سامانهٔ پیشنهادگر ترکیبی، مبتنی بر هستان‌شناسی برای مقابله با مشکل شروع سرد

پیام بحرانی[۱]، بهروز مینایی بیدگلی[۲]، حمید پروین*[۳و۴]، میترا میرزارضایی[۱] و احمد کشاورز[۵]

[۱]گروه مهندسی کامپیوتر، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران

[۲]دانشکده مهندسی کامپیوتر، دانشگاه علم و صنعت ایران، تهران، ایران

[۳]گروه مهندسی کامپیوتر، واحد نورآباد ممسنی، دانشگاه آزاد اسلامی، نورآباد ممسنی فارس، ایران

[۴]باشگاه پژوهشگران جوان و نخبگان، واحد یاسوج، دانشگاه آزاد اسلامی، یاسوج، ایران

[۵]گروه مهندسی برق، دانشکده مهندسی سیستم‌های هوشمند و علوم داده، دانشگاه خلیج فارس، بوشهر، ایران

چکیده

انتظار می‌رود سامانه‌های پیشنهادگر (RS) قلم‌های دقیق را به مصرف‌کنندگان پیشنهاد دهند. شروع سرد مهم‌ترین چالش در RSها است. RSهای ترکیبی اخیر، دو مدل پالایش محتوا پایه (ConF) و پالایش مشارکتی (ColF) را با هم ترکیب می‌کنند. در این پژوهش، یک RS ترکیبی مبتنی بر هستان‌شناسی معرفی می‌شود که در آن هستان‌شناسی در بخش ConF به‌کار رفته است، این در حالی است که ساختار هستان‌شناسی توسط بخش ColF بهبود داده می‌شود. در این مقاله، رویکرد ترکیبی جدیدی مبتنی بر ترکیب شباهت جمعیت‌شناختی و شباهت کسینوسی بین کاربران به‌منظور حل مشکل شروع سرد از نوع کاربر جدید، ارائه شده است. همچنین، رویکرد جدیدی مبتنی بر ترکیب شباهت هستان‌شناسی و شباهت کسینوسی بین اقلام به‌منظور حل مسأله شروع سرد از نوع قلم جدید، ارائه شده است. ایده اصلی روش پیشنهادی، گسترش پروفایل‌های کاربر/قلم بر اساس سازوکارهای مختلف برای ایجاد پروفایل با عملکرد بالاتر برای کاربران/قلم‌ها است. روش پیشنهادی در یک مجموعه‌داده واقعی ارزیابی شده است و آزمایش‌ها نشان می‌دهند که روش پیشنهادی در مقایسه با روش‌های پیشرفته RS، به‌خصوص در مواجهه با مسأله شروع سرد، عملکرد بهتری دارد.

واژگان کلیدی: سامانه پیشنهادگر، هستان‌شناسی، توسعه پروفایل، سامانه پیشنهادگر ترکیبی.

An Ontological Hybrid Recommender System for Dealing with Cold Start Problem

Payam Bahrani[1], Behrouz Minaei-Bidgoli[2], Hamid Parvin[3,4*], Mitra Mirzarezaee[1] & Ahmad Keshavarz[5]

[1]Department of Computer Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran

[2]School of Computer Engineering, Iran University of Science and Technology, Tehran, Iran

[3]Department of Computer Engineering, Nourabad Mamasani Branch, Islamic Azad University,
Nourabad Mamasani, Iran

[4]Young Researchers and Elite Club, Yasooj Branch, Islamic Azad University, Yasooj, Iran

[5]Department of Electrical Engineering, Faculty of Intelligent Systems Engineering and Data Science, Persian Gulf University, Bushehr, 75169, Iran

Abstract

Recommender systems that predict user ratings for a set of items are known as subset of information filtration systems. They help users find their favorite items from thousands of available items.

* نویسندهٔ عهده‌دار مکاتبات

* Corresponding author

• تاریخ ارسال مقاله: ۱۳۹۹/۹/۲۶ • تاریخ پذیرش: ۱۴۰۰/۳/۱ • تاریخ انتشار: ۱۴۰۱/۰۳/۳۱ • نوع مطالعه: پژوهشی

[Downloaded from jsdp.rcisp.ac.ir on 2026-09-05]

[DOI: 10.52547/jsdp.19.1.1]

One of the most important and challenging problems that recommendation systems suffer from is the problem of dispersion. This means that due to the scatter of data in the system, they are not able to find popular items with the desired reliability and accuracy. This is especially true when there are a large number of items and users in the system and the filled ratings are low. Another challenging problem that these systems suffer from is their scalability. One of the major problems with these systems is the cold start. This problem occurs due to the small number of items rated by the user, i.e. the scatter of users. This problem is divided into two categories: new user and new item. The main focus of this article is on the problem of the new user type. This problem occurs when a new user has just logged in and has not rated any item yet, or when the user has already logged in but has been less active in rating. The goal is to address these three challenges.

In this study, an ontology-based hybrid recommender system is introduced in which ontology is used in the content-based filtering section, while the ontology structure is improved by the collaborative filtering section. In this paper, a new hybrid approach based on combining demographic similarity and cosine similarity between users is presented in order to solve the cold start problem of the new user type. Also, a new approach based on combining ontological similarity and cosine similarity between items is proposed to solve the cold start problem of the new item type. The main idea of the proposed method is to extend users'/items' profiles based on different mechanisms to create higher-performance profiles for users/items.

The proposed method is evaluated in a real data set, and experiments show that the proposed method performs better than the advanced recommender system methods, especially in the case of cold start.

Keywords: Recommender System; Ontology; Profile Expansion; Hybrid Recommender System.

۱- مقدمه

سامانه‌های پیشنهادگر[1] که رتبه‌بندی کاربران را برای مجموعه‌ای از قلم‌ها (کتاب‌ها، قلم‌ها، اخبار، آهنگ‌ها و غیره) پیش‌بینی می‌کنند به‌عنوان زیرمجموعه سامانه‌های پالایش اطلاعات (IFS) شناخته می‌شوند [1,2].

RS به کاربران کمک می‌کند تا قلم‌های مورد علاقه خود را از بین هزاران قلم موجود پیدا کنند. درحقیقت، دلیل موفقیت این سامانه‌ها در تارنماهای تجاری، شخصی سازی پیشنهادهای ارائه‌شده به کاربران است [3].

پالایش‌کردن مشارکتی[2] یکی از متداول‌ترین روش ها برای اجرای RS است [4]. مقادیر رتبه‌بندی که توسط کاربر به اقلام مختلف اختصاص داده شده است، بردار امتیاز اقلام[3] (IRV) نامیده می‌شود. هر قلمی که توسط کاربر رتبه‌بندی نشده باشد، در عنصر مربوط به آن IRV مقدار Nullدارد. به این ترتیب، کاربران مشابه (آن دسته از کاربران با IRV یکسان با IRV کاربر هدف) به‌عنوان همسایگان کاربر هدف شناخته می‌شوند. کاربر هدف به‌عنوان کاربری شناخته می‌شود که به‌وسیله سامانه، به او پیشنهاد داده می‌شود؛ بنابراین، یک ColF RS فهرستی از قلم‌های مورد علاقه را در میان قلم‌های رتبه‌بندی‌شده توسط همسایگان که با علامت اختصاری N نشان داده می‌شود، به کاربر هدف ارائه می‌دهد. RSهای ColF به رویکردهای مبتنی بر حافظه و مبتنی بر مدل تقسیم می‌شوند [5]. ColF RS ها به‌طورمستقیم از ماتریس قلم-کاربر برای محاسبه پیش‌بینی‌ها و تولید پیشنهادهایی برای کاربر هدف استفاده می‌کنند [6]. در این روش‌ها از روش‌های آماری مانند شباهت پیرسون [7] و شباهت کسینوس [8] به‌منظور محاسبه شباهت بین کاربران و یافتن مجموعه همسایه کاربر هدف استفاده می‌شود [9]. در ColF مبتنی بر مدل، راهکار استفاده از سازوکارهای یادگیری ماشین، مانند خوشه‌بندی [10]، طبقه‌بندی بیزی [11] و الگوریتم ژنتیک [12]، برای ایجاد مدلی از داده‌های آموزشی در ماتریس قلم-کاربر است؛ علاوه‌براین، در مرحله آزمون، می‌توان از این مدل آموزش برای پیش‌بینی رتبه کاربر استفاده کرد.

یکی از مشکلات مهم و چالش‌برانگیز که سامانه‌های RS از آن رنج می‌برند، مشکل پراکندگی داده است. به این معنی، که به‌دلیل پراکنده‌بودن داده‌ها در سامانه، آنها قادر به یافتن موارد محبوب با اطمینان و دقت مطلوب نیستند [13]. درواقع، این امر به‌ویژه در شرایطی که تعداد زیادی قلم و کاربر در سامانه وجود دارد و نسبت امتیاز آنها توسط کاربران پایین است، پدیدار می‌شود؛ بنابراین، یافتن نزدیک‌ترین همسایگان معنی‌دار کاربر غیرممکن است. مشکل چالش‌برانگیز دیگر مربوط به RS امکان حضور کاربرانی با علایق غیرعادی است. کاربرانی که علایق غیرعادی دارند می‌توانند RS را گمراه و قلم‌های نامناسب را به کاربران دیگر پیشنهاد کنند.

[1] Recommender Systems
[2] ColFting Filtering
[3] Items' Rating Vector

[Downloaded from jsdp.rcisp.ac.ir on 2026-09-05]

[DOI: 10.52547/jsdp.19.1.1]

ماتریس قلم-کاربر یک ماتریس از اندازه $n \times m$ است که n نشان‌دهنده تعداد کاربران و m نشان‌دهنده تعداد قلم‌ها است. مدخل (i,j) این ماتریس، رتبه‌ای است که کاربر i-ام به قلم j-ام اختصاص می‌دهد. تعداد قلم هایی که توسط کاربر رتبه‌بندی شده به‌طورتقریبی کمتر از یک درصد از کل قلم‌های موجود در سامانه است؛ در نتیجه، ماتریس قلم-کاربر، یک ماتریس تُنُک با کمتر از یک درصد داده‌ها است. این مشکل تا حدودی در پالایش‌های ترکیبی RS قابل حل است. نویسندگان مرجع [14] مدلی را پیشنهاد کردند که ورودی‌های با مقدار $null$ در ماتریس قلم-کاربر به‌عنوان یک مرحله پیش‌پردازش با استفاده از یک شبکه عصبی پرسپترون چندلایه پر شود. آنها از روش تشابه قلمی استفاده کرده‌اند [14]. یک راه‌حل برای مقابله با مشکل پراکندگی، رویکرد کاهش ابعاد است [15]. به‌عنوان مثال، یک SVD[1] همراه با PSO[2] برای مواجهه با مسأله پراکندگی استفاده شده است. گفتنی است که SVD سربار بالایی دارد؛ بنابراین، برخلاف نتایج مؤثر، نمی‌توان از SVD در RSهای برخط استفاده کرد. برای سرعت‌بخشیدن به زمان اجرای SVD از PSO استفاده می‌شود [15].

مشکل چالش‌برانگیز دیگری که سامانه‌های RS از آن رنج می‌برند، مربوط به مقیاس‌پذیری آنها است. وقتی حدود صدها هزار کاربر وجود دارد، RS ColFهای مبتنی بر کاربر، کارآمد و مقیاس‌پذیر هستند؛ اما در سامانه‌های واقعی، بیش از میلیون‌ها کاربر وجود دارد؛ بنابراین RS ColFهای مبتنی بر کاربر، دیگر مقیاس‌پذیر و مفید نیستند. RSهای مبتنی بر قلم‌ها را می‌توان به‌عنوان یک گزینه جایگزین در نظر گرفت.

یکی از مهم‌ترین مشکلات سامانه‌های RS، شروع سرد[3] (CS) است [16]. مشکل CS به‌دلیل تعداد کم قلم های رتبه‌بندی‌شده توسط کاربر، یعنی پراکندگی IRV کاربران رخ می‌دهد. مشکلات CS به دو دسته تقسیم می شوند: کاربر جدید[4] (NU) و قلم جدید[5] (NI) [17]. تمرکز اصلی در این مقاله بر روی مشکل CS از نوع NU است. مشکلNU زمانی رخ می‌دهد که کاربر جدید به‌تازگی وارد سامانه شده و تاکنون هیچ قلمی را رتبه‌بندی نکرده یا زمانی‌که کاربر از قبل در سامانه حضور داشته باشد، اما فعالیت کمتری داشته و در نتیجه تعداد کمی از قلم‌ها را تاکنون رتبه‌بندی کرده است.

تاکنون کارهای زیادی برای حل مشکل CS در RS انجام شده است. راه‌حل‌های مسأله CS به‌طور کلی به دو گروه تقسیم می‌شوند. گروهی از راه‌حل‌ها، روش‌هایی هستند که از طریق بهبود RSهای سنتی با آن مقابله می‌کنند تا آنها را برای مقابله با مشکل CS آماده کند [18]. گروه دیگری از راه‌حل‌ها روش‌هایی هستند که با استفاده از RSهای ترکیبی با آن روبه‌رو می‌شوند [19]. کاربران جدید آنهایی هستند که در ابتدا تعداد کمی از قلم‌ها را رتبه‌بندی می کنند یا کاربرانی که هنوز هیچ قلمی را رتبه‌بندی نکرده‌اند. گفتنی است که آنها تقریباً یک نمایه[6] خالی (یا یک نمایه بسیار کوچک) دارند. بنابراین، RSها قادر به تشخیص ترجیحات خود نیستند و نمی‌توانند قلم‌های مناسب و مورد علاقه خود را به آنها ارائه دهند؛ از این‌رو، بیشتر پژوهش‌گران از RS های ترکیبی برای حل مسأله NUاستفاده کرده‌اند [20]. RSهای ترکیبی به‌طورمعمول ترکیبی از ColF با منابع داده اضافی هستند.

در این مقاله روش جدیدی ارائه می‌شود که، برای حل مشکل CS در RS پیشنهاد شده است. در سامانه پیشنهادی، فرآیند پیشنهاد شامل دو مرحله است. در مرحله نخست، مشخصات کاربر هدف و مشخصات قلم هدف توسعه می‌یابد. برای این منظور، از ترکیب شباهت کسینوس و شباهت جمعیتی به‌عنوان شباهت ترکیبی نهایی برای انتخاب نزدیکترین همسایگان کاربر هدف برای گسترش مشخصات آن استفاده می‌شود؛ بنابراین، نوآوری اساسی در روش پیشنهادی این است که، علاوه‌بر اطلاعات مربوط به رتبه‌بندی که به‌وسیلهٔ کاربران به قلم‌های مختلف داده می‌شود، از اطلاعات جمعیت‌شناختی آنها نیز به‌عنوان اطلاعات اضافی برای حل مشکل CS استفاده می‌شود. پس از توسعه پروفایل‌های کاربر و پروفایل‌های قلم، اگر کاربر هدف NUباشد، قلم‌های مورد نظر بر اساس مشخصات گسترش‌یافته برای کاربر هدف رتبه‌بندی می شوند. در غیراین‌صورت اگر قلم مورد نظر NI باشد، قلم هدف بر اساس مشخصات توسعه‌یافته قلم رتبه‌بندی می شود؛ به بیان دیگر، سامانهٔ پیشنهادگر یک پروفایل متشکل از اطلاعات گذشته درخصوص تمایلات کاربران تشکیل داده و نگهداری می‌کند تا در آینده بر اساس این پروفایل، اقلام دیگر را نیز به وی پیشنهاد دهد به‌گونه‌ای که این اقلام بیشترین نزدیکی را با اقلام مورد علاقه کاربر داشته باشد.

[1] Singular Value Decomposition
[2] Particle Swarm Optimization
[3] Cold Start
[4] New User
[5] New Item
[6] Index

[Downloaded from jsdp.rcisp.ac.ir on 2026-09-05]
[DOI: 10.52547/jsdp.19.1.1]

مقاله حاضر شامل پنج بخش که بخش نخست آن مقدمه و بیان مسأله است. در بخش دوم ادبیات پژوهش و مرور منابع مختلف آورده شده است. بخش سوم روش پیشنهادی را تشریح می‌کند. بخش چهارم آزمایش‌های مختلف و نتایج تجربی را برای روش پیشنهادی و مقایسه با سایر روش‌های رقیب نشان می‌دهد. درنهایت بخش پنجم نتیجه‌گیری کلی و پیشنهادهای آینده را بیان می‌کند.

۲- ادبیات پژوهش و کارهای مرتبط

در این بخش سعی می‌شود که مروری بر ادبیات پژوهش و کارهای گذشته صورت گیرد.

تاکنون، پژوهش‌های زیادی در مورد استفاده از منابع داده اضافی مانند اطلاعات جمعیت‌شناختی [21] کاربران برای حل مشکل CS انجام شده است. به‌عنوان مثال، پیشنهاد می‌شود برای یادگیری علایق کاربران و طبقه‌بندی قلم‌ها بر اساس علایق کاربران، برچسب‌گذاری مشارکتی در یک فیلترRS مشترک استفاده شود [22]. روش دیگری برای مقابله با مشکلات یادشده که به‌طورمعمول به‌عنوان راه‌حلی برای مشکل CS در RS استفاده می‌شود، خوشه‌بندی است. در خوشه‌بندی به کارگیری قلم‌ها، کاربران و یا به‌کارگیری هم‌زمان هر دو امکان‌پذیر است. این روش‌ها می‌توانند سرعت اجرای RSها را تسریع کنند.

روش جدیدی در [23] برای حل مسأله CS، که شامل سه مرحله مختلف است، بررسی شده است. در مرحله نخست، با استفاده از الگوریتم‌های طبقه‌بندی، کاربر جدید در گروه خاصی قرار می‌گیرد. در مرحله دوم، شباهت بین کاربر جدید و سایر کاربران گروه هدف بر اساس داده‌های جمعیتی کاربران محاسبه می‌شود. در مرحله سوم، با استفاده از گروه‌های مختلف ایجادشده برای کاربران، رتبه‌بندی برای کاربر جدید پیش‌بینی می‌شود. یک روش ترکیبی در [24] برای حل مسأله CS با استفاده از محتوای کاربرمحور و اطلاعات اجتماعی ارائه شده است. ایده اصلی این روش ساخت اطلاعات محتوا بر اساس مشخصات واژگان کلیدی مرتبط با قلم‌های مختلف و استفاده از این اطلاعات برای ایجاد پیشنهاد برای کاربران جدید است. روش‌های مختلفی برای ایجاد پروفایل کاربر بر اساس اطلاعات جمعیتی توسط [25] بررسی شده است. این روش‌ها ترکیبات مختلفی را شامل می‌شوند که شامل انواع ویژگی‌های مورد استفاده، انواع نمایش ویژگی‌ها و روش‌های نمایه‌سازی کاربران است. روش دیگری مبتنی بر اطلاعات جمعیتی برای حل مشکل CS برای کاربران جدید توسط [26] ارائه شده است. در این روش، به‌جای استفاده از نرخ‌هایی که کاربران به موارد مختلف می‌دهند، از اطلاعات جمعیت‌شناختی آنها برای تولید پیشنهادها استفاده می‌شود. برای این‌منظور، چارچوبی برای ارزیابی ویژگی‌های جمعیت‌شناختی مانند سن، جنسیت و شغل ارائه شده است. در RS کاربرمحور دیگر، راه‌حلی برای مشکل CS ارائه شده است [27]. به این ترتیب، کاربران مختلف بر اساس داده‌های جمعیتی خود طبقه‌بندی می شوند و سپس براساس دسته‌های جمعیتی آنها، پیشنهادهایی تولید می‌شود؛ بنابراین، با استفاده از اطلاعات جمعیتی کاربران، مشکل CS تا حد زیادی حل شده است.

برخلاف روش‌های یادشده، سایر روش‌ها از منابع داده اضافی برای حل مسأله CS از نوع NU استفاده نکرده اند. آنها فقط سعی می‌کنند مشکل CS را با وضعیت فعلی نمایه رتبه‌بندی کاربر بهبود بخشند. یک چارچوب مبتنی بر معیار شباهت محلی کاربران و همچنین یک معیار شباهت سراسری کاربران سعی در بهبود مشکل CS دارد [28]؛ علاوه‌براین، از انواع مختلف پراکندگی داده‌ها، مانند معیار پراکندگی کلی، معیار پراکندگی خاص کاربر[1] و غیره، برای حل مشکل پراکندگی داده‌ها و مسأله CS استفاده شده است. در این روش، به‌منظور پیش‌بینی رتبه یک قلم خاص برای یک کاربر معین، از ترکیب رتبه‌بندی های ایجادشده بر اساس شباهت‌های محلی و سراسری کاربران با درنظرگرفتن توزیع یکسان ماتریس رتبه‌بندی کاربران استفاده می‌شود. نویسندگان در مرجع [31-29] بر محدودیت‌هایی از جمله معیارهای شباهت سنتی همانند شباهت کسینوسی و پیرسون تمرکز کرده‌اند، و معیار تشابه اکتشافی جدیدی با عنوان "نزدیکی-تأثیر-محبوبیت"[2] یا به اختصار (PIP) را برای مقابله با مشکل CSارائه داده‌اند. این معیار شباهت شامل سه عامل است: (۱) عامل مجاورت (نشان‌گر فاصله بین دو رتبه‌بندی)، (۲) ضریب تأثیر (نشانگر شدت نفرت یا علاقه کاربران به قلم خاص) و (۳) عامل محبوبیت (نشان‌گر فاصله بین میانگین رتبه‌بندی دو کاربر تا قلم داده‌شده از میانگین امتیاز کل کاربران به قلم داده‌شده).

[1] User-Item Specific Sparsity Measure
[2] Proximity–Impact–Popularity

[Downloaded from jsdp.rcisp.ac.ir on 2026-09-05]
[DOI: 10.52547/jsdp.19.1.1]

در فیلتر مشارکتی، RS قلم‌ها را بر اساس رتبه‌بندی سایر کاربران رتبه‌بندی و سپس قلم‌ها را با بیشینه مقادیر رتبه به کاربر فعلی پیشنهاد می‌کند؛ بنابراین، در ابتدای سامانه، چون هنوز رتبه‌بندی قلم وجود ندارد، RS برای پیش‌بینی امتیازات با کیفیت بالا برای قلم‌ها با یک مشکل چالش‌برانگیز روبه‌رو است. این مشکل شروع سرد نامیده می‌شود. به‌طور کلی، وقتی قلم جدیدی به سامانه وارد می‌شود، RS با سردی قلم روبه‌رو می‌شود و آن قلم NI است و وقتی کاربر جدیدی به سامانه معرفی می‌شود، RS با شروع سرد کاربر روبه‌رو می‌شود و آن کاربر یک NU است [32]. NU کاربرانی هستند که به‌طور کلی کمتر از پنج قلم را ارزیابی می‌کنند [33]. برای پرداختن به CS، پژوهش‌گران از دو روش استفاده می‌کنند: (الف) استفاده از اطلاعات اضافی کاربران یا مواردی برای کمک به RS و (ب) استفاده از معیارهای شباهت مناسب‌تر، تا از اطلاعات قبلی در سامانه بیشینه بهره را ببریم. نویسندگان مرجع [34] پیشنهاد کردند که از قوانین ارتباط به‌عنوان منبع اطلاعات، برای گسترش مشخصات کاربر و حل مشکل شروع سرد از نوع NU استفاده شود [34]. نویسندگان مرجع [35] مدلی را پیشنهاد کردند که در آن می‌توان از طریق علایق نامرئی[1] کاربران به مدل‌های رفتاری اساسی کاربر دست یافت. آنها یک RS را پیشنهاد می‌دهند که یاد می‌گیرد و از این اطلاعات استفاده می‌کند.

۳- روش پیشنهادی

هستان‌شناسی ریشه‌ای یونانی دارد و از دو واژهٔ Onto به معنی هستان و Logia به معنی شناختن، سرچشمه گرفته است. تعریف هستان‌شناسی در دنیای رایانه از زبان Gruber و Studer بدین صورت است که هستان‌شناسی را یک توصیف واضح و رسمی از یک مفهوم مشترک، تعریف می‌کنند. می‌توان گفت که دو واژه واضح و رسمی‌بودن به این معنی است که برای رایانه قابل درک باشد.

اجزای اصلی هستان‌شناسی شامل مفاهیم، ویژگی ها و ارتباط بین آنها است. مفاهیم را در بعضی از مراجع کلاس یا رده تعبیر کرده‌اند. رده مجموعه یا گروهی از موجودات است که در یک یا چندین صف و ویژگی با هم مشترک هستند. به‌عنوان مثال در یک دانشگاه استاد و مدیر گروه (چنانچه استاد نباشد) جزء دسته کارمندان دانشگاه قرار می‌گیرند. به هر یک از اعضای رده یک نمونه گفته می‌شود. درواقع داشتن ویژگی‌های مشترک بین نمونه‌های رده‌ها، آن نمونه را جزئی از ردهٔ دیگری می‌کند. رابطه هم که نمونه‌هایی از یک رده را به نمونه‌هایی از یک کلاس دیگر مرتبط می‌کند. (به‌عنوان مثال روابط سلسله-مراتبی بین مفاهیم) به‌عبارت دیگر، اجزای اصلی هستان-شناسی شامل مفاهیم، ویژگی‌ها و ارتباط بین آنها است. یعنی می‌توان اشیای پیرامون خود را بر اساس ویژگی‌ها و خصایص‌شان دسته‌بندی کرد تا راحت‌تر بتوان در مورد آنها قوانینی را اعمال کرد.

در این بخش، روش جدیدی برای حل مشکل شروع سرد در سامانه پیشنهادگر ارائه شده است. در سامانه پیشنهادی، فرآیند پیشنهاد شامل دو مرحله است. در مرحله نخست، مشخصات کاربر هدف و مشخصات قلم هدف گسترش می‌یابد. برای این منظور، از ترکیب شباهت کسینوس و شباهت جمعیتی به‌عنوان شباهت ترکیبی استفاده شده است. شباهت‌سنجی ترکیبی پیشنهادی، برای انتخاب نزدیک‌ترین همسایگان کاربر هدف برای توسعه مشخصات آنها استفاده می‌شود؛ بنابراین، در این روش پیشنهادی، علاوه‌بر اطلاعات مربوط به رتبه‌بندی که توسط کاربران به قلم‌های مختلف داده می‌شود، از اطلاعات پروفایل آنها نیز به‌عنوان اطلاعات اضافی برای حل مشکل شروع سرد استفاده می‌شود. پس از توسعه پروفایل‌های کاربر و پروفایل‌های قلم، در مرحله دوم، اگر کاربر هدف NU باشد، قلم‌های مورد نظر بر اساس مشخصات توسعه‌یافته برای کاربر هدف رتبه‌بندی می شوند. در غیراین‌صورت اگر قلم مورد نظر NI باشد، قلم هدف بر اساس مشخصات توسعه داده شده رتبه‌بندی می‌شود. در زیربخش‌هایی که در ادامه آورده شده‌اند، هر مرحله با جزئیات شرح داده خواهد شد. شکل (۱) طرحی از روش پیشنهادی را نشان می‌دهد. همچنین، شکل (۲) شبه‌کد روش پیشنهادی را نشان می‌دهد.

۳-۱- راه‌کار میانگین رتبه‌بندی محلی

در این مرحله، یک روش ترکیبی مبتنی بر کسینوس و شباهت‌های جمعیت‌شناختی بین کاربران برای گسترش مشخصات کاربر ایجاد شده است. از اطلاعات موجود در ماتریس رتبه‌بندی کاربر برای محاسبه شباهت کسینوس استفاده می‌شود. همچنین، از اطلاعات جمعیت‌شناختی کاربران (شامل سن، جنسیت و شغل) برای محاسبه تشابه جمعیتی در بین کاربران استفاده می‌شود. به‌منظور تجزیه‌و تحلیل مشکل شروع سرد، از روش پیشنهادی استفاده می‌شود. در این روش، برای هر زوج ارزیابی‌شده از کاربران، همه قلم‌هایی که توسط هر دو رتبه‌بندی شده‌اند؛ انتخاب، سپس از معیار تشابه کسینوس برای محاسبه شباهت بین کاربران استفاده می‌شود، که در معادله (۱) ارائه شده است.

[1] Invisible Interests

[Downloaded from jsdp.rcisp.ac.ir on 2026-09-05]

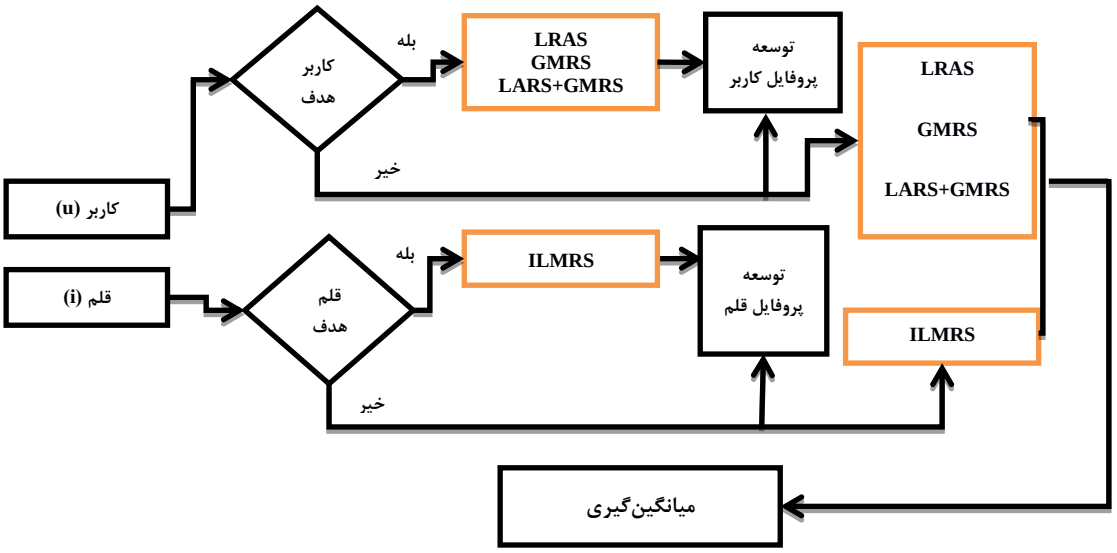
[DOI: 10.52547/jsdp.19.1.1]

$$US_{ij}^{R} = \frac{\sum_{k\in A_{ij}} R_{ik}R_{jk}}{\sqrt{\sum_{k\in A_{ij}} R_{jk}^{2}}\sqrt{\sum_{k\in A_{ij}} R_{jk}^{2}}} \qquad (۱)$$

در معادله (۱)، US_{ij}^{R} شباهت کسینوسی بین یک جفت از کاربران i-ام و j-ام است. کمینه و بیشینه مقادیر برای US_{ij}^{R} به‌ترتیب مقادیر صفر و یک است. R_{ij} بیان‌گر میزان نرخی است که کاربر i-ام به قلم j-ام می‌دهد. A_{ij} مجموعه شاخص‌های آن دسته از عناصر است که توسط کاربران i-ام و j-ام طبقه‌بندی و به‌شرح زیر تعریف می شود:

$$A_{ij} = \{k | R_{jk} \neq NaN \wedge R_{ik} \neq NaN\} \qquad (۲)$$

که NaN ناشناخته است.



(شکل-۱): شمای کلی از روش پیشنهادی

(Figure-1): The scheme of the proposed method

Input:

1. Let's $U_i.DI$ denotes demographic information of the ith user.
2. Let's R_{ij} denotes the rate value the ith user gives to the jth movie. Let's $R_{i:}$ denotes the rate values the ith user gives to the different movies; its transpose (i.e. $R_{i:}^{T}$) can be seen as a rate vector of the ith user. Let's $R_{:j}$ denotes the rate values the different users give to the jth movie.
3. Let's $I_i.L$ denotes the language of the ith movie.
4. Let's $I_i.D$ denotes the director of the ith movie.
5. Let's $I_i.W$ denotes the writer of the ith movie.
6. Let's $I_i.R$ denotes the average rate the ith item has gotten.
7. Let's $I_i.Rt$ denotes the runtime of the ith item.
8. Let's $I_i.Rd$ denotes the release data of the ith item.
9. Let's $I_i.C$ denotes the country of the ith item.
10. Let's $I_i.N$ denotes the number of rates the ith item has gotten.
11. Let's $I_i.P$ denotes the producer of the ith item.
12. Let's $I_i.G_j$ is an asymmetric Boolean variable indicating whether the ith item has the jth genre or not.
13. Let's $I_i.A_k$ denotes whether the kth famous actor is available in the actors' list of the ith movie or not. A famous actor is the one features in at least five movies. $I_i.A_k$ is an asymmetric Boolean variable.
14. Let l be expanding size of user profile or item profile

Output:

1. R matrix

Body:

1. $Z = Q = 25$
2. $N_{LARS} = N_{GMRS} = N_{GMRS} = l$
3. $US_{ij}^{D} = Sim(U_i.DI, U_j.DI)$ where US_{ij}^{D} is the users' similarity matrix defined based on users' demographic information.
4. $US_{ij}^{R} = \frac{\sum_{k\in A_{ij}} R_{ik}R_{jk}}{\sqrt{\sum_{k\in A_{ij}} R_{ik}^{2}\sum_{k\in A_{ij}} R_{jk}^{2}}}$ where US_{ij}^{R} is the users' similarity matrix defined based on users' ratings and $A_{ij} = \{k | R_{jk} \neq NaN \wedge R_{ik} \neq NaN\}$.
5. $IS_{ij}^{R} = \frac{\sum_{k\in B_{ij}} R_{ki}R_{kj}}{\sqrt{\sum_{k\in B_{ij}} R_{ki}^{2}\sum_{k\in B_{ij}} R_{kj}^{2}}}$ where IS_{ij}^{R} is the items similarity matrix defined based on users' ratings and $B_{ij} = \{k | R_{ki} \neq NaN \wedge R_{kj} \neq NaN\}$.
6. Let's IS_{ij}^{O} be the items' similarity matrix defined based on items' ontology information (features defined in lines 3-13 of input section).
7. Define $US_{ij} = \theta_u(|A_{ij}|) \times US_{ij}^{D} + \left(1 - \theta_u(|A_{ij}|)\right) \times US_{ij}^{R}$.
8. Define $IS_{ij} = \theta_i(|B_{ij}|) \times IS_{ij}^{O} + \left(1 - \theta_i(|B_{ij}|)\right) \times IS_{ij}^{R}$ is the items' similarity matrix defined based on items' ontology information.
9. Use LARS or GMRS or LARS+GMRS to determine R_{ij} and then expand user profile of the target user if the target user is NU

[Downloaded from jsdp.rcisp.ac.ir on 2026-09-05]

[DOI: 10.52547/jsdp.19.1.1]

10. Use ILMRS to determine R_{ij} and then expand the target item profile if the target item is a NI
11. If jth item is a cold item or NI
12. Use ILMRS to determine R_{ij} on the expanded profiles of items
13. Elseif ith user is a cold user or NU
14. Estimate R_{ij} using equation 6 on the expanded profiles of users either by GMRS or LARS or LARS+GMRS
15. Else
16. Estimate R_{ij} using ILMRS on the non-expanded profiles of items and with GMRS or LARS or LARS+GMRS on the non-

expanded profiles of users and then report average of the values obtained by two methods

(شکل-۲): شبه‌کد روش پیشنهادی

(Figure-2): The general framework of the proposed method

شباهت کاربران نیز بر اساس داده‌های جمعیت‌شناختی (به‌عنوان مثال سن، جنسیت و شغل) محاسبه می‌شود. شباهت‌های جمعیتی در معادله (۳) بین دو کاربر محاسبه می‌شود. در این رابطه، اگر مقدار یک صفت خاص برای دو کاربر یکسان باشد، مقدار شباهت این صفت خاص برای هر دو کاربر برابر است؛ درغیراین‌صورت، برای آنها برابر با صفر است. به‌عنوان مثال، اگر هر دو کاربر جنسیت یکسانی داشته باشند، مقدار یک، درغیراین‌صورت صفر است:

$$US_{ij}^{D} = Sim\left(U_i.DI, U_j.DI\right) = \frac{\sum_{k=1}^{d} w_k \times \pi\left(U_i.f_k, U_j.f_k\right)}{\sum_{k=1}^{d} \pi\left(U_i.f_k, U_j.f_k\right)} \quad (۳)$$

از این نظر، US_{ij}^{D} شباهت جمعیتی بین یک جفت از کاربران i-ام و j-ام است. حداقل و حداکثر مقادیر US_{ij}^{R} به ترتیب صفر و یک است. $U_i.f_k$ ویژگی k-ام جمعیت‌شناختی کاربر i-ام است. تعداد کل ویژگی‌های جمعیتی با d مشخص می‌شود. w_k مخفف وزن ویژگی جمعیتی k-ام در محاسبه تشابه جمعیتی است. همچنین $\pi(A,B)$، تابعی است که شباهت بین A و B را اندازه‌گیری می‌کند و براساس معادله (۴) تعریف می‌شود:

$$\pi\left(U_i.f_k, U_j.f_k\right) = \begin{cases} 0 & \left(f_k \text{ is nominal}\right) \wedge \left(U_i.f_k \neq U_j.f_k\right) \\ 1 & \left(f_k \text{ is nominal}\right) \wedge \left(U_i.f_k = U_j.f_k\right) \\ \dfrac{U_i.f_k - U_j.f_k}{\max\limits_{i_1} U_{i_1}.f_k - \min\limits_{i_1} U_{i_1}.f_k} & f_k \text{ is numeric} \end{cases} \quad (۴)$$

پس از محاسبه کسینوس و تشابهات جمعیتی در بین کاربران، از معادله (۵) برای محاسبه شباهت نهایی بین کاربران استفاده می‌شود:

$$US_{ij} = \theta_u\left(|A_{ij}|\right) \times US_{ij}^{D} + \left(1 - \theta_u\left(|A_{ij}|\right)\right) \times US_{ij}^{R} \quad (۵)$$

که در آن $\theta_u\left(|A_{ij}|\right)$ پارامتر تعیین‌کننده وابستگی شباهت نهایی بین کاربران به هر یک از کسینوس و شباهت‌های جمعیتی است. در این رابطه، مقدار $\theta_u\left(|A_{ij}|\right) = 1$ نشان‌گر وابستگی کامل شباهت نهایی به شباهت جمعیتی کاربران است و مقدار $\theta_u\left(|A_{ij}|\right) = 0$ وابستگی کامل شباهت نهایی به شباهت کسینوس کاربران را نشان می‌دهد. در مورد رابطه (۵)، می‌توان نتیجه گرفت که حتی اگر هر دو کاربر هیچ قلم رتبه‌بندی مشترکی نداشته باشند، می‌توان شباهت آنها با یکدیگر را بر اساس معیار تشابه جمعیتی محاسبه کرد؛ بنابراین، $\theta_u(n)$ بر اساس معادله (۶) تعریف می‌شود:

$$\theta_u(n) = \begin{cases} \min\left(\dfrac{n}{n+10}, 0.2\right) & n < 4 \\ \min\left(\dfrac{n}{n+100} + 0.2, 1.0\right) & n \geq 4 \end{cases} \quad (۶)$$

پس از محاسبه شباهت نهایی بین کاربران، Z کاربرانی که بیشترین شباهت را با کاربر هدف دارند؛ به‌عنوان مجموعه نزدیک‌ترین همسایگان کاربر هدف انتخاب می‌شوند. مقدار نرخ پیش‌بینی‌شده‌ای که کاربر i-ام (کاربر هدف) به قلم j-ام می‌دهد، با توجه به معادله (۷) برآورد می‌شود:

$$\hat{R}_{ij} = \frac{\sum_{k \in NN_i(Z)} US_{ik} \times R_{kj}}{\sum_{k \in NN_i(Z)} US_{ik}} \quad (۷)$$

$\hat{R}_{ij}$ مقدار نرخ پیش‌بینی شده‌ای است که کاربر i-ام (کاربر هدف) به قلم j می‌دهد، $NN_i(Z)$ شاخص‌های Z نزدیک‌ترین همسایه کاربر هدف، یعنی کاربر-کاربر را نشان می‌دهد؛ سپس، تعداد معینی از قلم‌ها که توسط نزدیک‌ترین کاربران به کاربر هدف (یعنی همسایگان آن) دارای رتبه بالایی هستند، بر اساس معادله (۷) اضافه می شوند. ما آن را راه‌کار میانگین رتبه‌بندی محلی $(LARS)$ نام می‌گذاریم؛ بنابراین، $N_{LARS}(=\mathrm{Z})$ قلم‌های مختلف برای توسعه مشخصات کاربر هدف انتخاب شده است. برای این منظور، قلم‌هایی که دارای بالاترین امتیاز توسط همسایگان Z از کاربر هدف هستند، انتخاب می‌شوند؛ بنابراین، در این راه‌کار، قلم‌های موجود بر اساس رتبه‌بندی پیش‌بینی‌شده‌ای که دریافت کرده‌اند، به‌ترتیب نزولی و قلم‌های N_{LARS} از ابتدای فهرست برای توسعه مشخصات کاربر هدف انتخاب می‌شوند.

[Downloaded from jsdp.rcisp.ac.ir on 2026-09-05]

[DOI: 10.52547/jsdp.19.1.1]

۲-۳- راه‌کار سراسری با بیشترین امتیاز

راه‌کار سراسری با بیشترین امتیاز (GMRS) راهکاری است که در آن قلم‌های مختلف N_{GMRS} برای توسعه مشخصات کاربر هدف انتخاب می‌شود [25]. برای این‌منظور، قلم‌هایی که بیشترین رتبه‌بندی را به‌وسیلهٔ تمام Z همسایگان کاربر هدف داشته باشند، انتخاب می‌شوند؛ بنابراین، در این راهکار، اقلام موجود بر اساس میانگین تعداد امتیازاتی که توسط Z نزدیکترین همسایه کاربر هدف دریافت کرده‌اند، به‌ترتیب نزولی مرتب می‌شوند و قلم‌های N_{GMRS} از ابتدای فهرست برای توسعه کاربر هدف انتخاب می‌شوند.

۳-۳- راه‌کار دارای امتیاز محلی

در این راه‌کار شباهت‌های قلم هدف با سایر قلم‌های موجود در سامانه محاسبه، سپس، برخی از مشابه‌ترین قلم‌ها برای پیش‌بینی رتبه‌بندی قلم هدف انتخاب می شوند. یک رویکرد ترکیبی مبتنی بر کسینوس و شباهت-های هستان‌شناسی بین اقلام ایجاد شده است. از اطلاعات موجود در ماتریس رتبه‌بندی کاربر برای محاسبه تشابه کسینوس بین اقلام استفاده می‌شود؛ همچنین، از اطلاعات کمکی (هستان‌شناسی) اقلام برای محاسبه شباهت جدید (هستان‌شناسی) در بین کاربران استفاده می‌شود. در این روش، برای هر جفت قلم ارزیابی‌شده، رتبه‌بندی کاربران برای هر دو انتخاب، سپس از معیار تشابه کسینوس برای محاسبه شباهت بین آن کاربرها همان‌طور که در معادله (۸) ارائه شده است، استفاده می‌شود:

$$IS_{ij}^{R} = \frac{\sum_{k \in B_{ij}} R_{ki} R_{kj}}{\sqrt{\sum_{k \in B_{ij}} R_{ki}^2 \sum_{k \in B_{ij}} R_{kj}^2}} \qquad (۸)$$

در معادله (۸)، IS_{ij}^{R} شباهت کسینوس بین یک جفت از اقلام i-ام و j-ام است. حداقل و حداکثر مقادیر IS_{ij}^{R} به ترتیب صفر و یک است. B_{ij} مجموعه شاخص‌های آن دسته از کاربران است که به هر دو قلم i-ام و j-ام امتیاز می‌دهند و به‌صورت زیر تعریف می‌شود:

$$B_{ij} = \{k | R_{ki} \neq NaN \wedge R_{kj} \neq NaN\} \qquad (۹)$$

که NaN ناشناخته است.

شباهت اقلام نیز بر اساس داده‌های کمکی محاسبه می‌شود. $I_i.L$ زبان قلم i-ام، $I_i.D$ کارگردان قلم i-ام، $I_i.W$ نویسنده قلم i-ام، $I_i.R$ میانگین نرخ قلم، $I_i.Rt$ زمان اجرا قلم i-ام، $I_i.Rd$ داده‌های انتشار قلم i-ام، $I_i.C$ کشور قلم i-ام را نشان می‌دهد. $I_i.N$ تعداد نرخ‌هایی را که قلم کسب کرده است. $I_i.P$ تولیدکننده قلم i-ام، $I_i.G_j$ یک متغیر نامتقارن بولی است که نشان می‌دهد قلم i-ام ژانر j-ام را دارد یا خیر. $I_i.A_k$ نشان‌دهنده این است که آیا بازیگر معروف k-ام در لیست بازیگران قلم i-ام موجود است یا خیر. بازیگر مشهور یکی از بازیگرهای حداقل پنج قلم است. $I_i.A_k$ یک متغیر نامتقارن بولی است. IS_{ij}^{O} ماتریس تشابه قلم است، که براساس اطلاعات کمکی (هستان‌شناسی) قلم‌ها تعریف شده است (ویژگی‌های تعریف‌شده در این پاراگراف). شباهت‌های هستان‌شناسی بین همه جفت‌ها با توجه به معادله (۱۰) محاسبه می‌شود:

$$IS_{ij}^{O} = Sim(I_i.O, I_j.O) = \frac{\sum_{k=1}^{o} \omega_k \times \pi(I_i.f_k, I_j.f_k)}{\sum_{k=1}^{o} \pi(I_i.f_k, I_j.f_k)} \qquad (۱۰)$$

در این رابطه، IS_{ij}^{O} شباهت جدید (هستان‌شناسی) بین یک جفت از قلم‌های i-ام و j-ام است. کمینه و بیشینه مقادیر IS_{ij}^{R} به‌ترتیب صفر و یک است. تعداد کل ویژگی های کمکی با o نشان داده می‌شود. ω_k مخفف وزن ویژگی کمکی k-ام در محاسبه شباهت جدید (هستان شناسی) است. همچنین، $\pi(A, B)$ تابعی است که شباهت بین A و B را اندازه‌گیری می‌کند و مانند معادله (۴) بر اساس معادله (۱۱) تعریف می‌شود:

$$\pi(I_i.f_k, I_j.f_k) = \begin{cases} 1 & (f_k \text{ is Asy Bool}) \wedge (I_i.f_k = I_j.f_k = 1) \\ 0 & f_k \text{ is Asy Bool} \\ 0 & (f_k \text{ is nominal}) \wedge (I_i.f_k \neq I_j.f_k) \\ 1 & (f_k \text{ is nominal}) \wedge (I_i.f_k = I_j.f_k) \\ \dfrac{I_i.f_k - I_j.f_k}{\max\limits_{i_1} I_{i_1}.f_k - \min\limits_{i_1} I_{i_1}.f_k} & f_k \text{ is numeric} \end{cases} \qquad (۱۱)$$

پس از محاسبه کسینوس و شباهت‌های جدید (هستان‌شناسی) بین اقلام، از معادله (۱۲) برای محاسبه شباهت نهایی بین قلم‌ها استفاده می‌شود:

$$IS_{ij} = \theta_i(|B_{ij}|) \times IS_{ij}^{O} + \left(1 - \theta_i(|B_{ij}|)\right) \times IS_{ij}^{R} \qquad (۱۲)$$

به‌طوری‌که $\theta_i(|B_{ij}|)$، مانند $\theta_u(|A_{ij}|)$، پارامتری است که وابستگی شباهت نهایی بین اقلام به هر یک از کسینوس و شباهت‌های جدید (هستان‌شناسی) را تعیین می‌کند. $\theta_i(|B_{ij}|)$ بر اساس معادله (۱۳) تعریف شده است:

$$\theta_i(n) = \min\left(\frac{n}{n + 1000}, 0.9\right) \qquad (۱۳)$$

پس از محاسبه شباهت نهایی بین قلم‌ها، Q قلمی که بیشترین شباهت را با قلم هدف داشته باشد، انتخاب شده و برای پیش‌بینی رتبه‌بندی استفاده می‌شود. اکنون، این اقلام که بیشترین شباهت را با قلم مورد نظر دارند، به-

[Downloaded from jsdp.rcisp.ac.ir on 2026-09-05]

[DOI: 10.52547/jsdp.19.1.1]

عنوان مجموعه نزدیک‌ترین همسایگان قلم هدف انتخاب می‌شوند. مقدار نرخ پیش‌بینی‌شده‌ای که کاربر i به قلم j می‌دهد طبق معادله (۱۴) برآورد می‌شود:

$$\hat{R}_{ij} = \frac{\sum_{k \in NN_u(Q)} IS_{jk} \times R_{ik}}{\sum_{k \in NN_u(Q)} IS_{jk}} \quad (۱۴)$$

$\hat{R}_{ij}$مقدار نرخ پیش‌بینی‌شده‌ای است که کاربر i به قلم j، یعنی کاربر هدف، می‌دهد. $NN_u(Q)$ مجموعه‌ای از تمام شاخص‌های Q نزدیکترین همسایه قلم هدف را نشان می‌دهد. این راه‌کار، به‌عنوان راه‌کار دارای امتیاز محلی (ILMRS) نامیده می‌شود؛ بنابراین، $N_{ILMRS}(= Q)$ قلم های مختلف برای پیش‌بینی رتبه‌بندی انتخاب می‌شود. برای این منظور، هر قلم j که با توجه به IS_{ij} در بالاترین رتبه N_{ILMRS} باشد، به نمایه قلم اضافه می‌شود.

۴- آزمایش‌ها و نتایج تجربی

۴-۱- مجموعه‌داده

مجموعه‌داده‌های معیار به‌منظور ارزیابی روش پیشنهادی، از مجموعه‌داده‌های Movielens 1m که می‌توان از تارنمای گروه لنز به‌صورت برخط بارگیری کرد، استفاده شده است. این مجموعه شامل یک میلیون رتبه‌بندی است که توسط ۶۰۴۰ کاربر به ۳۹۵۰ ویدیو داده می‌شود. رتبه‌بندی‌های موجود شامل مقیاس عددی پنج نمره‌ای است که، رتبه یک نشان‌دهنده امتیاز بسیار کم، رتبه دو نشان‌دهنده امتیاز کم، رتبه سه نشان‌دهنده امتیاز متوسط ، رتبه چهار نشان‌دهنده علاقه بالا و رتبه پنج نشان‌دهنده علاقه بسیار زیاد کاربران است. در این مجموعه‌داده، هر کاربر دست‌کم بیست قلم را رتبه‌بندی کرده است. یک دهم قلم‌ها، یعنی ۳۹۵ قلم، به‌عنوان قلم سرد استفاده می‌شود. اقلام سرد، اقلامی است که درنهایت بیست امتیاز کسب کند. ما به‌طور تصادفی صفر تا بیست رتبه واقعی برای هر یک از این قلم‌های سرد در مجموعه‌داده قرار می-دهیم. ما آن را به‌عنوان مجموعه‌داده ICS (Item Cold Start) نام‌گذاری می‌کنیم. یک‌دهم کاربران، یعنی ۶۰۴ کاربر، به‌عنوان کاربر سرد استفاده می‌شوند. ما آن را به عنوان مجموعه‌داده UCS (User Cold Start) نامگذاری می‌کنیم. در مجموعه‌داده مورد استفاده برای ارزیابی روش پیشنهادی، هر کاربر دست‌کم بیست قلم را ارزیابی کرده است؛ بنابراین، برای ارزیابی روش پیشنهادی در شرایط شروع سرد (به‌عنوان مثال مجموعه داده UCS)، تعدادی از قلم‌ها به‌طور تصادفی بر اساس روش پیشنهادی برای هر کاربر انتخاب می‌شوند. این تعداد قلم انتخاب‌شده کمتر از کل قلم‌های ارزیابی‌شده توسط هر کاربر است. برای هر کاربر، تعداد قلم‌های اولیه در نمایه برابر با مقادیر مختلف تا ده است. یک‌دهم رتبه‌بندی نیز به‌عنوان مجموعه‌آزمون عمومی در یک بستر آزمایش دیگر استفاده می‌شود. ما آن را به‌عنوان مجموعه‌داده UCS (User Cold Start) نام‌گذاری می‌کنیم. گفتنی است که ICS، UCSیا TNبه‌طور تصادفی سی بار تولید می‌شوند و در نتیجه، ما برای هر روش سی آزمایش مستقل بر روی ICS، UCS یا TN انجام می‌دهیم و نتیجه متوسط در این سی آزمایش مستقل بوده و به‌عنوان عملکرد روش در نظر گرفته شده است.

۴-۲- معیارهای ارزیابی

برای ارزیابی روش پیشنهادی، معیار میانگین خطای مطلق (MAE) با استفاده از معادله (۱۵) محاسبه می‌شود:

$$MAE = \frac{\sum_i \sum_j |R_{ij} - \hat{R}_{ij}|}{TestSize} \quad (۱۵)$$

در معادله (۱۵)، R_{ij} امتیاز واقعی کاربر i به قلم j است. $\hat{R}_{ij}$ رتبه پیش‌بینی‌شده کاربر i به قلم j است و $TestSize$تعداد کل رتبه‌بندی‌های پیش‌بینی شده است. همچنین، معیار تصحیح رتبه‌بندی (RC) با استفاده از معادله (۱۶) محاسبه می‌شود:

$$RC = \frac{\sum_i \sum_j \delta\left(R_{ij} == round(\hat{R}_{ij})\right)}{TestSize} \quad (۱۶)$$

اگر A درست باشد $\delta(A)$ یک است، در غیر این‌صورت صفر است. آزمون معناداری [36] روشی آماری برای تأیید اعتبار این است که تفاوت بین عملکرد دو یا چند روش از نظر آماری با سطح اطمینان (p-1) را نشان می‌دهد. آزمون آماری معناداری را می‌توان با معیارهای مختلف ارزیابی انجام داد. اصطلاح در سطح اطمینان p به معنی با احتمال pاست.

۴-۳- مقایسه و ارزیابی نتایج تجربی

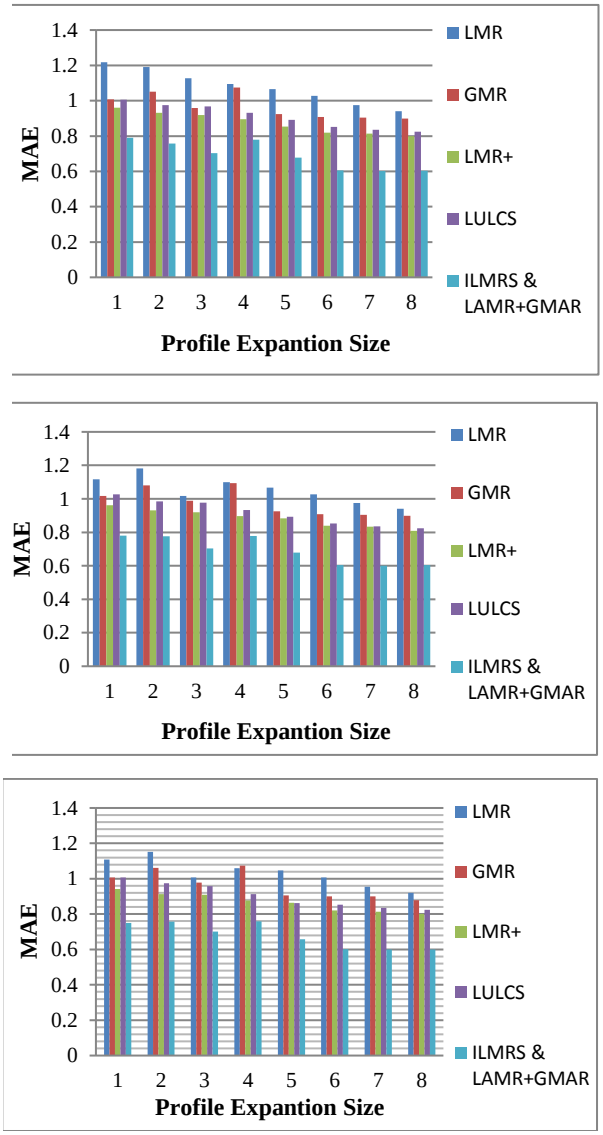
برای ارزیابی روش پیشنهادی، ما آن را با برخی از روش های اولیه خام مانند LMR، LULCS، GMR، LMR + (LMR با استفاده از اطلاعات جمعیت‌شناختی) و همچنین سامانه پیشنهادگر پالایش مشارکتی کاربر پایه (یا گسترش بدون نمایه به اختصار NoPE مقایسه می‌کنیم) همچنین، ما از روش‌های مختلف دیگری مانند غیرنرمال-

[Downloaded from jsdp.rcisp.ac.ir on 2026-09-05]

[DOI: 10.52547/jsdp.19.1.1]

سازی ConF RS [37]، RS بر اساس تجزیه مقدار منفرد (با SVD نشان داده شده) [38]، RS بر اساس محبوبیت (Pop) [39] و Top − N بر اساس هستان‌شناسی استفاده می‌کنیم (OTopN) [40]. همه این روش‌ها از پارامترهای پیش‌فرض خود استفاده می‌کنند [41-46]. این آزمون‌ها به‌طور متفاوت در سه مجموعه‌داده انجام شده است ICS، UCSو TN؛ علاوه‌براین، وزن هر داده جمعیتی از نظر سن، جنس و شغل به‌وسیلهٔ یک مجموعه اعتبارسنجی تنظیم می‌شود. همچنین، وزن هر اطلاعات هستان‌شناسی قلم‌ها با یک مجموعه اعتبارسنجی تنظیم می‌شود. پارامترهای Z و Q روی ۲۵ تنظیم شده‌اند، زیرا از نظر آزمایشی بهترین انتخاب است [36].

روش‌های استفاده شده در آزمایش‌ها با توجه به دو معیار مقایسه می‌شوند: (الف) MAE و (ب) RC (گفتنی است که روش پیشنهادی با توجه به نوع راه‌کار مورد استفاده برای توسعه مشخصات، نامگذاری شده است، در اینجا ILMRS & LAMR + GMAR نامیده شده است.) شکل (۳) نتایج مربوط به معیار MAE را برای روش پیشنهادی و سایر روش‌های مختلف نشان می‌دهد. همان طور که از نتایج شکل (۳) مشخص است، روش پیشنهادی به‌طورتقریبی در همه موارد بهترین عملکرد را دارد. برای وضوح بیشتر شکل (۴) نیز ترسیم شده است تا بتوان برتری روش پیشنهادی را نیز با هفت روش دیگر نشان داد (برای وضوح بهتر و مقایسه با روش‌های LMR، GMR، LMR+ و LULCS شکل (۴) ترسیم شد). در این بخش، روش‌های مختلف بر اساس دو معیار MAE و RC بنا شده است. همان‌طور که از نتایج شکل (۵) مشخص است، روش پیشنهادی از نظر دقت به‌طورتقریبی بهترین عملکرد را در همه موارد دارد. نتایج در شکل‌های (۳، ۴ و ۵) نشان می‌دهد که روش پیشنهادی بهترین عملکرد را بر اساس معیارهای RC و MAE در همه مجموعه‌داده‌ها دارد. همان‌طور که در شکل (۳) نشان داده شده است، بهترین مقدار MAE برای روش‌های مختلف RS زمانی به‌دست می‌آید که مقدار پارامتر $l = 10$ باشد. همچنین بهترین مقدار RC برای روش‌های مختلف RS زمانی به‌دست می آید که مقدار پارامتر $l = 50$ باشد. بهترین مقدار RC برای روش RS پیشنهادی زمانی به‌دست می‌آید که مقدار پارامتر $l \geq 10$ باشد؛ بنابراین، طبق نتایج در شکل‌های (۳، ۴ و ۵)، بهترین اندازه توسعه پروفایل ده است؛ بنابراین، اندازه توسعه پروفایل از اینجا به بعد ده خواهد بود، یعنی $l = 10$. شکل‌های (۳، ۴ و ۵) به‌ترتیب مقادیر MAEو RC را برای روش‌های مختلف RS نشان می‌دهد. آزمون آماری معناداری که بر روی نتایج شکل (۵) انجام شده است، مقدار p را ۰/۰۳۷۵ نشان می‌دهد. شکل (۶) دقت پیشنهادهای Top-N ارائه‌شده به‌وسیلهٔ روش‌های مختلف RS را در مجموعه داده ICS نشان می‌دهد. برای مقایسه روش پیشنهادی با سایر روش‌های به‌نسبه جدید و مشابه از نمودار شکل (۷) استفاده شده است. در این نمودار از معیاری به نام معیار فیشر استفاده شده است. هر چه مقدار این معیار بیشتر باشد، روش ارائه‌شده عملکرد بهتری دارد. روش‌های مقایسه‌شده در نمودار آورده شده‌اند.

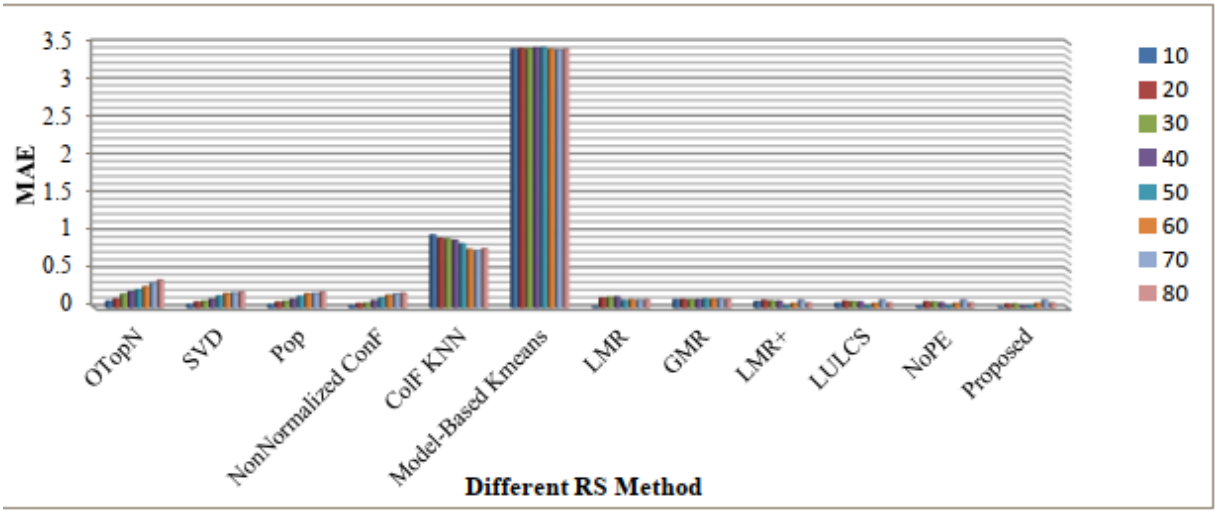


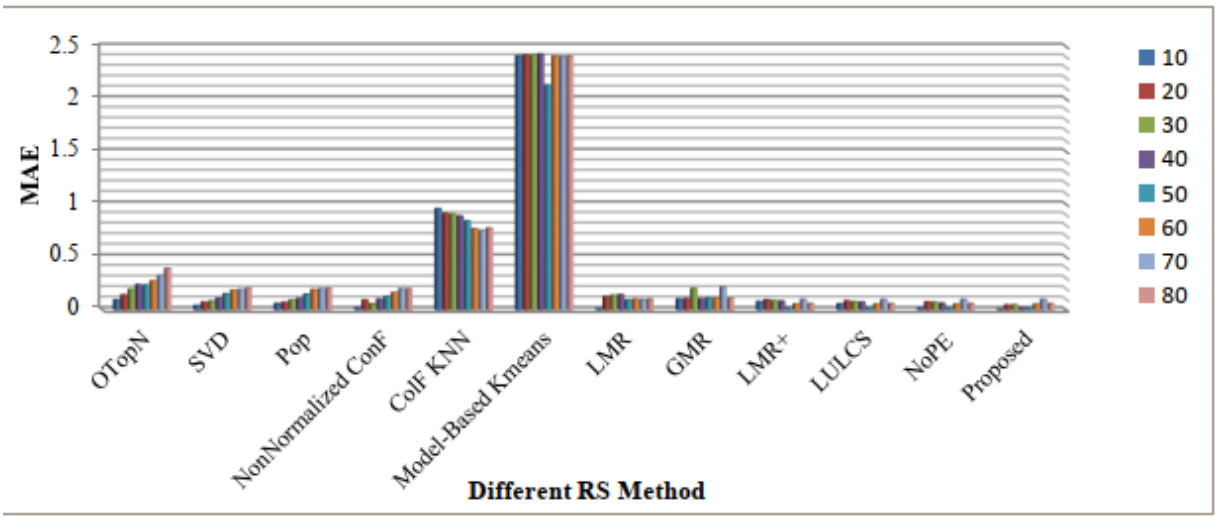
(شکل-۳): مقایسه عملکرد سامانه‌های پیشنهادگر مختلف بر حسب معیار MAE برای اندازه‌های مختلف گسترش پروفایل کاربر در (بالا) مجموعه داده TN، (وسط) مجموعه‌داده ICS و (پایین) مجموعه‌داده UCS

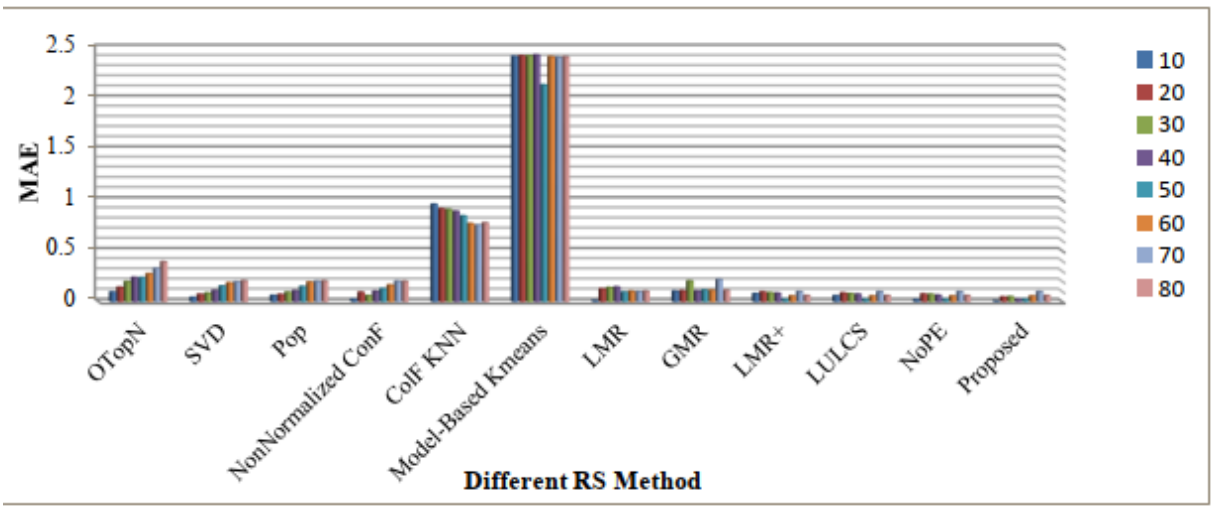
(Figure-3): The performance comparison of different RSs in terms of MAE for different profile expansion sizes on (top) TN dataset, (center) ICS dataset and (bottom) UCS dataset

[Downloaded from jsdp.rcisp.ac.ir on 2026-09-05]

[DOI: 10.52547/jsdp.19.1.1]





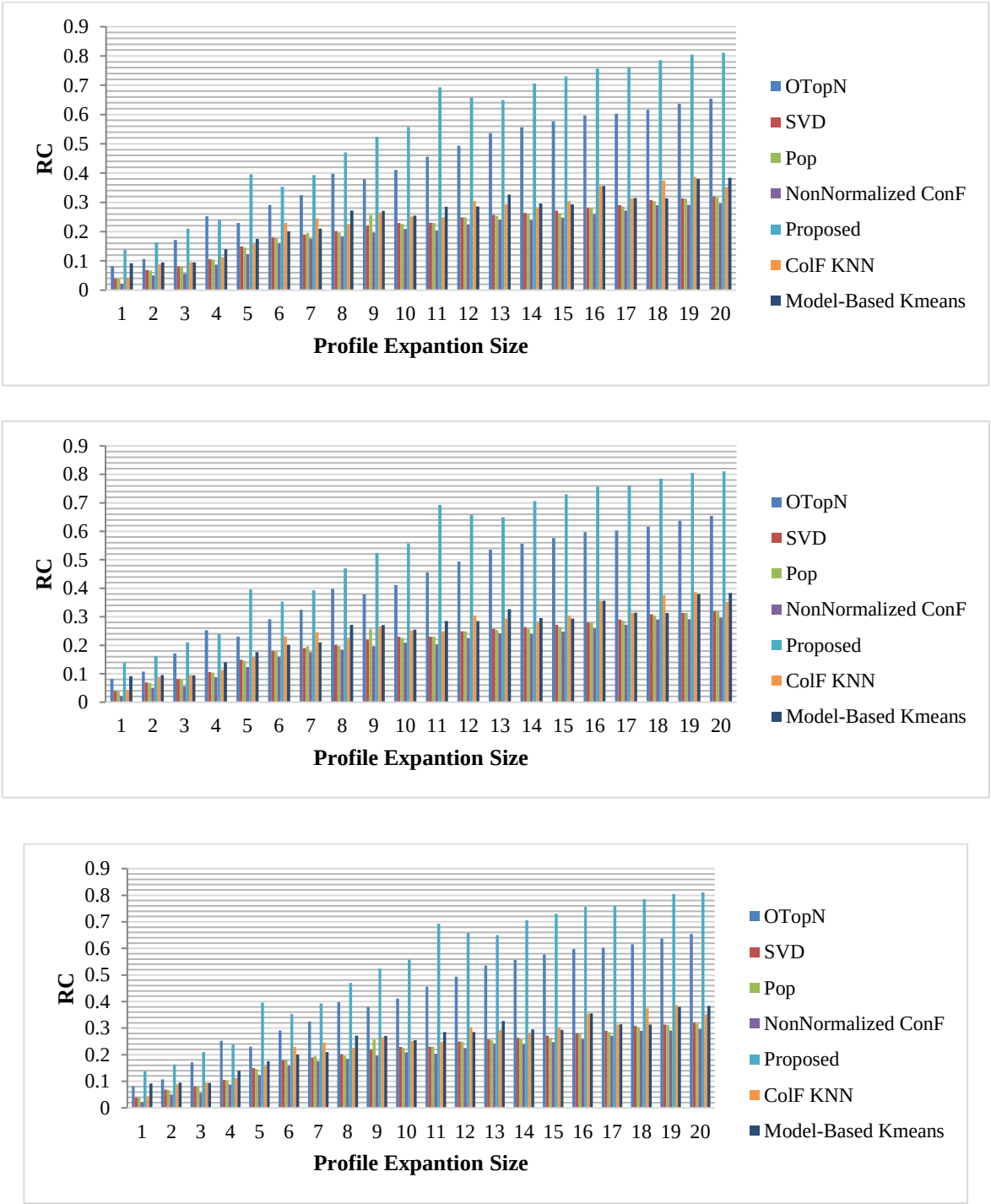


(شکل-۴): مقایسه عملکرد چند سامانه پیشنهادگر جدید با روش پیشنهادی بر حسب معیار MAE برای اندازه‌های مختلف گسترش پروفایل کاربر در (بالا) مجموعه‌داده TN، (وسط) مجموعه‌داده ICS و (پایین) مجموعه‌داده UCS

(Figure-4): Comparison of the performance of several new RSs with the proposed method in terms of MAE for different profile expansion sizes on (top) TN dataset, (center) ICS dataset and (bottom) UCS dataset

[Downloaded from jsdp.rcisp.ac.ir on 2026-09-05]

[DOI: 10.52547/jsdp.19.1.1]

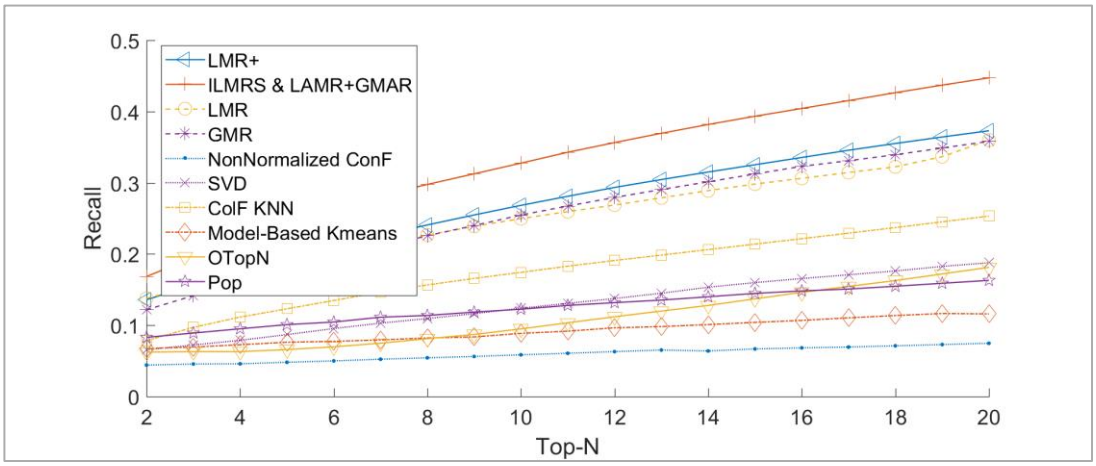


(شکل-۵): مقایسه عملکرد سامانه‌های پیشنهادگر مختلف بر حسب معیار RC برای اندازه‌های مختلف گسترش پروفایل کاربر در (بالا) مجموعه‌داده TN، (وسط) مجموعه‌داده ICS و (پایین) مجموعه‌داده UCS

(Figure-5): The performance comparison of different RSs in terms of RC for different profile expansion sizes on (top) TN dataset, (center) ICS dataset and (bottom) UCS dataset

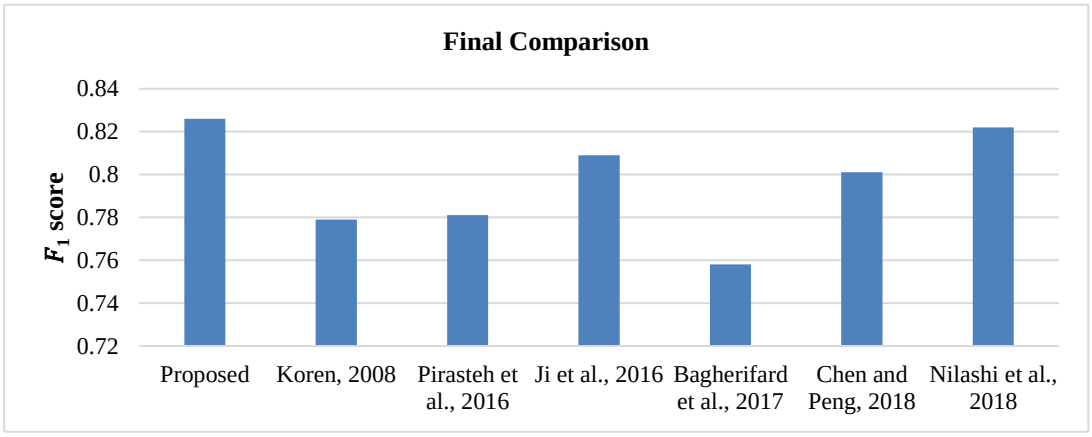
[Downloaded from jsdp.rcisp.ac.ir on 2026-09-05]

[DOI: 10.52547/jsdp.19.1.1]



(شکل-۶): مقایسه عملکرد سامانه‌های پیشنهادگر مختلف بر حسب معیار Recall برای اندازه‌های مختلف تعداد N پیشنهاد برتر، برای گسترش پروفایل کاربر با اندازه ۱۰ در مجموعه‌داده ICS

(Figure-6): The performance comparison of different RSs in terms of Recall of Top-N recommendations for profile expansion size of 10 on ICS dataset



(شکل-۷): مقایسه عملکرد سامانه‌های پیشنهادگر مختلف بر حسب میانگین معیار F_1 در تمام مجموعه‌داده‌ها

(Figure-7): The performance comparison of different RSs in terms of F_1 measure averaged on all datasets

۴-۴- آزمون آماری معناداری

آزمون فریدمن یک آزمون ناپارامتری است که برای مقایسه سه یا بیش از سه گروه وابسته که دست‌کم در سطح رتبه‌ای اندازه‌گیری می‌شوند، مورد استفاده قرار می گیرد. این آزمون می‌تواند در مورد داده‌های پیوسته (فاصله‌ای یا نسبی) نیز به‌کار برده شود، اما در هنگام محاسبه این داده‌ها نیز رتبه‌بندی آنها مدنظر قرار می گیرد. آزمون فریدمن معادل ناپارامتری آزمون F وابسته در تحلیل واریانس اندازه‌های تکراری است. در این حالت برای اجرای تحلیل واریانس داده‌های تکرارشده ضرورتی به‌وجود فرضیاتی مانند نرمال بودن توزیع، برابری واریانس‌ها و پیوسته بودن مقیاس وجود ندارد؛ بنابراین در تحلیل واریانس اندازه‌های تکراری چنانچه یک یا همه فرضیات ابتدایی یادشده رد شوند، از آزمون فریدمن استفاده می‌شود. فرضیه صفر (H0) در این آزمون بیان می‌کند که توزیع مشاهدات در سنجش‌های تکرارشده یکسان هستند. یا به‌عبارت دیگر میان توزیع‌های ایجاد شده در اثر سنجش‌های مکرر روی یک گروه و یا بین گروه‌های همتا در زمینه متغیر وابسته تفاوتی وجود ندارد. همچنین یک فرضیه جایگزین (H1) وجود دارد که با فرضیه صفر رقابت می‌کند و پیشنهاد می‌کند دست‌کم یکی از نتایج متفاوت است. در این کار، نتایج محاسبه‌شده با استفاده از روش پیشنهادی و سایر رویکردها، مقایسه و تحلیل می‌شود که آیا تفاوت بین نتایج با استفاده از آزمون t یا فریدمن از نظر آماری معنی‌دار است. آزمون فریدمن یک آزمون آماری است که برای ارزیابی این‌که تفاوت بین دو مجموعه‌داده تصادفی است یا از نظر آماری معنی‌دار است. محاسبه آماره فریدمن که آن را با x_r^2 نشان می‌دهند با استفاده از رابطه زیر امکان‌پذیر است:

$$x_r^2 = \frac{SS_{br}}{k(k+1)/12} \qquad (۱۷)$$

[Downloaded from jsdp.rcisp.ac.ir on 2026-09-05]

[DOI: 10.52547/jsdp.19.1.1]

که در آن SS_{br} مجموع مجذورات رتبه‌ای بین توزیع‌ها و kتعداد توزیع‌ها است که رتبه‌بندی در مورد آنها انجام می‌شود.

(جدول-۱): آزمون t و نتایج آزمون فریدمن برای دقت محاسبه‌شده با رویکرد پیشنهادی و رویکردهای دیگر
(Table-1): The t-test analysis between accuracies of different methods and the proposed method and their friedman-test analysis

$Friedman's\ test$	$t-test$						
	Nilashi et al.	**Chen and Peng.**	**Ji et al.**	**Pirasteh et al.**	**Koren.**	**Bagherifard et al.**	
–	0.9969	0.9960	0.9959	0.9951	0.9947	0.9901	**Mean**
1	1	1	1	1	1	1	**h**
–	3.6619	3.0530	6.3346	9.0459	12.3997	6.0955	**t - stat**
3.0774e−04	0.0026e−01	0.0037e−01	0.0032e−01	8.2749e−04	2.4317e−04	0.0037e−01	**p - value**

جدول (۱) نتایج آزمون t را با استفاده از نتایج محاسبه‌شده با رویکرد پیشنهادی و سایر رویکردها ارائه می‌دهد. برای این آزمون، فرضیه صفر H0 یعنی این‌که، معنی نتایج روش‌های مختلف برابر است و H1 یعنی این-که، معنی نتایج روش‌های مختلف نابرابر است. مقادیر $P\ (T \leq t)$ (مقدار p) بسیار نزدیک به صفر نشان می‌دهد که اطمینان ارزیابی بالاتر از ۹۹ درصد است؛ زیرا $P\ (T \leq t) \ll 0.05$ است. همچنین، مقادیر t-stat بزرگتر از صفر است و مقدار برگشتی $h = 1$ نشان می‌دهد که آزمون t فرضیه صفر را در سطح معنی‌داری 5% یعنی $\alpha = 0.05$ رد می‌کند. از این‌رو، رویکرد ما به‌طور قابل توجهی سایر رویکردها را بهبود بخشید. جدول (۱) نتایج به‌دست آمده از آزمون فریدمن را برای $\alpha = 0.05$ نشان می‌دهد. نتیجه توصیف می‌کند که همه رویکردها میانگین منحصربه‌فردی دارند و مقدار p بسیار نزدیک به صفر است که آزمون فرضیه صفر را رد می‌کند.

۵- نتیجه‌گیری

در این مقاله، رویکرد ترکیبی جدیدی مبتنی بر ترکیب شباهت جمعیت‌شناختی و شباهت کسینوس بین کاربران به‌منظور حل مشکل شروع سرد از نوع کاربر جدید و همچنین، رویکرد جدیدی مبتنی بر ترکیب شباهت هستان‌شناسی و شباهت کسینوس بین اقلام به‌منظور حل مسأله شروع سرد از نوع قلم جدید ارائه شده است. ایده اصلی روش پیشنهادی توسعه پروفایل‌های کاربر بر اساس راه‌کارهای مختلف برای ایجاد پروفایل با عملکرد بالاتر برای کاربران است. نتایج حاصل از آزمایش‌ها عملکرد بهتر روش پیشنهادی را نسبت به سایر روش‌ها نشان می‌دهد. یکی از پیشنهادهایی که می‌تواند برای کارهای آینده در نظر گرفته شود، استفاده از اطلاعات مختلف در مورد محتوای قلم‌ها و کاربران است. استفاده از این اطلاعات اضافی مربوط به قلم‌ها و کاربران می‌تواند باعث افزایش کارایی سامانه‌های پیشنهادگر، به‌ویژه در مواجهه با مسأله شروع سرد شود. تاکنون، ما چندین راه‌حل برای مشکل شروع سرد در سامانه‌های پیشنهادگر مورد بحث قرار داده-ایم. با این‌حال، استفاده از هستان‌شناسی قلم، هستان-شناسی کاربر، شباهت معنایی بهبودیافته به‌وسیلهٔ WordNet می‌تواند رهنمودهای آینده پژوهش باشد.

۶- مراجع 6- References

[1] R. Yera and L. Martinez, "Fuzzy tools in recommender systems: A survey," *International Journal of Computational Intelligence Systems*, vol. 10, pp. 776-803, 2017.

[2] M. D. Ekstrand and J. A. Konstan, "Recommender Systems Notation," 2019.

[3] M. Doja, "Recommender System for Personalized Adaptive E-learning Platforms to Enhance Learning Capabilities of Learners Based on their Learning Style and Knowledge Level," 2019.

[4] F. S. d. Aguiar Neto, "Pre-processing approaches for collaborative filtering based on hierarchical clustering," Universidade de São Paulo.

[5] M.-P. T. Do, D. V. Nguyen, and L. Nguyen, "Model-based Approach for Collaborative Filtering," 2019.

[Downloaded from jsdp.rcisp.ac.ir on 2026-09-05]

[DOI: 10.52547/jsdp.19.1.1]

[6] Y. Yang, Y. Xu, E. Wang, J. Han, and Z. Yu, "Improving existing collaborative filtering recommendations via serendipity-based algorithm," *IEEE Transactions on Multimedia*, vol. 20, pp. 1888-1900, 2017.

[7] S. K. Raghuwanshi and R. Pateriya, "Collaborative Filtering Techniques in Recommendation Systems," in *Data, Engineering and Applications*, ed: Springer, 2019, pp. 11-21.

[8] T. N. Duong, V. D. Than, T. H. Tran, Q. H. Dang, D. M. Nguyen, and H. M. Pham, "An Effective Similarity Measure for Neighborhood-based Collaborative Filtering," in *2018 5th NAFOSTED Conference on Information and Computer Science (NICS)*, 2018, pp. 250-254.

[9] J. Feng, X. Fengs, N. Zhang, and J. Peng, "An improved collaborative filtering method based on similarity," *PloS one*, vol. 13, pp. e0204003, 2018.

[10] P. Thakkar, K. Varma, V. Ukani, S. Mankad, and S. Tanwar, "Combining User-Based and Item-Based Collaborative Filtering Using Machine Learning," in *Information and Communication Technology for Intelligent Systems*, ed: Springer, 2019, pp. 173-180.

[11] Z. Yang, C. Fu, R. Lin, T. Peng, and Y. Tang, "Collaborative Filtering Recommendation Algorithm Based on AdaBoost-Naïve Bayesian Algorithm," in *International Conference on Human Centered Computing*, 2018, pp. 380-392.

[12] B. S. Neysiani, N. Soltani, R. Mofidi, and M. H. Nadimi-Shahraki, "Improve Performance of Association Rule-Based Collaborative Filtering Recommendation Systems using Genetic Algorithm," 2019.

[13] J. Borràs, A. Moreno, and A. Valls, "Intelligent tourism recommender systems: A survey," *Expert Systems with Applications*, vol. 41, pp. 7370-7389, 2014.

[14] S. Gong and H. Ye, "An item based collaborative filtering using bp neural networks prediction," in *2009 International Conference on Industrial and Information Systems*, 2009, pp. 146-148.

[15] A. Abdelwahab, H. Sekiya, I. Matsuba, Y. Horiuchi, S. Kuroiwa, and M. Nishida, "An efficient collaborative filtering algorithm using SVD-free Latent Semantic Indexing and particle swarm optimization," in *2009 International Conference on Natural Language Processing and Knowledge Engineering*, 2009, pp. 1-4.

[16] I. Viktoratos, A. Tsadiras, and N. Bassiliades, "Combining community-based knowledge with association rule mining to alleviate the cold start problem in context-aware recommender systems," *Expert Systems with Applications*, vol. 101, pp. 78-90, 2018.

[17] Z. Li, H. Zhao, Q. Liu, Z. Huang, T. Mei, and E. Chen, "Learning from history and present: Next-item recommendation via discriminatively exploiting user behaviors," in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2018, pp. 1734-1743.

[18] R. Logesh, V. Subramaniyaswamy, D. Malathi, N. Sivaramakrishnan, and V. Vijayakumar, "Enhancing recommendation stability of collaborative filtering recommender system through bio-inspired clustering ensemble method," *Neural Computing and Applications*, pp. 1-24, 2019.

[19] Y. Qian, Y. Zhang, X. Ma, H. Yu, and L. Peng, "EARS: Emotion-aware recommender system based on hybrid information fusion," *Information Fusion*, vol. 46, pp. 141-146, 2019.

[20] T. Mohammadpour, A. M. Bidgoli, R. Enayatifar, and H. H. S. Javadi, "Efficient clustering in collaborative filtering recommender system: Hybrid method based on genetic algorithm and gravitational emulation local search algorithm," *Genomics*, 2019.

[21] P. Valdiviezo-Díaz and J. Bobadilla, "A Hybrid Approach of Recommendation via Extended Matrix Based on Collaborative Filtering with Demographics Information," in *International Conference on Technology Trends*, 2018, pp. 384-398.

[22] M. Batet, A. Moreno, D. Sánchez, D. Isern, and A. Valls, "Turist@: Agent-based personalised recommendation of tourist activities," *Expert Systems with Applications*, vol. 39, pp. 7319-7329, 2012.

[23] D. Kotkov, J. A. Konstan, Q. Zhao, and J. Veijalainen, "Investigating serendipity in recommender systems based on real user feedback," in *Proceedings of the 33rd Annual ACM Symposium on Applied Computing*, 2018, pp. 1341-1350.

[24] M. Eirinaki, J. Gao, I. Varlamis, and K. Tserpes, "Recommender systems for large-scale social networks: A review of challenges and solutions," ed: Elsevier, 2018.

[25] M. Y. H. Al-Shamri, "User profiling approaches for demographic recommender systems," *Knowledge-Based Systems*, vol. 100, pp. 175-187, 2016.

[26] L. Safoury and A. Salah, "Exploiting user demographic attributes for solving cold-start problem in recommender system," *Lecture Notes on Software Engineering*, vol. 1, pp. 303-307, 2013.

[Downloaded from jsdp.rcisp.ac.ir on 2026-09-05]

[DOI: 10.52547/jsdp.19.1.1]

[27] M. M. Khan, R. Ibrahim, M. Younas, I. Ghani, and S. R. Jeong, "Facebook interactions utilization for addressing recommender systems cold start problem across system domain," *Journal of Internet Technology*, vol. 19, pp. 861-870, 2018.

[28] V. S. Dixit and P. Jain, "Recommendations with Sparsity Based Weighted Context Framework," in *International Conference on Computational Science and Its Applications*, 2018, pp. 289-305.

[29] H. J. Ahn, "A new similarity measure for collaborative filtering to alleviate the new user cold-starting problem," *Information Sciences*, vol. 178, pp. 37-51, 2008.

[30] V. Formoso, D. FernáNdez, F. Cacheda, and V. Carneiro, "Using profile expansion techniques to alleviate the new user problem," *Information processing & management*, vol. 49, pp. 659-672, 2013.

[31] R. Attar and A. S. Fraenkel, "Local feedback in full-text retrieval systems," *Journal of the ACM (JACM)*, vol. 24, pp. 397-417, 1977.

[32] A. M. Acilar and A. Arslan, "A collaborative filtering method based on artificial immune network," *Expert Systems with Applications*, vol. 36, pp. 8324-8332, 2009.

[33] G. Guo, "Improving the performance of recommender systems by alleviating the data sparsity and cold start problems," in *Twenty-Third International Joint Conference on Artificial Intelligence*, 2013.

[34] G. Shaw, Y. Xu, and S. Geva, "Using association rules to solve the cold-start problem in recommender systems," in *Pacific-Asia conference on knowledge discovery and data mining*, 2010, pp. 340-347.

[35] Q. Liu, E. Chen, H. Xiong, C. H. Ding, and J. Chen, "Enhancing collaborative filtering by user interest expansion via personalized ranking," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 42, pp. 218-233, 2011.

[36] J. Demšar, "Statistical comparisons of classifiers over multiple data sets," *Journal of Machine learning research*, vol. 7, pp. 1-30, 2006.

[37] G. Karypis, "Evaluation of item-based top-n recommendation algorithms," in *Proceedings of the tenth international conference on Information and knowledge management*, 2001, pp. 247-254.

[38] P. Cremonesi, Y. Koren, and R. Turrin, "Performance of recommender algorithms on top-n recommendation tasks," in *Proceedings of the fourth ACM conference on Recommender systems*, 2010, pp. 39-46.

[39] R. Bambini, P. Cremonesi, and R. Turrin, "A recommender system for an IPTV service provider: a real large-scale production environment," in *Recommender systems handbook*, ed: Springer, 2011, pp. 299-331.

[40] H. Cui, M. Zhu, and S. Yao, "Ontology-based Top-N Recommendations on new items with matrix factorization," *Journal of Software*, vol. 9, pp. 2026-2032, 2014.

[41] J.Zhong, , H. Xie, & F.L. Wang, "The research trends in recommender systems for e-learning: A systematic review of SSCI journal articles from 2014 to 2018", *Asian Association of Open Universities Journal*, vol.*14*(1), pp.12–27, 2019.

[42] V.Vanitha, P. Krishnan, "A modified ant colony algorithm for personalized learning path construction", *Journal of Intelligent & Fuzzy Systems*, vol. *37*(5), pp. 6785–6800.

[43] L. H. Son, Dealing with the new user cold-start problem in recommender systems: A comparative review. *Information Systems*, vol.*58*, pp. 87–104, 2016.

[44] N.Silva, D.Carvalho, A. C.Pereira, F. Mourão, & L.Rocha, "The pure cold-start problem: A deep study about how to conquer first-time users in recommendations domains", *Information Systems*, vol. *80*, pp. 1–12, 2019.

[45] L.Romero, C. Saucedo, M. L. Caliusco, & M.Gutiérrez, "Supporting self-regulated learning and personalization using ePortfolios: a semantic approach based on learning paths", *International Journal of Educational Technology in Higher Education*, vol.16(1), pp.16, 2019.

[46] J.R. Almeida, E.Monteiro, L.B.Silva, A.P.Sierra, J.L.Oliveira, "A Recommender System to Help Discovering Cohorts in Rare Diseases", *In Proceedings of the 2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS*), Rochester, MN, USA, pp.28–30, July, 2020.

پیام بحرانی دانش‌آموخته دوره دکترای تخصصی دانشگاه آزاد اسلامی واحد علوم و تحقیقات تهران در رشته مهندسی رایانه گرایش سامانه‌های نرم افزاری است.

زمینه‌های پژوهشی مورد علاقه ایشان مباحثی نظیر سامانه‌های امنیت اطلاعات، هستان‌شناسی و سامانه‌های پیشنهادگر است.

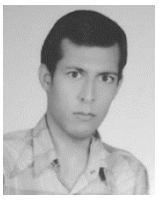
نشانی رایانامه ایشان عبارت است از:

bahranipayam@gmail.com

[Downloaded from jsdp.rcisp.ac.ir on 2026-09-05]

[DOI: 10.52547/jsdp.19.1.1]

بهروز مینایی بیدگلی دانش آموخته دانشگاه ایالتی میشیگان آمریکا در رشته علوم و مهندسی رایانه با تخصص هوش مصنوعی و داده‌کاوی است. وی در حال حاضر عضو هیأت علمی و دانشیار دانشکده مهندسی کامپیوتر دانشگاه علم و صنعت و رئیس دانشکده مهندسی کامپیوتر است. ایشان سرپرستی گروه پژوهشی فناوری‌های بازی‌های رایانه‌ای و نیز آزمایشگاه داده‌کاوی را به عهده دارد. محاسبات نرم، یادگیری ماشین، بازی‌های رایانه‌ای، داده‌کاوی، متن‌کاوی، و پردازش زبان طبیعی، زمینه‌های پژوهشی مورد علاقه ایشان است.
نشانی رایانامه ایشان عبارت است از:
b_minaei@iust.ac.ir



حمید پروین تحصیلات خود را در مقطع کارشناسی در دانشگاه چمران اهواز به پایان رساند. ایشان مدرک کارشناسی ارشد و دکترا را در دانشگاه علم و صنعت دریافت کردند و پس از آن به عضویت هیأت علمی دانشگاه آزاد اسلامی واحد نورآباد ممسنی درآمدند. وی هم‌اکنون در چندین واحد دانشگاهی در رشته رایانه مشغول به تدریس است. زمینه‌های پژوهشی وی مباحثی نظیر الگوریتم‌های بهینه‌سازی، طبقه‌بندی و خوشه‌بندی داده‌ها است.
نشانی رایانامه ایشان عبارت است از:
parvin@iust.ac.ir

میترا میرزارضایی استادیار دانشکده فنی و مهندسی گروه رایانه دانشگاه آزاد اسلامی واحد علوم و تحقیقات تهران هستند. زمینه‌های پژوهشی مورد علاقه ایشان، یادگیری ماشین و شناسایی الگوها است.
نشانی رایانامه ایشان عبارت است از:
mirzarezaee@srbiau.ac.ir

احمد کشاورز مدارک کارشناسی و کارشناسی ارشد خود را به‌ترتیب در سال‌های ۱۳۸۰ و ۱۳۸۳ از دانشگاه شیراز و تربیت مدرس در رشته مهندسی برق و مخابرات سیستم دریافت کرد. ایشان درجه دکترای خود را در سال ۱۳۸۷ از دانشگاه تربیت مدرس در رشته مخابرات سیستم دریافت کرده است. وی هم‌اکنون دانشیار گروه مهندسی برق دانشگاه خلیج فارس است. زمینه‌های پژوهشی مورد علاقه ایشان عبارت است از: سنجش از دور، پردازش تصاویر پزشکی، ماشین بینایی و هوش مصنوعی.
نشانی رایانامه ایشان عبارت است از:
a.keshavarz@pgu.ac.ir

[Downloaded from jsdp.rcisp.ac.ir on 2026-09-05]

[DOI: 10.52547/jsdp.19.1.1]

[Downloaded from jsdp.rcisp.ac.ir on 2026-09-05]

[DOI: 10.52547/jsdp.19.1.1]