‌بندی، خوشه‌بندی، قواعد تلازمی[1]، درخت تصمیم و شبکه‌های عصبی اشاره کرد. بنابراین، با توجه به نیاز مبرم مدیران ارشد و نخبگان به اطلاع از نیاز کاربران در عرصه‌های مختلف،جهت برنامه‌ریزی‌های کوتاه‌مدت و بلندمدت، این پژوهش به بررسی مدل‌سازی رفتار اطلاع‌یابی کاربران پایگاه‌های اطلاعاتی با روش شبکهٔ عصبی و توجه بر تعاملات پیشین آن‌ها در حین جست‌وجو پرداخته است.

از سوی دیگر، با ارائهٔ مدل‌های جامع و آزمون‌شده، این امکان برای کتابخانه‌ها و مراکز اطلاع‌رسانی فراهم می‌شود که سریع‌تر از پیش، نظارت دقیق‌تر و صحیح‌تری به اهداف، برنامه‌ها، کیفیت خدمات و منابع، نیازهای کاربران، چگونگی رفتار اطلاع‌یابی آن‌ها، تعیین نقاط قوت و ضعف مجموعه، گردآوری منابع و ارائهٔ خدمات جدید و تحول رویّه‌ها داشته باشند و بتوانند به فرایند جست‌وجو و کشف مدل‌های گوناگون، مختصرسازی‌ها و اخذ مقادیری از مجموعه‌ای از داده‌های معلوم بپردازند.

مدیران کتابخانه تنها زمانی می‌توانند به موفقیت در پاسخگویی به نیازهای مراجعه‌کنندگان خود امید داشته باشند که بتوانند رفتارهای آیندهٔ آن‌ها را پیش‌بینی کنند و نیاز کاربران و راه‌های مناسب برای رفع آن را شناسایی نمایند. برای این منظور، کتابخانه‌ها می‌بایست داده‌های کامل از فعالیت‌های گذشتهٔ کاربران خود داشته باشند تا بتوانند از آن‌ها در ارتقای کیفیت خدمات و منابع خود استفاده نمایند.

۲. پیشینه

تکنیک‌های داده‌کاوی در داخل کشور به‌طور عمده در سازمان‌های تجاری، بانک‌ها، صنایع و شرکت‌های خصوصی به‌کار گرفته شده است. در پژوهش‌های مبتنی بر این رویکرد اتکای اصلی پژوهشگران بر دسته‌بندی مشتریان بر اساس پروفایل‌ها و رفتارهای آنان در استفاده از محصولات و خدمات سازمان‌های مختلف صورت گرفته است. پژوهش‌های انجام‌شده در خارج از کشور که مبتنی بر این رویکرد هستند، به‌طور عمده به بخش‌بندی بازار، به‌ویژه مدل‌های ایجاد وفاداری از طریق دسته‌بندی پرداخته و

1. rules of communication

با استفاده از این کار توانسته‌اند کارایی خدمات مشتریان را افزایش و هزینه‌های عملیاتی و بازاریابی را کاهش دهند و در نهایت، به افزایش سود کل و رضایت مشتریان دست یابند. البته، مطالعات دیگری در مورد روش‌های اطلاع‌یابی کاربران در جست‌وجوی اطلاعات انجام شده است.

«ژانگ» و همکاران مدل درک رفتار اطلاع‌یابی کاربران در هنگام کلیک را ارائه دادند. آن‌ها دریافتند که کلیک کاربر با ترجیح کاربر مطابقت ندارد و علایق کاربر به‌طور مرتب تغییر می‌کند. این مدل در واقع، روشی را برای برآورد علاقه و رضایت کاربر نسبت به تنوع عناوین فراهم می‌کند (Zhang et al. 2018). «دو دیگو و اسک» بیان می‌کنند که تحقیقات مبتنی بر جست‌وجوی وب باید نظر کاربر را نیز در نظر بگیرد. بر این اساس رویکرد تنظیم لیست‌های نتایج جست‌وجوی تولیدشده توسط موتورهای جست‌وجوگر وب را پیشنهاد می‌دهند (de Diego & Eske 2019). «ژو» و همکاران بیان می‌کنند که روش‌های قدیمی، مسئلهٔ جست‌وجو را به‌عنوان روشی برای انتخاب و رتبه‌بندی اسناد بررسی می‌نمایند. در این مطالعه با تمرکز بر موضوع جست‌وجو، یک روش جست‌وجو مبتنی بر یادگیری عمیق با مدل‌سازی جست‌وجو برای موضوعات ارائه می‌شود (Zhou et al. 2019).

با توجه به اینکه بازخورد تأثیرگذار، خواه مثبت یا منفی، بسیار کمک‌کننده است و مزایای بسیاری برای ارسال‌کننده، دریافت‌کننده و سازمان‌ها خواهد داشت، پژوهش‌های زیادی از جمله Lavrenko & Croft (2001); Salton & Buckley (1990); Robertson & Jones (1976); Jones & Diaz (2007); Paramita, Sanderson & Clough (2009); Radlinski & Dumais (2006); Tao and Zhai (2006); Sanderson et al. (2009); در مدل‌های بازیابی خود از بازخورد استفاده کرده‌اند.

در اغلب مدل‌های زبانی ارائه‌شده توسط محققان، بازخورد الگوریتم توسط مدل‌های پرسش زبانی (Tao & Zhai 2006; Sanderson & Tang 2009; Clough 2009)، یا مدل ربط (Jones & Diaz 2007) انجام شده است. «کوهن» و «روی» در تحقیقات خود بیان کردند که بازخورد فعال به‌طور معمول در آموزش ماشین کارایی خوبی دارد (Cohn 1996; Roy 2001). بسیاری از محققان در زمینهٔ حذف نتایج زاید فعالیت داشته‌اند (Bookstein 1983; Carbonell & Goldstein 1998; Chen Karger 2006; Tang & Sanderson 2010; Yao et al. 2013). پژوهشگرانی همچون «یااو و چن» در صدد بودند تنوع را بالا برده و از حشو و زواید

صرفنظر نمایند (Yao et al. 2013). آن‌ها در بحث تنوع، از ریخت‌شناسی برای طبقه‌بندی پرسش‌ها و اسناد و ایجاد یک مجموعه جواب استفاده نموده‌اند.

«بدر، محمداسماعیل و حنیف» در پژوهش خود به دسته‌بندی کاربران هدف کتابخانهٔ مرکزی «دانشگاه صنعتی اصفهان» از لحاظ شناسایی رفتارها و نیازهای اطلاعاتی آنان پرداخته‌اند. آن‌ها دریافتند که بیشترین منابع اطلاعاتی مورد نیاز کاربران کتابخانه، ترجمهٔ کتاب‌های لاتین و کمترین نیاز آنان گزارش‌ها و طرح‌های پژوهشی است. مؤثرترین خدمات در رفع نیازهای اطلاعاتی آنان، توانایی کتابداران و متخصصان اطلاع‌رسانی در بازیابی اطلاعات و آگاهی آنان از منابع و خدمات اطلاعاتی و کم‌اثرترین خدمات،آشنایی با انواع منابع و خدمات اطلاعاتی در زمینهٔ رشتهٔ تخصصی خود و امکان استفاده از منابع اطلاعاتی به‌صورت فایل‌های دیجیتال و بدون نیاز به حضور در کتابخانه شناسایی شده است (۱۳۹۶).

«محمداسماعیل و نعیمی» در پژوهش خود به بررسی خوشه‌بندی و مطالعهٔ تطبیقی نیازهای اطلاعاتی و رفتار اطلاع‌یابی دانش‌پژوهان حوزه و دانشگاه خراسان رضوی با استفاده از الگوریتم شبکهٔ عصبی پرداختند. مؤثرترین مؤلفه در هدف و انگیزهٔ اطلاع‌یابی دانش‌پژوهان حوزه انجام فعالیت‌های پژوهشی، فرهنگی و مذهبی و در دانش‌پژوهان دانشگاه ارتقای کیفیت تدریس ارزیابی بود (۱۳۹۵).

«راد» در تحقیق خود جهت بررسی رفتارها و رویکردهای اطلاع‌یابی دانشجویان تحصیلات تکمیلی دانشگاه آزاد اسلامی هنگام استفاده از شبکهٔ جهانی وب دریافتند که رفتار اطلاع‌یابی جامعهٔ پژوهش مطابق الگوی «الیس»[1] است، لیکن تفاوت معناداری میان رفتار دانشجویان در گروه‌های مختلف آموزشی وجود دارد (۱۳۸۸).

«یاری و احمدی» به‌منظور ارائهٔ تصویری روشن، علمی و واقعی از وضعیت رفتار اطلاع‌یابی در ایران متون منتشر شده در باب مباحث و بنیان‌های نظری و پژوهش‌های انجام‌شده در حوزهٔ رفتار اطلاع‌یابی در ایران را با رویکردی تحلیلی و کیفی بررسی کرده‌اند. بررسی‌های آن‌ها مبیّن آن است که در طی ۱۷ سال که از طرح بحث رفتار اطلاع‌یابی در ایران می‌گذرد، از لحاظ کیفی و کمی روند انتشار منابع رشد چشمگیری داشته است. کمتر از ۴ درصد از متون منتشرشده در حوزهٔ رفتار اطلاع‌یابی ترجمه شده‌اند.

1. Ellis

بیشتر منابع این حوزه در قالب پژوهش به بررسی رفتار اطلاع‌یابی افراد پرداخته‌اند (۱۳۹۳).

«نعیمی، محمداسماعیل و حیدری» با استفاده از شبکهٔ عصبی به بررسی تعیین رفتار اطلاع‌یابی دانش‌پژوهان سطح چهار حوزهٔ علمیهٔ خراسان رضوی پرداختند. آنها دریافتند که برای رفع نیازهای اطلاعاتی دانش‌پژوهان در حوزه، تدارک منابع مورد نیاز در کنار فراهم‌آوری زمینهٔ تصمیم‌گیری اطلاعات‌محور برای مدیران کتابخانه‌ها مثمرثمر خواهد بود (۱۳۹۷).

«فوهر» با استفاده از جست‌وجوی تعاملی در سیستم‌های بازیابی اطلاعات، یک مدل کلی برای تشخیص قدم بعدی سیستم با توجه به عملکرد قبلی سیستم ارائه داد (Fuhr 2008).

در کنار یادگیری ماشین از دههٔ ۹۰ میلادی، کار بر روی تحلیل اجزای اصلی در کاهش تعداد دادگان توسط Billsus & Pazzani (1998) و بهبود دسته‌بندی داده‌های ورودی با کمک خوشه‌بندی توسط Kohrs & Merialdo (1999) انجام شد.

«جاهرر، توشر و لگنشتاین» برای بازیابی اطلاعات از روش‌های مختلف از جمله رگرسیون خطی، kامین نزدیک‌ترین همسایگی (KNN)[1] و درخت تصمیم‌گیری استفاده نمودند. البته، مدل شبکهٔ عصبی به‌کاررفته یک مدل ساده انتخاب شده بود. در این مدل چگونگی بهبود مدل‌سازی دادگان ورودی متناسب با مدل شبکهٔ عصبی مورد بحث قرار نگرفت (Jahrer, Töscher & Legenstein ۲۰۱۰). همچنین، Hauff (2010); Jones & Diaz (2007) از شبکه‌های عصبی به‌جای رگرسیون خطی برای بازیابی بهتر استفاده نمودند.

در زمینهٔ بازیابی اطلاعات یک سیستم یکپارچه که بتواند با تکنیک‌های مختلف کار کند و نیز بتواند درخواست‌های مختلف را هوشمندانه تشخیص دهد، لازم است (Shani and Gunawardana 2011) سیستم‌های بازیابی اطلاعات هم‌اکنون از نظر تجاری و جامعهٔ تحقیقاتی مورد توجه هستند. در بسیاری از موارد کاربری که مایل به استفاده از یک پایگاه است باید بین مجموعه‌ای از رویکردهای مورد نظر یکی را انتخاب کند. اولین قدم برای انتخاب یک پایگاه این است که تصمیم گرفته شود روی کدام ویژگی‌های برنامه تمرکز شود. در واقع، پایگاه‌ها دارای خواص مختلفی هستند که ممکن است بر تجربهٔ کاربر تأثیر بگذارند. برنامهٔ متداول در بازیابی اطلاعات شامل جست‌وجوی محتوا،

1. k-nearest neighbour

طبقه‌بندی متن، مدیریت محتوا و پرسش و پاسخ است. در بیشتر این موارد رویکردهای آماری و یا یادگیری ماشین به‌کار می‌رود.

تحقیقات نشان داده است که روش شبکهٔ عصبی برای داده‌ها با حجم کمتر، نتایج بهتری تولید می‌کند (Abu Abbas 2008).

تحقیقات متعدد انجام‌شده در زمینهٔ بازخورد در بازیابی اطلاعات، نمایانگر اهمیت موضوع است. همچنین، این پژوهش‌ها نشان می‌دهند که استفاده از بازخورد موجود در سیستم بسیار کارآمد است.

۳. روش پژوهش

مطالعهٔ حاضر نوعی پژوهش اکتشافی است که داده‌های مورد نظر در آن با استفاده از مشاهدهٔ رفتار کاربران پایگاه‌های اطلاعاتی «پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)» گردآوری شده است. با این توصیف، پژوهش حاضر یک مطالعهٔ کاربردی موردی است که در آن از تکنیک‌های شبکه‌های عصبی برای تعیین مدل رفتار اطلاع‌یابی کاربران استفاده شده است.

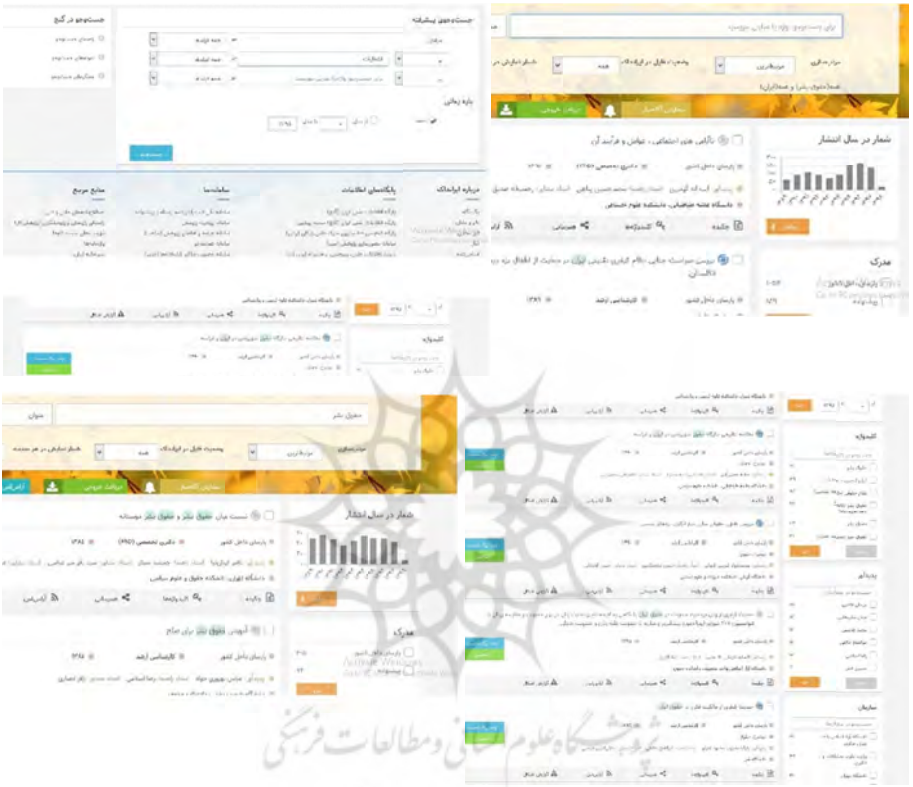
این مطالعه در چندین گام مختلف و متوالی انجام می‌شود. در گام اول، نظام‌های بازیابی اطلاعات موجود در «پژوهشگاه» بررسی شدند. در مرحلهٔ دوم، روابط تعاملی موجود مورد بررسی قرار گرفت و بر این اساس، مدلی ثانویه مبتنی بر تعامل کاربران ارائه شد. در مرحلهٔ سوم، از طریق مشاهده و بررسی رفتارهای اطلاعاتی کاربران مدلی از بازیابی اطلاعات آن‌ها ترسیم گردید. برای اعتباریابی این مدل نیز از نظر متخصصان استفاده شد.

جامعهٔ مورد مطالعه، کاربران بخش جست‌وجو در پایگاه «(گنج) پژوهشگاه علوم و فناوری اطلاعات (ایرانداک)» است. در این مطالعه سعی شد روند بازیابی اطلاعات توسط افراد مورد بررسی قرار گیرد. با توجه به حجم عظیم و متنوع کاربران، محدودهٔ مطالعاتی با استفاده از روش اشباع داده به‌عنوان استاندارد طلایی انجام گرفت.

گردآوری اطلاعات در این پژوهش برای رسیدن به اهداف مورد نظر با مشاهدهٔ رفتار کاربران انجام شده است. در ادامه، ۱۲۰ لاگ جست‌وجو با استفاده از نرم‌افزار «اسکرین کپچر»[1] بررسی گردیده است (شکل ۱). این نرم‌افزار امکان ذخیرهٔ اطلاعات شامل حرکت

1. screen capture

نشانگر، انتخاب فهرست (منو)، نتایج نمایش داده‌شده و تمام فعالیت‌های روی اسکرین را فراهم می‌نماید. بنابراین، می‌توان تمام فعالیت‌های کاربر در فرایند جست‌وجو را برای تعیین جنبه‌های مختلف استراتژی جست‌وجوی اطلاعات مد نظر قرار داد.



شکل ۱. اسکرین‌شات از برخی از تصاویر ثبت‌شده

این مطالعه با همکاری ۳۰ نفر از دانشجویان تحصیلات تکمیلی با پیشینهٔ تخصص در جست‌وجوی اطلاعات (دارای گواهی شرکت در کارگاه مهارت در جست‌وجوی اطلاعات) انجام شد. از این ۳۰ شرکت‌کننده، ۱۵ نفر زن و ۱۵ نفر مرد از گروه سنی ۳۵-۲۹ بودند. ارزیابی گواهی‌های شرکت در کارگاه‌های جست‌وجوی اطلاعات نشان داد که مشارکت‌کنندگان، جست‌وجوگران ماهری هستند. با توجه به وجود دانش سیستمی (مهارت‌های جست‌وجو) در کاربران در این مطالعه، می‌توان به سه روش پرس‌وجو، مرور، و کلیک کردن، اطلاعات بازخورد کاربر در جست‌وجو را دریافت نمود و از آن در

آموزش شبکهٔ عصبی بهره برد. اطلاعات بازخورد کاربران دربارهٔ جملات چکیدهٔ مطالب انتخاب‌شده، در شبیه‌سازی استفاده شده است.

ویژگی‌های اطلاعات دریافتی از بازخورد کاربران که مورد توجه پژوهشگران اطلاع‌رسانی است، به اجمال بیان می‌شود. این موارد با بررسی‌های انجام‌شده در این مطالعه به‌دست آمده و برخی از آن‌ها در شبیه‌سازی الگوریتم به‌کار رفته است.

اطلاعات دریافتی از بازخورد کاربران

کاربران به‌طور معمول، تنها با بررسی عنوان و خلاصهٔ نشان داده‌شده در نتیجهٔ جست‌وجو دربارهٔ مرتبط‌بودن یا نبودن نتایج تصمیم‌گیری می‌کنند، به‌طوری که حتی گاهی نیازی به دیدن متن اصلی نیست. این جنبه از رفتار کاربران با استفاده از ویژگی‌های حاصل از نحوهٔ پرس‌وجو و رابطهٔ آن با خلاصهٔ نمایش داده‌شده از متن، به‌دست می‌آید. این ویژگی شامل میزان همپوشانی کلمات پرس‌وجو و عنوان نتیجه، نسبت کلمات مشترک بین پرس‌وجو و خلاصهٔ نتیجه و از این دست است که در شبیه‌سازی نیز مورد بحث قرار گرفته است.

جنبه‌های سادهٔ تعامل کاربر با صفحات وب می‌تواند به‌عنوان بازخوردهای ضمنی استفاده شود. برای مثال، ویژگی‌هایی مانند زمان ساکن ماندن روی صفحه یا روی دامنه و میزان انحراف از زمان ماندن مورد انتظار برای یک پرس‌وجو از مواردی است که می‌توان از آن‌ها بهره برد.

کلیک‌های کاربران یکی از شاخه‌های خاص در تعامل کاربر با موتور جست‌وجوست. به‌عنوان مثال، برای هر زوج پرس‌وجو تعداد کلیک‌هایی که روی هر نتیجه انجام شده و اینکه آیا نتایج قبلی یا بعدی یک نتیجه توسط کاربر کلیک شده یا خیر، به‌عنوان بازخورد ضمنی قابل استفاده است. مرور مستندات به‌خوبی مسیر مرور کاربر در وب را نشان می‌دهد و در نتیجه می‌تواند برای تعیین مرتبط‌بودن یک مستند به پرس‌وجو به‌خوبی عمل کند. هرچه یک مستند بیشتر مرور شده باشد و زمان مرور آن بیشتر شود، احتمال مرتبط‌بودن آن با پرس‌وجو بیشتر است.

جدول ۱، این موارد را به تفصیل برای برخورداری از بررسی رفتار اطلاع‌یابی کاربران نشان می‌دهد.

جدول ۱. ویژگی‌های دریافت اطلاعات بازخورد کاربران

ویژگی‌های سیر کلیک‌ها	
موقعیت	موقعیت سند در میان نتایج
تناوب کلیک	تعداد کلیک‌ها برای این پرسش
احتمال کلیک	احتمال یک کلیک برای این پرسش
بعدی کلیک شده است	۱ اگر برروی بعدی کلیک شده، ۰ در غیر این صورت
قبلی کلیک شده است	۱ اگر بر روی قبلی کلیک شده، ۰ در غیر این صورت
بالا کلیک شده است	۱ اگر بالا کلیک شده، ۰ در غیر این صورت
پایین کلیک شده است	۱ اگر پایین کلیک شده، ۰ در غیر این صورت
ویژگی‌های مرورکردن	
زمان ساکن ماندن	زمان ساکن ماندن درصفحه
زمان ساکن ماندن بعد از انتخاب	زمان کلی ساکن ماندن بعد از جست‌وجو
لینک پیگیری‌شده	۱ اگر لینک پیگیری نتیجه باشد، ۰ در غیر این صورت
میانگین زمان	میانگین زمان ساکن ماندن بر روی یک صفحه
ویژگی‌های پرس‌وجو	
همپوشانی عنوان	تشابه کلمات بین پرسش و عنوان
همپوشانی چکیده	تشابه کلمات بین پرسش و توضیحات
همپوشانی پرسش–دامنه	تشابه کلمات بین پرسش و دامنه
طول پرسش	تعداد کلمات در پرسش

برای بررسی ایدهٔ به‌کارگیری هوش مصنوعی در بازیابی اطلاعات و رفتار کاربران در حین جست‌وجو، مشخصات مختلفی که در بررسی رفتار کاربران سودمند است، در جدول ۱، آمده است. همچنین، به‌منظور برآورد کارایی این ایده، ویژگی‌های پرس‌وجو برای دستیابی به ارتباط کلمات کلیدی کاربران و نتایج بازیابی‌شدهٔ متناسب با آن در یک نرم‌افزار شبیه‌سازی شده است. با توجه به نتایج این پژوهش ادامهٔ بررسی ایدهٔ حاضر برای مطالعات آتی با به‌کارگیری دیگر پارامترهای عنوان‌شده در این مقاله برای بازیابی رفتار کاربران بسیار مفید خواهد بود.

در ادامه، جایگاه قرارگیری کلمات کلیدی واردشده توسط کاربر در جمله مورد توجه قرار می‌گیرد. در واقع، بنابر مطالعهٔ Zheng et al. (2017) برای تعیین میزان صحت

انتخاب مناسب نتیجهٔ بازیابی‌شده، موقعیت و شباهت کلمات موجود در جملات چکیدهٔ مطالب نسبت به کلمات کلیدی ورودی کاربر سنجیده شده است.

مجموعهٔ دادهٔ ورودی شبکه، شامل چکیدهٔ نشانه‌گذاری‌شده از مقالاتی از پایگاه دادهٔ «ایرانداک» است و در آن چکیدهٔ هر مقاله به جملاتی مجزّا تفکیک شده است. برای نمونه، دو کلمهٔ کلیدی ورودی توسط کاربر را به‌صورت توکن "keywordA" و "keywordB" در نظر بگیرید. میزان معیار شباهت از ضرب داخلی بردار کلمات در جمله با بردار کلمات ورودی و به‌کارگیری یک تابع بیشینه هموار برای تبدیل این معیار به مقادیر احتمالاتی محاسبه می‌شود. بنابراین، این معیار نشان‌دهندهٔ رابطهٔ هر کلمه با کلمات موجود خواهد بود.

بر مبنای مطالعهٔ Zheng et al. (2017) مقدار شباهت با پارامتر Sml_i از رابطهٔ زیر محاسبه شده است:

$$Sml_i = softmax(\frac{u_{w_i}.u_{e_j}}{m}) \quad (۱)$$

که در این رابطه $u_{e_j}, j \in \{1,2\}$ بردار کلمات کلیدی ورودی و u_{w_i} بردار کلمات مربوط به لغات جملات است که در یکدیگر ضرب داخلی شده‌اند. m اندازهٔ ابعاد بردار کلمات است. رابطهٔ فوق شباهت هر یک از کلمات جمله را با کلمهٔ ورودی نشان می‌دهد.

پس از محاسبهٔ شباهت کلمهٔ nام با کلمهٔ ورودی اول به‌صورت Sml_i^1 و شباهت کلمهٔ nام با کلمهٔ ورودی دوم به‌صورت Sml_i^2 از متوسط این دو مقدار برای بیان میزان شباهت کلمهٔ nام در جمله با کلمات ورودی به‌صورت Sml_i یاد می‌شود. همچنین، پس از مشاهدهٔ جمله به‌صورت کامل، معنا و جایگاه کلمات آن نسبت به کلمات کلیدی مورد نظر مورد توجه قرار می‌گیرد. در واقع، موقعیت کلمات مختلف در کنار میزان شباهت آن‌ها نسبت به کلمات ورودی، اهمیت کلمات و در پی آن جملات متن را در هماهنگی با کلمات جست‌وجوشدهٔ ورودی نشان می‌دهد.

برای دستیابی به نتایج بهتر، اطلاعات مربوط به موقعیت نیز در کنار اطلاعات مربوط به شباهت کلمات به‌کار رفته است. رابطهٔ زیر ترکیب این دو مقدار را در تابع نهایی نشان می‌دهد.

$$\gamma_i = softmax(\frac{W_1 d_i^1 + W_2 d_i^2 + W_{sml} Sml_i}{3}) \quad (۲)$$

در رابطهٔ فوق Sml_i, d_i^1 و d_i^2 به‌ترتیب، نمایندهٔ معیار شباهت، فاصلهٔ کلمهٔ nام در جمله با کلمهٔ کلیدی اول و فاصلهٔ کلمهٔ nام در جمله با کلمهٔ کلیدی دوم است. W_{sml}, W_1, W_2 اوزانی

هستند که به کمک a مدل آموزش داده می‌شود. برای محاسبهٔ وزن نهایی کلمات جمله، خروجی شبکهٔ عصبی LSTM[1] نیز در رابطهٔ فوق اضافه خواهد شد. بنابراین، پارامتر λ_i برای هر کلمه به کمک روابط مربوط به شبکهٔ LSTM به‌صورت زیر تعریف می‌شود:

$$\lambda_i = \text{Softmax}\ \left[V \cdot Tanh\left(W_h.h_i + \gamma_i + b\right)\right] \qquad (۳)$$

λ_i معرف میزان اهمیت هر گام زمانی است و برای محاسبهٔ میانگین وزندار بردار ویژگی خروجی LSTM به‌کار خواهد رفت.

در مدل به‌کاررفته برای این مطالعه، بردارهای ویژگی حاوی اطلاعات هر جمله به‌عنوان ورودی به‌کار می‌رود. مهم‌ترین ویژگی ورودی، کلمات موجود در هر جمله است. در این ویژگی هر جمله T توسط بردار $\{wt_1, wt_2, \cdots, wt_n\}$ که در آن wt_i کلمهٔ i ام جمله است، معرفی می‌شود.

برای ایجاد ویژگی مربوط به کلمات، همهٔ جملات بررسی شده و یک مجموعهٔ کامل از کلمات موجود فراهم می‌شود. بر اساس این مجموعه هر کلمه از جمله به یک عدد صحیح در بردار متناظر می‌گردد. ویژگی نقش زبانی کلمات در جمله به کمک کتابخانهٔ NLTK[2] در زبان «پایتون» فراهم شده است. ورودی تابع نهایی جملات، متن، و خروجی آن نقش زبانی کلمات بر اساس جملات ورودی است. در نتیجه، نقش هر یک از کلمات به یک مقدار عددی نگاشت می‌شود. به‌طور معمول، موقعیت کلمات نسبت به کلمات کلیدی در جمله، در بار معنایی جمله و اهمیت آن تأثیر به‌سزایی دارد. این ویژگی نشان‌دهندهٔ فاصلهٔ هر کلمه در جمله نسبت به کلمات کلیدی است که با اختلاف تعداد کلمات از کلمهٔ کلیدی نسبت به هر یک از کلمات جمله تعیین می‌شود. فاصلهٔ کلماتی که بعد از کلمهٔ کلیدی قرار می‌گیرند، با علامت مثبت، و فاصلهٔ کلماتی که قبل از کلمهٔ کلیدی قرار می‌گیرند، با علامت منفی در نظر گرفته می‌شود. این ویژگی به‌صورت بردار $\{d_1^k, d_2^k, d_3^k, \ldots., d_n^k\}$ بیان خواهد شد، که در آن k=1,2 شمارهٔ مربوط به کلمهٔ کلیدی، n تعداد کلمات در جمله و d_i^k فاصلهٔ کلمهٔ iام در جمله از کلمهٔ کلیدی ورودی k است.

برای ورود داده‌های موجود به مدل از لایهٔ جاگذاری Word2Vec که در کتابخانهٔ Keras معرفی شده، استفاده شده است که کمک خواهد کرد مقادیر گسستهٔ نمایانگر مفاهیم از پیش تعیین‌شده، به‌صورت یک بردار پیوسته از اعداد ارائه شوند. در این حالت، مفاهیم

1. Long Short-Term Memory
2. Natural Language Toolkit

نزدیک به یکدیگر در فضای برداری نیز فاصلهٔ کمی نسبت به یکدیگر خواهند داشت. با قرار گرفتن بردارهای ورودی به‌دنبال یکدیگر، بردار ورودی تشکیل می‌شود که در آن هر ستون، ویژگی‌های مربوط به هر یک از کلمات را نشان می‌دهد.

بردار ویژگی ایجادشده به شبکهٔ BLSTM[1] وارد شده و پس از دستیابی شبکه به ارتباطات بین کلمات در مکان‌های مختلف، این اطلاعات به‌صورت یک بردار، ویژگی خروجی مدل را ایجاد خواهد کرد. در شبکهٔ BLSTM برای هر شبکه یک گام زمانی در نظر گرفته می‌شود که برای تمامی ورودی‌ها ثابت است. از این رو، لازم است با اضافه نمودن صفر به انتهای جملات، ابعاد آن‌ها را با بزرگ‌ترین جمله هماهنگ کرد.

در ادامه، به‌منظور آموزش، روابط موجود در داده‌های ورودی در بلوک بازگشتی از شبکهٔ LSTM دوطرفه با احتمال حذف تصادفی ۵۰ درصد در حالت بازگشتی استفاده شده است. همچنین، الگوریتم نزول در امتداد گرادیان برخط برای بهینه‌سازی پارامترها و تابع آنتروپی برای تابع هزینه به‌کار رفته است.

در مدل به‌کاررفته اندازه دسته‌ای[2] برابر ۲۰۰، تعداد دوران[3] برابر ۳۰۰، اندازهٔ LSTM برابر ۲۰۰، اندازهٔ بردار کلمات ۲۰۰ و اندازهٔ بردار فاصله ۱۰ در نظر گرفته شده است.

به‌منظور ارزیابی نتایج در روش پیشنهادی برای امکان پیشبرد این ایده مقدار پارامتر از مقادیر مثبت صحیح، منفی صحیح، مثبت غلط و منفی غلط استفاده شده است. بنابراین، مقادیر پارامتر F بر حسب دقت P و یادآوری R به‌صورت زیر محاسبه شده است:

$$F1 = 2\frac{P \times R}{P + R} \quad (۴)$$

$$P = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \quad (۵)$$

$$R = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \quad (۶)$$

در جدول زیر، مقدار سه پارامتر عنوان‌شده برای روش پیشنهادی با مطالعات قبلی مقایسه شده است. همان‌طور که مشاهده می‌شود، روش ارائه‌شده نسبت به روش‌های پیشین بهبود داشته است.

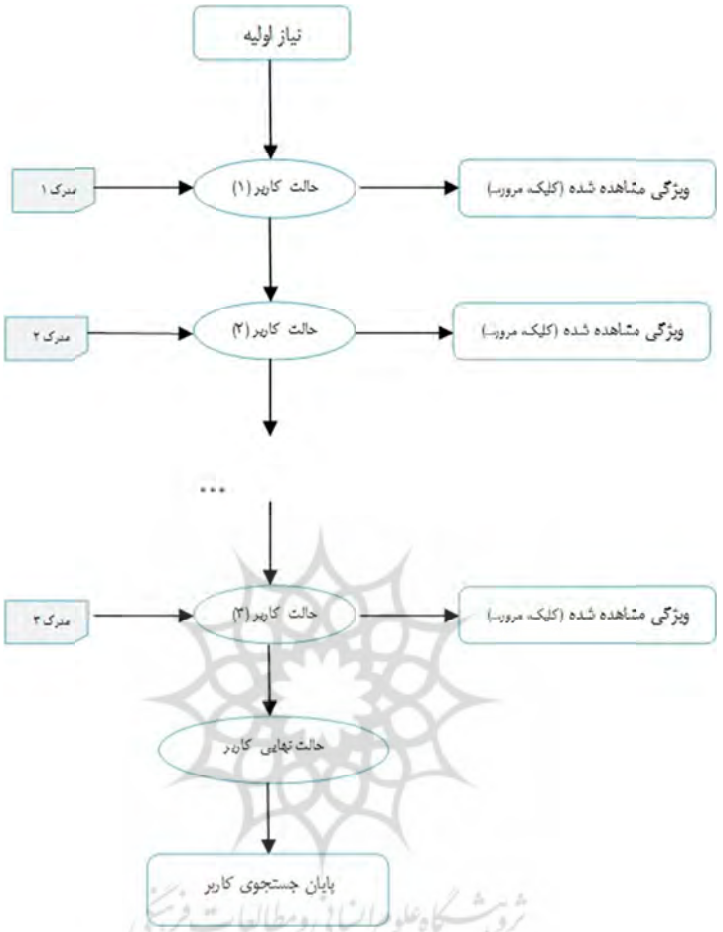
1. Bidirectional Long Short-Term Memory (BLSTM) 2. batch size 3. epoch

جدول ۲. مقادیر پارامترهای F, P و R در مطالعات پیشین و روش پیشنهادی

روش پیشنهادی	Yadav et al. (2021)	Zhang et al. (2018a)	
۷۷/۵	۷۲/۵۷	۷۳/۷	F1
۸۲/۵	۷۶/۵۲	۷۵	P
۷۳/۲	۶۹/۰۱	۷۲/۵	R

٤. یافته‌ها

پس از دستیابی به پارامترهای مناسب در الگوی جست‌وجوی افراد که به تفصیل در جدول ۱، ارائه شده، برخی از موارد این پارامترها در شبیه‌سازی مورد استفاده قرار گرفته‌اند. از آنجا که ایدهٔ استفاده از هوش مصنوعی در بازیابی اطلاعات نتایج خوبی به‌همراه دارد، این است که بررسی دیگر پارامترهای مؤثر در رفتار بازیابی اطلاعات توسط کاربران به گسترش این ایده کمک شایانی خواهد کرد. شکل ۳، مدل تعامل کاربران را در جست‌وجوی متوالی برای بازیابی نتیجهٔ معتبر توصیف می‌کند. کلمات کلیدی و یا پرسش نشان‌دهندهٔ نیازهای اطلاعاتی کاربر است. کاربر پس از بیان نیاز اطلاعاتی خود و ارسال این کلمات یا پرسش، لیستی از نتایج را دریافت خواهد کرد. این لیست شامل اقلام بازیابی‌شده در سیستم است و همهٔ موارد ممکن است تناسب کامل با نیاز کاربر را نداشته باشد و آن زمانی است که فرض کنیم کاربر به‌طور پیوسته لیست نتایج را از بالا به پایین بررسی کند. او در واقع، عملکرد ذهنی خود را نشان می‌دهد و آن آگاهی از موقعیت و کارهایی است که انجام می‌دهد. موارد ارائه‌شده در جدول ۱، عمل‌های رفتاری است که کاربر در حین جست‌وجو از آن بهره می‌برد. کاربر بر اساس ذهنیت خود مدرک را بررسی می‌کند. یا چکیدهٔ آن را مطالعه می‌کند، آن را ذخیره کرده و یا از آن پرینت تهیه می‌کند و یا در صورت عدم نیاز، از آن صرفنظر می‌کند. در واقع، کاربر با بررسی مدارک حالت فکری خود را نشان می‌دهد.



شکل ۳. مدل تعامل کاربران

می‌توان این‌طور عنوان کرد که بازیابی اطلاعات در مرکز یک مثلث قرار گرفته و عوامل خارجی، کاربر و تولیدکنندهٔ منبع سه‌گوشهٔ مثلث هستند. با توجه به نوع زیرساخت تولید منبع اطلاعاتی، توسعه و امکانات منبع بازیابی اطلاعات به‌دست‌آمده متغیر خواهد بود. همچنین، عواملی مانند اینکه کاربر با چه ویژگی‌های شخصیتی و محیطی و چه نیازی به منبع مورد نظر مراجعه کرده است نیز تعیین‌کننده هستند. عوامل خارجی مؤثر در تولید منبع و کاربر از جمله عوامل اجتماعی و فرهنگی از دیگر عوامل مؤثر در بازیابی اطلاعات هستند. بنابراین، جست‌وجوی اطلاعات با توجه به عوامل خارجی، استراتژی‌های به‌کاررفته در طراحی و سازماندهی مجموعه و عوامل شخصی و زمان انجام خواهد شد.

۵. بحث

الگـوی اغلـب جسـت‌وجوهای انجام‌شـده در بیشـتر مـوارد به‌صـورت «کلـی بـه جزئـی» و «تخصصـی» اسـت. شـرکت‌کنندگان جست‌وجویشـان را بـا اطلاعـات کلـی ماننـد معرفـی و بررسـی حقایـق شـروع نمـوده و سـپس، هـر یـک بـر روی جنبه‌هـای خاصـی از نتایـج تمرکـز نمودنـد. در بعضـی مـوارد، کاربـران در حیـن جسـت‌وجو ایده‌هـای جدیـدی به‌دسـت آوردنـد. نتایـج به‌دسـت‌آمده در فراینـد جسـت‌وجو بـرای کاربـران بـه فرمول‌بنـدی مجـدد پرسـش‌های ایـن افـراد منجـر گردیـد.

بسـیاری از شـرکت‌کنندگان جسـت‌وجوی خـود را هنگامـی کـه صفحـه‌ای شـامل منابـع فـراوان بـود، متوقـف نمـوده و به‌طـور معمـول، تصمیـم گرفتنـد کـه در زمانـی دیگـر از اطلاعـات اسـتفاده کننـد. بـرای مثـال، یکـی از کاربـران عنـوان نمـود کـه مـدرک به‌دسـت‌آمده در جسـت‌وجو مدرکـی اسـت کـه از آن در زمـان دیگـر پرینـت خواهـد گرفـت و یـا ایـن امـکان وجـود داشـت کـه مطلبـی را بـرای مطالعـه در آینـده ذخیـره نماینـد، زیـرا دسـتیابی بـه اطلاعـات بـرای آن‌هـا محدودیـت زمانـی و فوریـت نداشـت. بنابرایـن، ایـن گـروه از کاربـران تلاشـی بـرای به‌دسـت آوردن ایـدهٔ دقیقـی در هنـگام جسـت‌وجو نداشـتند و اطلاعـات آن‌هـا در داده‌هـای مـورد اسـتفاده در نظـر گرفتـه نشـد. بـرای مثـال، وقتـی یکـی از کاربـران یـک صفحـه بـا چندیـن منبـع مرتبـط یافـت، بـر روی هیچ‌کـدام کلیـک نکـرد، بلکـه از نظـر وی ایـن صفحـه شـروعی بـرای یافتـن اطلاعـات بیشـتر بـود. پاسـخ بازیابی‌شـده بـا توجـه بـه تجربیـات آن‌هـا، اطلاعـات معتبـری بـرای ادامـهٔ جسـت‌وجو بـود.

جسـت‌وجوگر دیگـر ابتـدا عبـارت «حقـوق بشـر + ایـران» را در موتـور جسـت‌وجو تایـپ نمـود و سـپس، صفحـهٔ نتایـج جسـت‌وجو را در زمـان کمـی ملاحظـه نمـود. تـلاش بعـدی وی بـا وارد نمـودن پرسـش‌های مشـابه انجـام شـد. او سـپس، در عبـارت جسـت‌وجو «دهـه اخیـر» را اضافـه نمـود. در تـلاش آخـر، وی بـه موتـور جسـت‌وجو برگشـت و ترکیـب پرسـش مشـابهی را امتحـان نمـود.

بـا توجـه بـه نتایـج به‌دسـت‌آمده، پیشـینهٔ کلـی کاربـران، دانـش موضوعـی، محـدودهٔ زمـان و ابزارهـای در اختیـار آن‌هـا پاسـخ مـورد نظـر کاربـران را تحـت تأثیـر قـرار می‌دهـد.

۶. نتیجه‌گیری

استخراج قوانین ناشناخته با استفاده از شبکهٔ عصبی به تصمیم‌گیری بهینه و اطلاعات‌محور منجر می‌شود. تاکنون، در سازمان‌ها و کتابخانه‌ها بیشتر از تصمیم‌گیری بر مبنای تجربه استفاده شده است. در دنیای امروز که اطلاعات در آن قدرت اصلی محسوب می‌شود، تلاش بر این است که تصمیم‌گیری بر پایهٔ اطلاعات‌محوری جایگزین تصمیم‌گیری بر پایهٔ تجربه‌گرایی گردد. این است که مدیران کتابخانه‌ها و مراکز اطلاع‌رسانی وظیفه دارند که با استفاده از تصمیمات اطلاعات‌محور، نیازهای اطلاعاتی و رفتار اطلاع‌یابی مراجعه‌کنندگان خود را شناسایی و پیش‌بینی نموده و در جهت تأمین به‌جا و به‌موقع آن‌ها گام بردارند و هدف اصلی، که جلب رضایت کاربران است، را برآورده نمایند.

با توجه به تجزیه و تحلیل یافته‌ها و نتایج به‌دست آمده پیشنهاد می‌شود رفتار اطلاعاتی با به‌کارگیری پارامترهای سنجیده‌شده در این مطالعه و با استفاده از شبکهٔ عصبی، جهت شناخت دقیق‌تر مهارت‌ها، موانع، اهداف و انگیزهٔ اطلاع‌یابی و تعیین و پیش‌بینی منابع و خدمات اطلاعاتی و راه‌های دسترسی به اطلاعات انجام پذیرد و نتایج به‌دست‌آمده با نتایج پژوهش حاضر مقایسه شود.

با توجه به نتایج به‌دست‌آمده به نظر می‌رسد که برای افزایش سطح مهارت اطلاع‌یابی کاربران فعلی و آینده، دورهٔ آموزشی تخصصی بازیابی اطلاعات برای همهٔ طبقات کاربران برگزار گردد. همچنین، در راستای افزایش سطح دسترسی کاربران پیشنهاد می‌شود تمهیداتی جهت تهیه، افزایش، به‌روز رسانی و غنای منابع اطلاعاتی الکترونیکی و آموزش استفاده از این منابع برای تمام کاربران با مرتبه‌های علمی مختلف اندیشیده شود.

در پژوهش حاضر پارامتر تعیین‌کننده F1، دقت P و R به کمک نتایج انتخاب‌شده توسط الگوریتم برابر ۷۳/۷، ۷۵ و ۷۲/۵ محاسبه شده و با نتایج مطالعات اخیر مقایسه شده است. بررسی‌های انجام‌شده نشان می‌دهد که ایدهٔ پیشنهادی در پیشبرد بازیابی اطلاعات به‌کمک شبکه‌های عصبی به‌خوبی نتیجه می‌دهد. به‌کارگیری پارامترهای عنوان‌شده در این مطالعه می‌تواند در تکمیل شبیه‌سازی ارائه‌شده با نتایج بهتر کمک نماید.

فهرست منابع

بدر، عاطفه، صدیقه محمداسماعیل، و حنیف حیدری. ۱۳۹۶. استفاده از تکنیک داده‌کاوی جهت دسته‌بندی کاربران هدف کتابخانه مرکزی دانشگاه صنعتی اصفهان (مطالعهٔ انگیزه‌ها و رفتارهای اطلاع‌یابی آنان). *پژوهشنامه پردازش و مدیریت اطلاعات* ۳۳ (۱): ۲۷۱-۲۹۴.

راد، ایرج. ۱۳۸۸. رفتار اطلاع‌یابی دانشجویان تحصیلات تکمیلی دانشگاه آزاد اسلامی در استفاده از شبکه جهانی وب. *کتابداری و اطلاع‌رسانی* ۱۲(۳ (مسلسل ۴۷)): ۱۴۱-۱۶۸.

محمداسماعیل، صدیقه، و جمیله نعیمی. ۱۳۹۵. بررسی تطبیقی نیازهای اطلاعاتی و رفتار اطلاع‌یابی دانش‌پژوهان حوزه و دانشگاه استان خراسان رضوی با رویکرد شبکه عصبی. *تعامل انسان و اطلاعات* ۳ (۴): ۶۰-۷۳.

نعیمی، جمیله، صدیقه محمداسماعیل، و حنیف حیدری. ۱۳۹۷. تعیین نیازهای اطلاعاتی و رفتار اطلاع‌یابی دانش‌پژوهان حوزه علمیه خراسان رضوی با رویکرد شبکه عصبی. *علوم و فنون مدیریت اطلاعات* ۴ (۱): ۹۱-۱۱۸.

یاری، شیوا، و حمید احمدی. ۱۳۹۳. مروری بر متون رفتار اطلاع‌یابی در ایران. *پژوهشنامه پردازش و مدیریت اطلاعات* ۳۰ (۱): ۱۷۳-۱۸۷.

References

Abu Abbas, Osama. .2008. Comparisons between data clustering algorithms. *The International Arab Journal of Information Technology* 5 (3): 320-325.

Al-mumen, N, & A. Morris. 2012. Modelling information-seeking Behaviour of graduate students at Kuwait University. Retrieved from www.Emeraldinsight.com (accessed Jan 20, 2016).

Badr, A., S. Mohamad Esmaeeli, & H. Heidari. 2018. Applying Data Mining Technique in order to categorize the Target Users of the Central Library of Isfahan University of Technology (Studying the Motives and Information Seeking Behaviors of Them). *Information Sciences & Technology* 33 (1): 275-294. Magiran.com/p177922

Billsus, D., and M. J. Pazzani. 1998. Learning collaborative information filters. *In Proceedings of the Fifteenth International Conference on Machine Learning (ICML)* Vol. 98, pp. 46-54. USA.

Biranvand, Ali 1392. Investigating the information-seeking behaviors of faculty members of Payame Noor University of Jahrom. *Information Management and Knowledge* 1 (1): 101-110.

Bookstein, A. 1983. Information retrieval: A sequential learning process. *Journal of American Society (ASIS)* 34 (5): 331–342.

Borlund, P. 2010. The Concept of Relevance in IR. *Journal of the American Society for Information Science and Technology* 54 (10): 913-925.

Buresova, I, & M. Simikova. 2012. Information Behaviour of Gifted Children – The Qualitative Study. *Procedia-social and behavioral Science*s 8 (69): 242-246.

Carbonell, J., and J. Goldstein. 1998. The use of MMR, diversity-based reranking for reordering documents and producing summaries. In Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval (pp. 335-336). Melbourne Australia August 24 - 28, 1998.

Chen, H., and D.R. Karger. 2006. Less is more: Probabilistic models for retrieving fewer relevant documents. *ACM SIGIR* 429–436. The 29th Annual International SIGIR Conference. Seattle Washington USA August 6 - 11, 2006429–436.

Cohn, D. A. 1996. Active learning with statistical models. *Journal of artificial intelligence research* 4: 129-145.

De Diego, A. & V. Eske. 2019. Öffentliches Wasserrecht im Rahmen von Zivilrechtsstreitigkeiten. Zeitschrift für Deutsches und Europäisches Wasser-, *Abwasser- und Bodenschutzrecht* 8 (4): 234-240.

Fuhr, N. 2008. A probability ranking principle for interactive information retrieval. *Information Retrieval* 11 (3): 251-265.

Hauff, C. 2010. Predicting the effectiveness of queries and retrieval systems. SIGIR (Special Intrest Group on Information Retireval) Forum (Vol. 44, No. 1, p. 88). Germany.

Hayati, Zahir. 1393. Segmenting public library clients based on their needs using the network: (3) Artificial neural, hierarchical analysis and Kano model. *Information Research and Public Libraries* 20: 534-513

Heidelberg Shani, G., and A. Gunawardana. 2011. Evaluating recommendation systems. *In Recommender systems handbook* (pp. 257-297). Boston, MA: Springer.

Jahrer, M., A. Töscher, and R. Legenstein. 2010. Combining predictions for accurate recommender systems. In Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining, KDD "10, 693–702, New York, NY, USA. ACM.

Jarvelin, K. 1986. Information Processing and Management: an *International Journal* 22 (6): 523-548.

Järvelin, K., and P. Ingwersen. 2004. Information Seeking Research Needs Extension towards Tasks and Technology. *Information Research* 10 (1): 212.

Jones, R., and F. Diaz. 2007. Temporal profiles of queries. *ACM Trans. Information Systems* 25 (3): 14.

Joshi, P. A, S. M. Nikose. 2013. Information seeking Behaviours of Users: A Case Study of Private Higher Technical Libraries in Chandrapur District. Retrived from http:// Eprints.rclis.org/3794. (accessed Jan. 20, 2016).

Khoshbaf, Mehdi. 1392. Investigating the information retrieval behavior of virtual graduate students of Imam Reza University based on Ellis model. Master Thesis. Imam Reza University of Mashhad.

Kohrs, A., & B.Merialdo. 1999. Improving collaborative filtering with multimedia indexing techniques to create user-adapting web sites. In *Proceedings of the seventh ACM international conference on Multimedia* (*Part 1*) (pp. 27-36). New York, US.

Lavrenko, V., and W. B. Croft. 2001. Relevance-based language models. *ACM SIGIR* (*Special Interest Group on Information Retrieval*) (2): 120–127.

Liu, X., & W. B. Croft. 2004. Cluster-based retrieval using language models. In SIGIR '04: *Proceedings of the* 27th *annual international ACM SIGIR conference on research and development in information retrieval*. 186–193. New York, NY: ACM.

Lu, H., M. Zhang & S. Ma. 2018. Between Clicks and Satisfaction: Study on Multi-Phase User Preferences and Satisfaction for Online News Reading. *In The* 41*st International ACM SIGIR Conference on Research & Development in Information Retrieval* (pp. 435-444). ACM. New York, US.

Mohamad Esmaeeli, S., & J.Naeimi. 2017. Investigation through and Clustering the Information Needs and Information Seeking Behavior of Seminary and University Students of Khorasan-e- Razavi with Neural Network Analysis, *Human Information Interaction* 3 (4): 1. Magiran.com/p1517874

Naeimi, J., & S. Mohamad Esmaeeli. 2016. The Assessment of Information- Seeking Behavior of Medical Sciences University Researchers of Khorasan Razavi using Neural Network Approach, *Library and Information Research Journal* 6 (2): 80-96. Magiran.com/p1625373

_____ & H. Heidari. 2018. The Assessment of Information- Seeking Behavior of Khorasan Razavi Seminary Students with Neural Network Approach, *Journal of Sciences and Techniques of Information Management* 4 (1): 91-118. Magiran.com/p1834056

Paramita, M. L., M. Sanderson, & P. Clough. 2009. Diversity in photo retrieval: overview of the ImageCLEFPhoto task 2009. In *Workshop of the Cross-Language Evaluation Forum for European Languages* (pp. 45-59). Springer, Berlin, Heidelberg.

Qasemokhani, Sakineh. 1393. Investigating the information seeking behavior of the elderly in public libraries in Khoy. Master Thesis, Tehran Teacher Training University.

Rad, I., 2009. Information seeking behavior of Islamic Azad University graduate students in using the World Wide Web, *Journal of Library and Information Science* 12 (3): 3.

Radlinski, F., and S. Dumais. 2006. Improving personalized web search using result diversification. In Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval (pp. 691-692). Washington, USA.

Robertson, S. E., and K. S. Jones. 1976. Relevance weighting of search terms. *Journal of the American Society of Information Science* 27 (3): 129–146.

Roy, N. 2001. Toward optimal active learning through Montecarlo estimation of error reduction. *International Conference on Machine Learning* (*ICML*). Williamstow, USA. pp. 441-448.

Salton, G., and C. Buckley. 1990. Improving retrieval performance by relevance feedback. *Journal of Information Science* 41 (4): 288–297.

Sanderson, M., J. Tang, T. Arni, & P. Clough. 2009. What else is there? search diversity examined. In *European Conference on Information Retrieval* (pp. 562-569). Springer, Berlin, Heidelberg.

Shani, G., and A. Gunawardana. 2011. Evaluating recommendation systems. *In Recommender systems handbook* (pp. 257-297). Boston, MA: Springer.

Tang, J., and M. Sanderson. 2010. *Evaluation and user preference study on spatial diversity*. In *European Conference on Information Retrieval* (pp. 179-190). Springer, Berlin, Heidelberg.

Tao, T. and C. Zhai. 2006. Regularized estimation of mixture models for robust pseudo-relevance feedback. In *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 162-169). USA.

Ticineto Clough, P. 2009. The new empiricism: Affect and sociological method. *European Journal of Social Theory* 12 (1): 43-61.

Yadav, S., S. Ramesh, S. Saha, and A. Ekbal. 2021. Relation extraction from biomedical and clinical text: Unified multitask learning framework. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*. doi: 10.1109/TCBB.2020.3020016.

Yan, X., J. Guo, S. Liu, X. Cheng, and Y. Wang. 2013. Learning topics in short texts by non-negative matrix factorization on term correlation matrix. In *Proceedings of the* 2013 *SIAM International Conference on Data Mining* (pp. 749-757). Society for Industrial and Applied Mathematics. Germany.

Yao, C., J. Bu, C. Wu, and G. Chen. 2013. Semi-supervised spectral hashing for fast similarity search. *Neuro computing* 101 (1): 52-58.

Yari, S., & H. Ahmadi. 2014. A Review on Information Seeking Behaviour Literature in Iran, *Information Sciences & Technology* 30 (1): 173-197. Magiran.com/p1338101

Zhang, Y., V. Zhong, V. Chen, A. Angeli, and C. Manning. 2017. Position-aware Attention and Supervised Data Improve Slot Filling. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, 35–45. USA.

Zhang, Y., H. Lin, Z. Yang, J. Wang, S. Zhang, & L. Yang. 2018a. A hybrid model based on neural networks for biomedical relation extractionJournal of Biomedical Informatics, 81: 83–92.

Zhang, Y., S. Liu, J. Tan, G. Jiang, & Q. Zhu. 2018b. Effects of risks on the performance of business process outsourcing projects: The moderating roles of knowledge management capabilities. *International Journal of Project Management* 36 (4): 627–639.

Zhao, J, & Y. Yun. 2009. A proximity language model for information retrieval. In SIGIR '09: *Proceedings of the* 32nd *international ACM SIGIR conference on research and development in information retrieval*. 291-298. New York, NY: ACM.

Zheng, W., H. Lin, L. Luo, Z. Zhao, Z. Li, Y. Zhang, & J. Wang. 2017. An attention-based effective neural model for drug-drug interactions extraction. *BMC Bioinformatics* 18 (1): 445-451.

Zhou, N., J. Du, X. Yao, W. Cui, Z. Xue, & M. Liang. 2019. A content search method for security topics in microblog based on deep reinforcement learning. *World Wide Web* 23 (1): 75-101

فتانه وهابی

دانشجوی دکتری کتابداری و اطلاع‌رسانی از پژوهشگاه علوم و فناوری اطلاعات ایران است.

بازخورد رفتاری، بازیابی اطلاعات، شبکه عصبی، کاربر، موتور جست‌وجو از جمله علایق پژوهشی وی است.

سعید اسدی

دارای مدرک تحصیلی دکتری فناوری اطلاعات از دانشگاه کوئینزلند است. ایشان هم‌اکنون دانشیار گروه علم اطلاعات و دانش‌شناسی دانشگاه شاهد است.

علم اطلاعات، ارتباطات علمی، کتابشناسی از جمله علایق پژوهشی وی است.

سهیل گنجه‌فر

دارای مدرک تحصیلی دکتری برق در رشتهٔ برنامه‌ریزی درسی از دانشگاه تربیت مدرس است. ایشان هم‌اکنون استاد گروه مهندسی برق دانشگاه علم و صنعت است.

شبکه‌های عصبی، سیستم‌های هوشمند و کنترل سیستم‌های حرکت از راه دور از جمله علایق پژوهشی وی است.