سامانهٔ پیشنهاددهندهٔ ترکیبی با کمک واریانس رتبه‌بندی اقلام

پیام بحرانی[۱]، بهروز مینایی بیدگلی[۲]، حمید پروین*[۳]، میترا میرزارضایی[۴] و احمد کشاورز[۵]

[۱و۴]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

[۳]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

مدل نزدیک‌ترین همسایگی (KNN) و سامانه‌های توصیه‌گر مبتنی بر این مدل (KRS) از موفق‌ترین سامانه‌های توصیه‌گر در حال حاضر در دسترس هستند. این روش‌ها شامل پیش‌بینی رتبه‌بندی یک آیتم بر اساس میانگین رتبه‌بندی آیتم‌های مشابه است. میانگین رتبه‌بندی آیتم‌های مشابه، با درنظرگرفتن تشابه تعریف شده، میانگین امتیازی را به هر آیتم، به‌عنوان ویژگی به آن خواهد داد. در این مقاله KRS ایجادشده با ترکیب رویکردهای زیر ارائه شده است: (الف) استفاده از میانگین و واریانس رتبه‌بندی اقلام به‌عنوان ویژگی‌های آیتم، برای یافتن موارد مشابه در (IKRS)؛ (ب) استفاده از میانگین و واریانس رتبه‌بندی کاربر به‌عنوان ویژگی‌های کاربر برای یافتن کاربران مشابه با KRS کاربرپسند (UKRS)؛ (ج) استفاده از میانگین وزنی برای تلفیق رتبه‌بندی کاربران/آیتم‌های همسایه. (د) استفاده از یادگیری جمعی. سه روش پیشنهادی EVMBR، EWVMBR و EWVMBR-G در این مقاله پیشنهاد داده شده است. هر سه روش مبتنی بر کاربر بوده، که در آن‌ها از فاصله VM به‌عنوان معیار تفاوت بین کاربران/آیتم‌ها، برای یافتن کاربران/آیتم‌های همسایه استفاده و سپس به‌ترتیب از میانگین غیروزنی، وزنی و وزنی بر اساس مدل ترکیبی کوواریانس کامل گوسین، برای پیش‌بینی رتبه‌بندی کاربر ناشناخته استفاده می‌شوند. هر سه روش مبتنی بر کاربر بوده، که در آن‌ها از فاصله VM به‌عنوان معیار تفاوت بین کاربران/آیتم‌ها، برای یافتن کاربران/آیتم‌های همسایه استفاده و سپس میانگین به‌ترتیب از میانگین غیروزنی، وزنی، وزنی بر اساس مدل ترکیبی کوواریانس کامل گوسین رتبه‌بندی، برای پیش‌بینی رتبه‌بندی کاربر ناشناخته استفاده می‌شوند. ارزیابی‌های تجربی نشان می‌دهد که سه روش پیشنهادی EVMBR، EWVMBR و EWVMBR-G، که از یادگیری جمعی استفاده می‌کند، دقیق‌ترین روش در بین روش‌های ارزیابی شده است. بسته به مجموعه داده، روش پیشنهادی EWVMBR-G موفق به دست‌یابی به بیست تا سی درصد خطای مطلق کمتر از MBR اصلی شده است. از نظر زمان اجرا، روش‌های پیشنهادی قابل مقایسه با MBR و بسیار سریع‌تر از روش slope-one و روش‌های توصیه‌گر KNN مبتنی بر کسینوس یا پیرسون هستند.

واژگان کلیدی: K-نزدیک‌ترین همسایه، رتبه‌بندی، واریانس، سیستم پیشنهاددهنده.

Hybrid Recommender System Based on Variance Item Rating

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

[1,4]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

[3]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

* Corresponding author

*نویسندهٔ عهده‌دار مکاتبات

• تاریخ ارسال مقاله: ۱۴۰۰/۴/۴ • تاریخ پذیرش: ۱۴۰۱/۵/۲ • تاریخ انتشار: ۱۴۰۱/۱۰/۳ • نوع مطالعه: پژوهشی

[DOI: 10.52547/jsdp.19.3.147]

Abstract

K-nearest neighbors (KNN) based recommender systems (KRS) are among the most successful recent available recommender systems. These methods involve in predicting the rating of an item based on the mean of ratings given to similar items, with the similarity defined by considering the mean rating given to each item as its feature. This paper presents a KRS developed by combining the following approaches: (a) Using the mean and variance of item ratings as item features to find similar items in an item-wise KRS (IKRS); (b) Using the mean and variance of user ratings as user features to find similar users with a user-wise KRS (UKRS); (c) Using the weighted mean to integrate the ratings of neighboring users/items; (d) Using ensemble learning. Three proposed methods EVMBR, EWVMBR and EWVMBR-G are presented in this paper. All three methods are user-based, in which VM distance is used as a measure of the difference between users / items, to find neighboring users / items, and then the weighted average is weighted, respectively. Also, weights based on the Gaussian combined covariance model are used to predict unknown user ratings. Our empirical evaluations show that the proposed method EVMBR, EWVMBR and EWVMBR-G, which utilizes ensemble learning, are the most accurate among the methods evaluated. Depending on the dataset, the proposed method EWVMBR-G managed to achieve 20 to 30 percent lower mean absolute error than the original MBR. In terms of runtime, the proposed methods are comparable to the MBR and much faster than the slope-one method and the cosine- or Pearson-based KNN recommenders.

Keywords: K-Nearest Neighbor, Rating, Variance, Recommender System.

۱- مقدمه

قطعاً یکی از اهداف همیشگی و مهم سیستم‌های رایانه‌ای ارائه خدمات مطلوب و ایجاد تجارب لذت بخش برای کاربران هنگام استفاده از سامانه می‌باشد [1]. این هدف منجر به انجام پژوهش‌های متعددی در زمینه‌های مختلف مرتبط با علوم رایانه شده است. گرچه مطلوبیت سامانه از نظر کاربر به اشکال مختلفی نمایان می‌شود، اما آنچه در همه موارد ثابت است، اصل حصول رضایت کاربران است. به‌عنوان مثال، به‌کارگیری روش‌های یادگیری ماشین و تشخیص آماری الگو به همراه روش‌های پردازش صوت در یک سامانه پاسخ‌گویی خودکار می‌تواند توهم ارتباط کاربران با عامل انسانی را فراهم ساخته و با ارائه دسترسی در تمامی ساعات شبانه روز و پاسخ‌گویی بلادرنگ، منجر به کسب رضایت کاربران شود. این روند به شکل چرخه‌ای ظاهر می‌شود که در آن به‌طور مداوم خدمات ارائه‌شده بهبود یافته و به‌تبع آن نیازها و یا توقعات کاربران نیز افزایش می‌یابد [1].

سامانه توصیه‌گر (RS) با تحلیل رفتار کاربر خود، اقدام به توصیه مناسب‌ترین اقلام (داده، اطلاعات، کالا و...) می‌کند. این سامانه رویکردی است که برای مواجهه با مشکلات ناشی از حجم فراوان و رو به رشد اطلاعات ارائه شده است و به کاربر خود کمک می‌کند تا در میان حجم عظیم اطلاعات، سریع‌تر به هدف خود نزدیک شود. سامانه‌های توصیه‌گر قابلیت ارائه توصیه‌های شخصی‌شده را به تک‌تک کاربرانی دارند که از بین حجم بالای اطلاعات به‌دنبال نوعی خاص از اطلاعات مرتبط با اولویت‌هایشان هستند. این سامانه‌ها با توانایی‌ای که در جمع‌آوری اطلاعات از رفتار و حرکات کاربران، دسته‌بندی و تفسیر آن‌ها دارند، امکانی فراهم آورده‌اند که کاربران با صرف زمان و انرژی کمتر به اطلاعات مناسب‌تری دسترسی پیدا کنند و برای شناسایی اولویت‌های کاربر و استفاده از این اطلاعات در ساخت توصیه‌ها نیاز به واسطی هوشمند دارند. درواقع نیاز به سامانه‌ای است که کاربرانش را می‌فهمد، به بیان دیگر تلاش بر این است که سامانه به‌اندازهٔ کافی هوشمند طراحی شود، که بتواند دربارهٔ کاربران خود اطلاعات به‌دست آورد، اطلاعات را دسته‌بندی کند و با استناد به آن‌ها توصیه‌های مناسبی ارائه کند.

سامانه‌های پیشنهاددهنده از مؤثرترین عناصری هستند که با تعبیه در یک سامانه در راستای بهبود تجربه کاربر در بهره بردن از آن سامانه مورد استفاده قرار می‌گیرند [2]. به‌عنوان مثال در حوزه تجارت الکترونیک این سامانه‌ها به کاربران در جهت کوتاه‌کردن زمان جستجو برای خرید کالا با کاهش تعداد گزینه‌های انتخابی (که منطبق بر سلایق آن‌ها باشد) کمک می‌کنند، برای مثال سایت آمازون کالاهایی را بر اساس داده‌های رتبه‌بندی‌شدهٔ پیشین کاربران، به آن‌ها پیشنهاد می‌دهد [2].

از اواسط دهه ۹۰ میلادی و اندکی پس از آن، با ارائه عبارت "سامانه پیشنهاددهنده" توسط رزنیک[1] و واریان[2]، توجه روزافزون پژوهش‌گران به سامانه‌های

[1] Resnick
[2] Varian

[DOI: 10.52547/jsdp.19.3.147]

پیشنهاددهنده معطوف شد و روزبه‌روز بر زمینه‌های کاربرد این سامانه‌ها افزوده شده است [4-2]. درواقع تا چند سال پیش، از دید بیش‌تر کاربران، تنها دسترسی به اطلاعات از طریق وب جهانی اتفاقی آرمانی قلمداد می‌شود و موجب رضایت بیشینه‌ای آن‌ها می‌شد؛ اما امروزه، عرضه مطلوب‌ترین آیتم‌ها که بیشترین قرابت را با روحیات کاربران دارد (در یک تارنمای فروشگاه برخط) یک نیاز ضروری محسوب می‌شود. کاربری که در تارنمایی نظیر آمازون[1] به دنبال کتابی در مورد یک مطلب است، اگر پیشنهادهایی را با بیشترین احتمال پذیرش از طرف وی، دریافت نکند، به‌احتمال زیاد به سمت تارنمایی که قادر به انجام چنین کاری است، سوق داده خواهد شد؛ مسأله‌ای که برای بیش‌تر کسب و کارهای الکترونیکی موفق، استفاده از بهترین سامانه‌های پیشنهاددهنده را به امری حیاتی بدل کرده. درحقیقت، از همین رو است که به‌نوعی سردمداران توسعه و استفاده از سامانه‌های پیشنهاددهنده، پژوهش‌گران و یا سازمان‌های وابسته به تجارت الکترونیک هستند؛ بنابراین سامانه‌های پیشنهاددهنده جزء ارزشمندترین اجزاء به شمار می‌روند که می‌توانند انتظار کاربر و تجربه وی را در هنگام کار با سامانه‌های عمومی[2] (در حالی که درون سامانه تعبیه شده‌اند[3]) به‌طور قابل توجهی افزایش دهند [1].

انتظار می‌رود سامانه‌های توصیه‌گر (RS) دو نوع کار را انجام دهند: پیش‌بینی و توصیه. در وظایف پیش‌بینی، سامانه‌ها براساس اطلاعات موجود مانند سابقه تنظیمات موجود، سعی در پیش‌بینی علاقه کاربر هدف به یک محصول یا آیتم جدید دارند. در این زمینه، کاربر هدف، کاربری است که پیشنهاد جدید برای آن در حال تولید و فرض بر این است که تنظیمات برگزیده منعکس‌کننده علاقه یا رضایت از محصول /کالای مورد نظر است. وظایف توصیه شامل تولید فهرستی از محصولات/ آیتم‌هایی است که به‌احتمال ترجیح کاربر هستند. روش‌های مختلفی برای برنامه‌های پیش‌بینی و توصیه تاکنون ارائه شده که مهم‌ترین اقدامات عملکردی در این زمینه، سرعت و دقت عمل است.

در سامانه‌ای با مقدار زیادی تصمیم‌گیری احتمالی، طراحی زیرسامانه‌ای که بتواند به مردم در تصمیم‌گیری خود کمک کند، اجتناب‌ناپذیر است. چنین سامانه‌هایی که سامانه توصیه‌کننده نامیده می‌شوند، قسمت‌های اساسی تارنماهای مهمی مانند Netflix، Amazon، Pandora، Facebook، YouTube و غیره هستند. انواع مختلفی از سامانه‌های توصیه‌کننده وجود دارد، مانند: سامانه‌های مبتنی بر فیلتر مشارکتی [1]، گراف مبتنی بر دانش [2]، فیلتر مبتنی بر محتوا [3,4] و سامانه‌های ترکیبی توصیه‌کننده [5]. در سامانه‌های پیشنهادی مبتنی بر فیلتر مشارکتی، برای تخمین رتبه‌بندی کاربر بر روی یک آیتم، از رتبه‌بندی سایر کاربران روی آن آیتم و همچنین از رتبه‌بندی آن کاربر روی سایر آیتم‌ها استفاده می‌شود. دو دسته اصلی سامانه‌های پیشنهادی مبتنی بر فیلتر مشارکتی، روش‌های مبتنی بر مدل و مبتنی بر حافظه هستند.

روش‌های مبتنی بر مدل مانند روش تجزیه‌های ماتریسی [6] با بهره‌گیری از الگوریتم‌های یادگیری ماشین و روش‌های داده‌کاوی، مدل‌های پیش‌بینی را ایجاد می‌کنند. این روش‌ها به‌طورمعمول قابل تفسیر نیستند و دارای یک مرحله آموزش بسیار زمان‌بر هستند اما عملکرد خوبی دارند. بنابراین، برخی از پژوهش‌گران سعی می‌کنند؛ زمان آموزش را در روش‌های مبتنی بر مدل زمان آموزش را کاهش دهند [7].

از آن‌جاکه همسایگی ایده اصلی برای پیش‌بینی رتبه‌بندی بدون مشاهده، در روش‌های فیلتر مشارکتی مبتنی بر حافظه است، این روش‌ها به‌عنوان روش‌های همسایه‌محور شناخته می‌شوند [8]. سادگی و توضیح از مزایای روش‌های فیلترکردن مبتنی بر حافظه است؛ اما این روش‌ها از پراکندگی ماتریس رتبه‌بندی رنج می‌برند؛ علاوه‌بر این، پیچیدگی محاسباتی فاز آزمون، چالش اصلی دیگر روش‌های فیلتر مبتنی بر حافظه از هنگام بهره‌برداری است [9].

روش‌های مبتنی بر حافظه (MClFRS) پیش‌بینی‌هایی را بر اساس شباهت کاربران یا آیتم‌ها انجام می‌دهند، اما دقت پیش‌بینی آن‌ها به دلیل پراکندگی ماتریس رتبه‌بندی[4] RM مشکوک است.

سامانه‌های توصیه‌گر مبتنی بر حافظه با m کاربر و n آیتم به‌طورمعمول برای ذخیره اطلاعات رتبه‌بندی به فضای $O(mn)$ نیاز دارند. در الگوریتم‌های فیلتر مشارکتی مبتنی بر آیتم (CF)، بردار ویژگی هر آیتم دارای طول m است و زمان $O(m)$ برای محاسبه شباهت بین دو آیتم با استفاده از فاصله پیرسون یا کسینوس طول می‌کشد. در [5]، یک الگوریتمCF کارآمد بر اساس معیار جدیدی به

[1] Amazon
[2] General Systems
[3] Embedded Inside
[4] Rating Matrix (RM)

[DOI: 10.52547/jsdp.19.3.147]

نام فاصله M(MD) پیشنهاد شده که به‌عنوان تفاوت بین میانگین رتبه‌بندی دو آیتم تعریف می‌شود. در مرحله اولیه‌سازی، میانگین رتبه‌بندی اقلام را محاسبه کرده و آن‌ها را در دو بردار ذخیره می‌کند که به فضای $O(m)$ نیاز دارد. مقاله [5]، فاصله M را تعریف کرده و الگوریتم MBR را پیشنهاد کرده است. فاصله M بسیار ساده و الگوریتم MBR به‌طور قابل توجهی کارآمدتر از روش‌های موجود است؛ بنابراین، برای داده‌های پویا در مقیاس بزرگ مناسب است. در مقایسه با الگوریتم‌های موجود، الگوریتم MBR نیز دقت خوبی از خود نشان داده است.

با توضیحات بالا می‌توان نتیجه گرفت که، یکی از موفق‌ترین انواع توصیه‌کننده‌ها، RSهایی هستند که مبتنی بر فاصله[1] هستند. مقدار MD برای دو آیتم، تفاوت بین میانگین رتبه‌بندی آن‌ها است. هر مقدار مجهول در RM برابر با میانگین رتبه‌بندی گروهی از بیشتر آیتم‌های مشابه فرض می‌شود، که در آن شباهت بر اساس MD تعریف شده است. RSهایی که از MD استفاده می‌کنند بسیار سریع‌تر از بقیه هستند [6]، زیرا شباهت دو آیتم را تنها با تجزیه و تحلیل دو مقدار اسکالر محاسبه می‌کند، در حالی که روش‌های دیگر برای این منظور باید دو بردار را تجزیه و تحلیل کنند. هدف از روش پیشنهادی این مقاله ارائه روشی است که مبتنی بر رویکرد MD است. درواقع روش ارائه‌شده در این مقاله همه مزایای روش مبتنی بر MD را دارد و توانسته با ترکیب سه رویکرد مناسب، معایب این روش را نیز از بین ببرد. نوآوری اساسی این مقاله به‌صورت زیر تشریح شده است:

در این مقاله یک رویکرد مبتنی بر MD با عنوان RSMD ایجادشده با ترکیب رویکردهای زیر ارائه شده است: (الف) استفاده از میانگین و واریانس رتبه‌بندی آیتم به‌عنوان ویژگی‌های آیتم برای تعیین آیتم‌های مشابه با KNN مبتنی بر کالا. (ب) استفاده از میانگین و واریانس رتبه‌بندی کاربر به‌عنوان ویژگی‌های کاربر برای تعیین کاربران مشابه با KNN مبتنی بر کاربر؛ (ج) استفاده از فرمول میانگین وزنی برای تلفیق رتبه‌بندی کاربران/آیتم‌های همسایه. (د) استفاده از یادگیر جمعی.

در ادامه این مقاله، بخش بعدی مقدمه‌ای بر RSها (به‌ویژه مواردی که از MD استفاده می‌کنند) ارائه می‌دهد. بخش ۳ RS پیشنهادی را ارائه می‌دهد. بخش ۴ نتایج ارزیابی‌های انجام‌شده با مجموعه داده‌های واقعی را ارائه می‌دهد و بخش نهایی مقاله شامل نتیجه‌گیری می‌باشد.

[1] M-distance (MD)

۲- ادبیات پژوهش و کارهای مرتبط

از معیارهای مختلفی مانند ضریب همبستگی پیرسون و تشابه کسینوس می‌توان برای تعیین نزدیک‌ترین همسایگان استفاده کرد. K-نزدیک‌ترین سامانه‌های توصیه‌شده همسایه‌محور که از ضریب همبستگی پیرسون و شباهت کسینوس استفاده می‌کنند، به‌ترتیب P-kNN و C-kNN نامیده می‌شوند [10,11]. یک بررسی جامع از اقدامات شباهت برای فیلتر مشارکتی به‌تازگی منتشر شده است [12] که اقدامات مختلف شباهت مانند ضریب همبستگی پیرسون، همبستگی کسینوس، ضریب همبستگی پیرسون محدود [13]، ضریب همبستگی پیرسون مبتنی بر تابع سیگموئید [13] و معیار فاصله مینکوفسکی را مقایسه می‌کند.

فیلترهای مشارکتی مبتنی بر کاربر (UBClFRS[2]) و فیلترهای مشارکتی مبتنی بر کالا (IBClFRS[3])، دو روش ساده پیش‌بینی و توصیه هستند که می‌توانند دقت قابل قبولی را ارائه دهند. IBClFRS در ساده‌ترین شکل، رتبه‌بندی یک کالا را بر اساس میانگین رتبه‌بندی k آیتم با بیشترین رتبه‌بندی مشابه پیش‌بینی می‌کند، که این موضوع می‌تواند یک فرآیند زمان‌بر باشد [14]. در این روش، تشابه دو آیتم با معیارهایی مانند تشابه کسینوس یا همبستگی پیرسون اندازه‌گیری می‌شود. در [15]، یک روش اندازه‌گیری شباهت جدید برای بهبود صحت توصیه در حل مشکل راه‌اندازی کاربر جدید (NUCS[4]) ارائه شد. در این کار، به‌طور تجربی نشان داده شد که اندازه‌گیری پیشنهادی از سایر اقدامات مشابهت مانند هم کسینوس و همبستگی پیرسون بهتر عمل می‌کند. در [16]، روش جدیدی به‌نام فیلترکردن مشارکتی معکوس (RClFRS[5]) پیشنهاد شد. RClFRS شامل یافتن نزدیک‌ترین k همسایگان اقلام دارای رتبه‌بندی و سپس پیش‌بینی رتبه‌بندی کالای بدون رتبه‌بندی بر اساس مشابه‌ترین آیتم‌های موجود در مرحله نخست است. برای بهبود دقت، این الگوریتم رتبه‌بندی‌های نادرست را شناسایی و آن‌ها را از پیش‌بینی خارج می‌کند. این روش بسیار سریع‌تر از RClFRS فیلترکننده مشترک معمولی است زیرا همیشه آیتم دارای امتیاز کمتر از آیتم‌های بدون رتبه است.

[2] User-based Collaborative Filtering RS
[3] Item-based Collaborative Filtering RS
[4] New User Cold-Starting
[5] Reversed Collaborative Filtering RS

[DOI: 10.52547/jsdp.19.3.147]

روش‌های ClFRS مبتنی بر حافظه (MClFRS[1]) بر اساس شباهت کاربران یا آیتم‌ها پیش‌بینی می‌کنند، اما به‌دلیل پراکنده‌بودن ماتریس رتبه‌بندی (RM)، صحت پیش‌بینی آن‌ها مشکوک است. در [17]، این تعریف با تعریف رتبه نهایی به‌عنوان ترکیبی وزنی از سه پیش‌بینی بر اساس رتبه‌بندی کاربران مشابه، رتبه‌بندی آیتم‌های مشابه و رتبه‌بندی آیتم‌های مشابه توسط کاربران مشابه، بهبود یافت.

در [18]، معیار جدیدی به‌نام MD[2] برای اندازه‌گیری شباهت آیتم‌ها ارائه شد. مقدار MD برای دو آیتم تفاوت بین میانگین رتبه‌بندی آن‌ها است. فرض می‌شود که هر مقدار ناشناخته در RM برابر با میانگین امتیاز گروهی از آیتم‌های مشابه است، که در آن شباهت بر اساس MD تعریف شده است. RSهایی که از MD استفاده می‌کنند، بسیار سریع‌تر از بقیه هستند [19]، زیرا با تجزیه و تحلیل تنها دو مقدار مقیاس، شباهت دو آیتم را محاسبه می‌کند، درحالی‌که سایر روش‌ها برای این منظور باید دو بردار را تجزیه و تحلیل کنند.

SORS شامل استفاده از یک رگرسیون خطی مستقیم برای پیش‌بینی مقادیر ناشناخته در RM است. در [20]، فرم گسترده‌ای از این روش ارائه شد. این الگوریتم ابتدا برای هر کاربر چندین نزدیک‌ترین همسایه را پیدا می‌کند که آستانه شباهت به آن کاربر را دارند (کاربران مختلف می‌توانند تعداد همسایگان متفاوتی داشته باشند)، سپس اقلام رتبه‌بندی‌نشده هر کاربر را فقط بر اساس نزدیک‌ترین همسایگان مشخص‌شده در مرحله قبلی (برخلاف کل داده‌ها) رتبه‌بندی می‌کند. نسخه‌های افزایشی SORSs[3] [21] برای برنامه‌هایی که اطلاعات کاربر یا آیتم به‌طور دائمی تکمیل می‌شوند، مناسب‌تر هستند.

در روشی که تحت عنوان MCOCR[4] شناخته می‌شود، ابتدا کاربرانی که علایق مشترک و آیتم‌های مورد علاقه خود را دارند در همان خوشه قرار می‌گیرند؛ سپس یک ClFRS به‌صورت تکراری روی کاربران و آیتم‌های هر خوشه انجام می‌شود [22].

برخلاف ClFRS، که هر آیتم بر اساس امتیازات داده‌شده توسط کاربران دیگر رتبه‌بندی می‌شود، فیلترکردن محتوای مبتنی بر محتوا (CBFRS[5]) شامل رتبه‌بندی آیتم‌ها بر اساس ویژگی‌های آن‌ها و میزان علاقه کاربران به این ویژگی‌ها است [23]. مزیت CBFRS توانایی آن‌ها در پیش‌بینی میزان علاقه کاربر به یک آیتم بدون نیاز به جمع‌آوری امتیازات داده‌شده توسط سایر کاربران به آن آیتم است. در RS ارائه شده در [24]، از CBFRS برای افزایش ClFRS استفاده می‌شود. به‌طور دقیق‌تر، این RS از امتیازات داده‌شده توسط کاربران مشابه و همچنین علاقه کاربر به ویژگی‌های آیتم استفاده می‌کند که بر اساس رتبه‌بندی داده‌شده توسط آن کاربر به سایر آیتم‌ها تعیین می‌شود.

در بیشتر RSها، پیش‌بینی این‌که آیا کاربر به آیتم یا پیشنهادی علاقه‌مند است یا خیر، می‌تواند به‌عنوان یک مسئله طبقه‌بندی جفت (PWCP) یا مسأله تصمیم‌گیری دوبه‌دو (PWDMP[6]) مدل شود. با این حال، استفاده از مسأله تصمیم‌گیری سه‌گانه TWDMP[7] [25-27] به ما امکان می‌دهد یک انتخاب سوم را خارج از تصمیم دودویی توصیه یا عدم توصیه یک آیتم به کاربر خاص در نظر بگیریم. به‌عبارت دیگر، این روش فقط درصورتی‌که اطمینان زیادی از علاقه کاربر داشته باشد، یک آیتم را به کاربر توصیه می‌کند. در غیراین‌صورت اگر اطمینان زیادی از علاقه کاربر نداشته باشد، آن آیتم را توصیه نمی‌کند. در غیراین‌صورت، گزینه سوم می‌تواند درخواست نظارت بیشتر از کاربر باشد [28].

در فاکتورهای ماتریسی (MFRS[8]) [29]، ویژگی‌های نهفته کاربران و آیتم‌ها به‌گونه‌ای تعیین می‌شود که محصول داخلی بردار ویژگی پنهان کاربر و بردار ویژگی نهان یک آیتم، با امتیاز داده شده کاربر به آن آیتم برابر باشد. در [30]، پس از استخراج ویژگی‌های نهفته به‌دست آمده از ماتریس (MF)، از یک ماشین بردار پشتیبان (SVM) مدل [31] برای طبقه‌بندی مورد/ کاربر

[1] Memory-based ClFRS
[2] M-distance
[3] The Slope-One RSs
[4] Multiclass CoClustering RS

[5] Content-based Filtering RS
[6] Pairwise Decision-Making Problem
[7] Triple-Wise Decision-Making Problem
[8] Matrix Factorization RSs

[DOI: 10.52547/jsdp.19.3.147]

به دو دسته مورد نظر و غیر ارجح استفاده شد. در [32]، یک شبکه عمیق کانولوشن برای افزایش توصیه‌های MFRS استفاده شد. نقش شبکه کانولوشن در این MFRS تسهیل استخراج ویژگی‌های نهفته اقلام بر اساس نظرات نوشته‌شده توسط کارشناسان یا مشتریان در مورد آن‌ها است.

یکی از موفق‌ترین انواع توصیه‌کنندگان، روش RSMD است. در این مقاله RSMD ایجاد شده با ترکیب رویکردهای زیر ارائه شده است: (الف) استفاده از میانگین و واریانس رتبه‌بندی آیتم به‌عنوان ویژگی‌های آیتم برای تعیین آیتم‌های مشابه با KNN مبتنی بر کالا. (ب) استفاده از میانگین و واریانس رتبه‌بندی کاربر به‌عنوان ویژگی‌های کاربر برای تعیین کاربران مشابه با KNN مبتنی بر کاربر؛ (ج) استفاده از فرمول میانگین وزنی برای تلفیق رتبه‌بندی کاربران/آیتم‌های همسایه. (د) استفاده از یادگیر جمعی. ارزیابی‌های انجام شده با سه مجموعه داده واقعی نشان می‌دهد که در میان روش‌های پیشنهادی، EVMBR و EWVMBR، که از یادگیری جمعی استفاده می‌کنند، دقیق‌ترین هستند. دقیق‌ترین روش پیشنهادی، EWVMBR، بیست تا سی درصد میانگین خطای مطلق کمتری نسبت به سیستم MBR دارد. از نظر زمان اجرا، روش‌های پیشنهادی قابل مقایسه با MBR هستند و بسیار سریع‌تر از تکنیک slope-one و توصیه‌کنندگان KNN مبتنی بر کسینوس و پیرسون هستند.

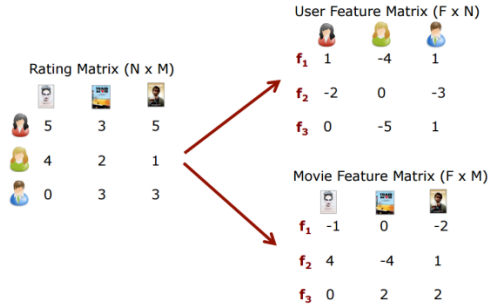
۳- روش‌های پیشنهادی

در این مقاله هشت سامانه پیشنهادی مبتنی بر MBR ارائه شده است. این سامانه‌ها با استفاده از ترکیبات مختلف و رویکردهای زیر توسعه یافته‌اند:

- استفاده از میانگین و واریانس رتبه‌بندی آیتم‌ها به‌عنوان ویژگی‌های آیتم برای یافتن آیتم‌های مشابه در رویکرد نزدیکترین آیتم‌های همسایه
- استفاده از میانگین و واریانس رتبه‌بندی کاربران به‌عنوان ویژگی‌های کاربر برای یافتن کاربران مشابه در رویکرد نزدیکترین کاربران همسایه
- استفاده از فرمول میانگین وزنی برای تلفیق رتبه‌بندی کاربران/ آیتم‌های همسایه
- استفاده از یادگیر جمعی

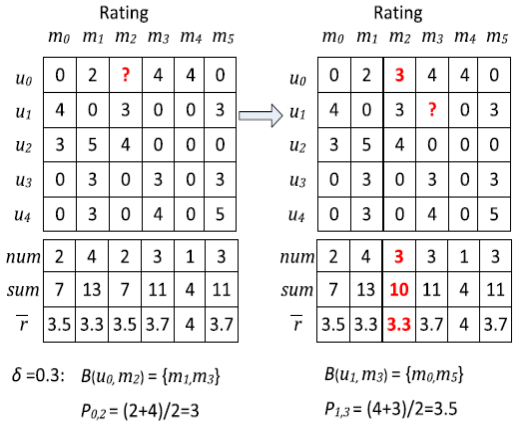
۳-۱- نمایش ماتریس

در فاکتورسازی ماتریس، کاربران و آیتم‌ها هر دو به‌عنوان بردارهای پنهان در یک فضای k بُعدی مشترک، $\mathbb{R}^k$ نشان داده می‌شوند، به‌طوری‌که کاربر i به‌عنوان یک بردار پنهان $u_i \in \mathbb{R}^k$ و آیتم j به‌عنوان یک بردار پنهان $v_j \in \mathbb{R}^k$ نشان داده می‌شود. پیش‌بینی این‌که آیا کاربر i آیتم j را دوست دارد یا خیر، با حاصل‌ضرب درونی بین بازنمایی‌های پنهان آن‌ها، $r_{ij} = u_i^T v_j$ ارائه می‌شود. برای استفاده از فاکتورسازی ماتریس برای فیلترکردن مشارکتی، نمایش‌های نهفته کاربران و آیتم‌ها باید با یک ماتریس مشاهده‌شده از رتبه‌بندی‌ها آموخته شوند. شکل‌های (۱ و ۲) نمایش فاکتورسازی ماتریس را برای مجموعه داده MovieLens100k نشان می‌دهند.



(شکل-۱): نمایش فاکتورسازی ماتریس برای مجموعه‌داده MovieLens100k

(Figure-1): Demonstration of matrix factorization for the MovieLens100k dataset



(شکل-۲): نمایش فاکتورسازی ماتریس در روش MBR برای مجموعه‌داده MovieLens100k

(Figure-2): Demonstration of matrix factorization in the MBR method for the MovieLens100k dataset

۳-۲- روش MBR

RSهایی که از MD استفاده می‌کنند بر اساس KNN هستند که آیتم‌های مشابه با استفاده از میانگین امتیاز داده‌شده به هر آیتم به‌عنوان ویژگی آن شناسایی می‌شوند [33,34]. در این روش، میانگین امتیاز آیتم x با معادله (۱) تعیین می‌شود:

$$\bar{R}_{:x} = \frac{\sum_{i=1}^{u} \delta(R_{ix})}{\sum_{i=1}^{u} \pi(R_{ix} \neq Unknown)} \qquad (۱)$$

[DOI: 10.52547/jsdp.19.3.147]

که در آن u تعداد کاربران در جدول آیتم/کاربر است (به‌عنوان مثال RM) R_{ix} امتیاز داده‌شده توسط کاربر i به آیتم x است (اگر کاربر i به آیتم x هیچ رتبه‌بندی ندهد R_{ix} مشخص نیست)، $\bar{R}_{:x}$ میانگین امتیاز داده‌شده به آیتم x است، $\delta(y)$ در معادله (۲) و $\pi(y)$ در معادله (۳) تعریف می‌شود:

$$\delta(y) = \begin{cases} y & y \neq Unknown \\ 0 & o.w. \end{cases} \quad (۲)$$

$$\pi(y) = \begin{cases} 1 & y = true \\ 0 & o.w. \end{cases} \quad (۳)$$

MD برای آیتم‌های x_1 و x_2 مطابق با معادله (۴) محاسبه می‌شود.

$$MD(x_1, x_2) = \left|\bar{R}_{:x_1} - \bar{R}_{:x_2}\right| \quad (۴)$$

شبه‌کد RSMD در الگوریتم (۱) ارائه شده است. برای پیش‌بینی رتبه‌ای که کاربر i قرار است به آیتم x بدهد (به‌عنوان مثال R_{ix})، این الگوریتم مشابه‌ترین آیتم همسایه x را که توسط کاربر i رتبه‌بندی شده است جستجو می‌کند. اگر حداقل یک همسایه از این نمونه وجود داشته باشد، R_{ix} به‌عنوان میانگین رتبه‌بندی آیتم همسایه توسط کاربر i تنظیم می‌شود. در غیراین‌صورت، $\bar{R}_{:x}$ به‌عنوان میانگین امتیاز داده‌شده توسط کاربران مختلف به آیتم‌های متفاوت داده می‌شود. به‌عبارت دیگر:

$$\hat{R}_{ix} = \begin{cases} \frac{\sum_k \pi(md_{ik} \leq \delta)\, R_{ix}}{\sum_k \pi(md_{ik} \leq \delta)} & \sum_k \pi(md_{ik} \leq \delta) \geq 1 \\ \bar{R}_{:x} & o.w. \end{cases} \quad (۵)$$

$\hat{R}_{ix}$ رتبه‌بندی پیش‌بینی‌شده است که کاربر i قرار است به آیتم x بدهد.

Algorithm 1: Pseudo Code of RSMD [5]
Inputs:
δ: Neighborhood threshold
s: User-item table
m: Number of items
u: Target user index
i: Target item index
Outputs:
p_{ui}: Predicted rating of item i by user u

/* **Step 1:** Determine the neighbors of item i */
$nb = 0$ // Number of neighbors of item i
$nbsum = 0$ // Sum of ratings given to the neighbors of item i
for $k = 1$ to m
/ item i cannot be a neighbor of itself /
if $(k == i)$ then
continue
endif
/any item that is rated 0 cannot be a neighbor/
if $(s_{uk} == 0)$ then
continue;
endif
$md_{ik} = abs(\mu_i - \mu_k)$
if $(md_{ik} \leq \delta)$ then
$nb = nb + 1$
$nbsum = nbsum + s_{uk}$
$endif$
$endfor$

/* **Step 2:** Predict the rating of item i */
$if\ (nb \geq 1)\ then$

$p_{ui} = \frac{nbsum}{nb}$
else
$p_{ui} = \mu_i$
$endif$
end

(شکل-۳): شبه کد الگوریتم RSMD
(Figure-3): RSMD algorithm pseudo-code

۳-۳- روش‌های مبتنی بر آیتم

در RSMD، میانگین امتیازات داده‌شده به هر آیتم به‌عنوان ویژگی آن آیتم برای تعیین آیتم‌های مشابه استفاده می‌شود.

در ادامه، ما چندین روش مبتنی بر آیتم را ارائه می‌دهیم که بر اساس این روش بنا شده است.

۱) استفاده از واریانس

واریانس رتبه‌بندی آیتم u از معادله زیر به‌دست آورده می‌شود:

$$v_i = \frac{\sum_{u=1}^{n}(s_{ui} - \mu_i)^2}{|\{s_{ui} \mid s_{ui} \neq 0, 1 \leq u \leq n\}| - 1} \quad (۶)$$

که در آن s_{ui} امتیاز داده شده توسط کاربر u به آیتم i است. همان‌طور که مشاهده می‌شود، فرمول میانگین رتبه نامشخصی را که مقدار صفر به آن‌ها داده می‌شود، نادیده می‌گیرد. در اینجا μ_i، میانگین رتبه‌ای است که توسط کاربران مختلف به آیتم i داده می‌شود.

تعریف: "فاصله VM" از آیتم‌های i و k به این صورت تعریف می‌شود:

$$vmd_{i,k} = \text{abs}(\mu_i - \mu_k) + \alpha \times abs(v_i - v_k) \quad (۷)$$

که $\alpha \geq 0$ وزن اهمیت واریانس نسبت به میانگین ظرفیت آن‌ها به‌عنوان ویژگی‌های آیتم است.

[DOI: 10.52547/jsdp.19.3.147]

۲) رتبه‌بندی وزنی همسایگان

در MBR، رتبه کاربر ناشناخته بر اساس میانگین غیروزنی رتبه‌بندی چندین آیتم همسایه تعیین می‌شود. برای افزایش دقت MBR، پیشنهاد می‌شود، این میانگین غیروزنی با یک میانگین وزنی جایگزین شود، یا به‌طور دقیق‌تر، پیش‌بینی رتبه آیتم i توسط کاربر u با معادله زیر محاسبه شود:

$$p_{ui} = \begin{cases} \frac{\sum_{\{k|vmd_{i,k} \le \delta\}} w_k \mu_k}{\sum_{\{k|vmd_{i,k} \le \delta\}} w_k} & |\{k|vmd_{i,k} \le \delta\}| > 0, \\ \mu_i & otherwise, \end{cases} \quad (۸)$$

که در آن $w_k > 0$ وزن گوسی آیتم k است، که توسط رابطه (۹) محاسبه می‌شود:

$$w_k = \exp\left(\frac{-vmd_{i,k}}{\sigma^2}\right) \quad (۹)$$

پارامتر $\sigma > 0$ عرض تابع گوسی است. هرچه فاصله VM بین آیتم‌های i و k بیشتر باشد $(vmd_{i,k})$، وزن آیتم k (w_k) در پیش‌بینی رتبه‌بندی آیتم i بیشتر است. پارامتر σ کنترل می‌کند که سرعت سهم وزنی w_k با افزایش فاصله VM چقدر سریع کاهش می‌یابد.

۳-۴- روش‌های مبتنی بر کاربر

اگرچه MBR یک روش مبتنی بر آیتم است، اما می‌تواند به‌عنوان مبانی روش‌های مبتنی بر کاربر همان‌طور که در ادامه توضیح داده شده است، نیز استفاده شود.

۱) استفاده از واریانس

در اینجا، واریانس رتبه‌بندی آیتم i به‌شرح زیر تعیین می‌شود:

$$\tilde{\mu}_u = \frac{\sum_{i=1}^{m} s_{ui}}{|\{s_{ui} \mid s_{ui} \neq 0, 1 \le i \le m\}|} \quad (۱۰)$$

$$\tilde{v}_u = \frac{\sum_{i=1}^{m} (s_{ui} - \tilde{\mu}_u)^2}{|\{s_{ui} \mid s_{ui} \neq 0, 1 \le i \le m\}| - 1} \quad (۱۱)$$

که در آن m تعداد آیتم‌های موجود در جدول آیتم/کاربر است.

تعریف: "فاصله VM" از آیتم‌های u و t به این صورت تعریف می‌شود:

$$\widetilde{vmd}_{u,t} = \text{abs}(\tilde{\mu}_u - \tilde{\mu}_t) + \alpha \times abs(\tilde{v}_u - \tilde{v}_t) \quad (۱۲)$$

که $\alpha \ge 0$، وزن اهمیت واریانس نسبت به میانگین ظرفیت آن‌ها به‌عنوان ویژگی‌های کاربر است.

۲) بهبود رتبه‌بندی وزنی همسایگان

برای بهبود دقت MBR، همانند قبل، میانگین غیروزنی با میانگین وزنی جایگزین می‌شود. در اینجا، رتبه‌بندی آیتم i توسط کاربر u با معادله زیر پیش‌بینی شده است:

$$p_{ui} = \begin{cases} \frac{\sum_{\{t|\widetilde{vmd}_{u,t} \le \delta\}} \tilde{w}_t \tilde{\mu}_t}{\sum_{\{t|\widetilde{vmd}_{u,t} \le \delta\}} \tilde{w}_t} & |\{t|\widetilde{vmd}_{u,t} \le \delta\}| > 0, \\ \tilde{\mu}_u & otherwise, \end{cases} \quad (۱۳)$$

که $\tilde{w}_t > 0$ وزن گوسی کاربر t است، که:

$$\tilde{w}_t = \exp\left(\frac{-\widetilde{vmd}_{u,t}}{\sigma^2}\right) \quad (۱۴)$$

۳-۵- تلفیق رتبه‌بندی‌های پیش‌بینی‌شده با روش‌های مبتنی بر آیتم و کاربر

خروجی‌های یک سامانه پیش‌بینی را می‌توان با کمک یادگیر جمعی بهبود داد. این روش شامل استفاده از چند پیش‌بینی برای به‌دست‌آوردن آرایه‌ای از نتایج و سپس ادغام این نتایج در پیش‌بینی نهایی است. در این مقاله، روش‌های مبتنی بر آیتم و مبتنی بر کاربر که در بخش‌های قبلی شرح داده شده‌اند، به‌عنوان مؤلفه‌های یادگیر جمعی استفاده می‌شوند. در این روش، رتبه نهایی میانگین وزنی رتبه‌بندی‌های تخمین زده‌شده توسط هر روش است. در اینجا، وزن احتمال دقت پیش‌بینی‌کننده است که با استفاده از داده‌های ارزیابی تخمین زده می‌شود.

در پایان، می‌توان هشت سیستم پیشنهادی در این مقاله را به‌صورت زیر خلاصه کرد:

- VMBR-I: روشی مبتنی بر آیتم، که شامل استفاده از فاصله VM به‌عنوان معیار تفاوت بین آیتم‌ها برای یافتن آیتم‌های همسایه و سپس استفاده از میانگین غیروزنی رتبه‌بندی این آیتم‌ها برای پیش‌بینی کاربر ناشناس است.
- VMBR-U: روشی مبتنی بر کاربر، که شامل استفاده از فاصله VM به‌عنوان معیار تفاوت بین کاربران برای یافتن کاربران همسایه و سپس استفاده از میانگین غیروزنی رتبه‌بندی این کاربران برای پیش‌بینی کاربر ناشناس است.
- WVMBR-I یا روش وزنی VMBR-I: روشی مبتنی بر آیتم که در آن از فاصله VM به‌عنوان معیار تفاوت بین

[DOI: 10.52547/jsdp.19.3.147]

آیتم‌ها برای یافتن آیتم همسایه استفاده می‌شود و سپس از میانگین وزنی رتبه‌بندی این آیتم‌ها استفاده می‌شود تا رتبه‌بندی کاربر ناشناخته را پیش‌بینی کند.

- WVMBR-U یا روش وزنی VMBR-U: روشی مبتنی بر کاربر که در آن از فاصله VM به‌عنوان معیار تفاوت بین کاربران برای یافتن کاربران همسایه استفاده می‌شود و سپس از میانگین وزنی رتبه‌بندی این کاربران برای پیش‌بینی رتبه‌بندی کاربر ناشناخته استفاده می‌کند.
- WMBR-I یا MBR-I وزن‌دار: روشی مبتنی بر اقلام، که شامل استفاده از فاصله M به‌عنوان معیار تفاوت بین آیتم‌ها برای یافتن آیتم همسایه و سپس استفاده از میانگین وزنی رتبه‌بندی این آیتم‌ها برای پیش‌بینی است. رتبه‌بندی کاربر ناشناخته از آنجا که MBR خود یک روش مبتنی بر آیتم است، در اینجا، MBR-I به MBR اصلی اشاره دارد.
- WMBR-U یا MBR-U وزنی: روشی مبتنی بر کاربر، که شامل استفاده از فاصله M به‌عنوان معیار تفاوت کاربر برای یافتن کاربر همسایه و سپس استفاده از میانگین وزنی رتبه‌بندی این کاربران برای پیش‌بینی است.
- EVMBR یا VMBR جمعی: روشی مبتنی بر کاربر، که در آن از فاصله VM به‌عنوان معیار تفاوت بین کاربران/آیتم‌ها برای یافتن کاربران/آیتم‌های همسایه استفاده شده و سپس میانگین غیروزنی رتبه‌بندی این کاربران برای پیش‌بینی رتبه‌بندی کاربر ناشناخته استفاده می‌شود.
- EWVMBR یا WVMBR جمعی: روشی مبتنی بر کاربر، که در آن از فاصله VM به‌عنوان معیار تفاوت بین کاربران/آیتم‌ها، برای یافتن کاربران/آیتم‌های همسایه استفاده می‌شود و سپس میانگین وزنی رتبه‌بندی این کاربران/آیتم‌ها، برای پیش‌بینی رتبه‌بندی کاربر ناشناخته استفاده می‌شود.

۳-۶- روش جمعی EWVMBR-G وزنی با مدل ترکیبی کوواریانس کامل گوسین

مدل آمیخته گاوسی به ترکیب خطی توابع توزیع گاوسی متعدد اشاره دارد. GMM[1] می‌تواند متناسب با هر نوع توزیعی باشد که به‌طورمعمول برای حل مواردی که داده‌های یک مجموعه شامل چندین توزیع مختلف است، مورد استفاده قرار می‌گیرد.

- EWVMBR-G جمعی: روشی مبتنی بر کاربر، که در آن از فاصله VM به‌عنوان معیار تفاوت بین کاربران/آیتم‌ها، برای یافتن کاربران/آیتم‌های همسایه استفاده می‌شود و سپس میانگین وزنی رتبه‌بندی این کاربران/آیتم‌ها بر اساس مدل ترکیبی کوواریانس کامل گوسین، برای پیش‌بینی رتبه‌بندی کاربر ناشناخته استفاده می‌شود. مدل ترکیبی کوواریانس کامل گوسین، یک مدل بسیار منعطف است که به این صورت تعریف می‌شود:

$$p(x) = \frac{1}{\sqrt{2\pi\tau}} e^{-0.5(\frac{x-\mu}{\tilde{v}_u})^2} = N(\mu - \tilde{v}_u{}^2) \qquad (۱۵)$$

شکل توسعه‌یافته این رابطه برای بردار تصادفی مانند x به‌صورت زیر است:

$$p(x) = \frac{1}{\sqrt[D/2]{2\pi}|\Sigma|^{1/2}} e^{-0.5(x-\tilde{v}_u)^T \Sigma^{-1}(x-\mu)} \qquad (۱۶)$$

که در این رابطه x بردار تصادفی، μ بردار میانگین متغیرهای تصادفی و Σ ماتریس کوواریانس متغیرهای تصادفی است.

۴- آزمایش‌ها و نتایج تجربی

در این بخش نتایج ارزیابی روش‌های پیشنهادی و مقایسه‌های انجام شده با MBR [18]، نزدیک‌ترین همسایگان (P-kNN)، نزدیک‌ترین همسایگان مبتنی بر کسینوس (C-kNN) [19]، و روش‌های slope-one [20] ارائه شده است. مجموعه داده‌های واقعی مورد استفاده برای این ارزیابی، مجموعه داده MovieLens100k با ۹۴۳ کاربر و ۱۶۸۲ فیلم، مجموعه داده DouBan با ۲۹۶۵ کاربر و ۳۹۶۹۵ فیلم و مجموعه داده EachMovie با ۷۲۹۱۶ کاربر و ۱۶۲۸ فیلم است. ارزیابی‌ها با استفاده از یک رایانه شخصی با پردازنده ۳/۱ گیگاهرتز، ۱۲ گیگابایت حافظه و سیستم عامل ویندوز ۱۰ انجام شده است.

الگوریتم‌های پیشنهادی با دیگر روش‌های رقیب ارزیابی شده‌اند. معیارهای عملکرد در پیش‌بینی رتبه‌بندی، میانگین خطای مطلق (MAE) و خطای میانگین مربعات (RMSE) هستند. در جدول (۱)، MAE روش‌های پیشنهادی برای بهترین مقدار آستانه (δ) با روش‌های دیگر مقایسه شده است. اعداد پررنگ بهترین نتایج به‌دست‌آمده برای هر مجموعه داده است.

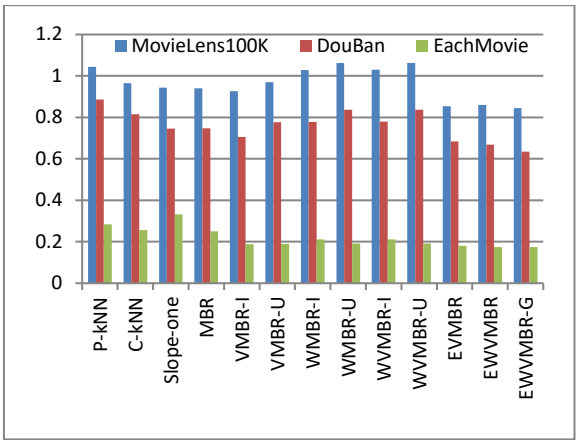
همان‌طور که در جدول (۱) نشان داده شده است، برای هر سه مجموعه‌داده، روش EWVMBR-G کمترین میزان MAE را در بین روش‌های پیشنهادی دارد.

[1] Gaussian Mixture Model

[DOI: 10.52547/jsdp.19.3.147]

همچنین، برای هر سه مجموعه‌داده، همه روش‌های پیشنهادی دارای خطای کمتری نسبت به روش‌های MBR، P-kNN، C-kNN و slope-one هستند. بسته به مجموعه‌داده‌ها، روش EWVMBR توانسته بیست تا سی درصد MAE پایین‌تر از MBR به‌دست آورد.

RMSE روش‌های پیشنهادی برای بهترین مقدار δ در شکل (۴) نشان داده شده است. همان‌طور که مشاهده می‌شود، برای مجموعه‌های داده، روش‌های EVMBR، EWVMBR و EWVMBR-G دارای کمترین RMSE در بین روش‌های پیشنهادی هستند. برای هر سه مجموعه‌داده، این روش‌ها همراه با VMBR-I دارای خطای کمتری نسبت به روش‌های MBR، P-kNN، C-kNN و slope-one هستند. بسته به مجموعه داده‌ها، روش‌های EVMBR، EWVMBR و EWVMBR-G ده تا بیست درصد RMSE پایین‌تر از MBR داشته‌اند.



(شکل-۴): ارزیابی مقدار RMSE روش‌های مختلف

(Figure-4): RMSE of the evaluated methods

(جدول-۱): ارزیابی مقدار MAE روش‌های مختلف

(Table-1): MAE of the evaluated methods

Method	MovieLens100K	DouBan	EachMovie
P-kNN	0.8363	0.7089	0.2277
C-kNN	0.7487	0.6366	0.1980
Slope-one	0.7421	0.5902	0.2900
[35]	0.6214	0.6137	0.1477
[36]	0.6971	0.5740	0.1398
[37]	0.7741	-	-
[38]	0.6733	-	-
MBR	0.7389	0.5869	0.1933
VMBR-I	0.6702	0.5296	0.1410
VMBR-U	0.7062	0.5794	0.1503
WMBR-I	0.6175	0.4673	0.1294
WMBR-U	0.6572	0.5119	0.1437
WVMBR-I	0.6165	0.4637	0.1294
WVMBR-U	0.6565	0.5116	0.1435
EVMBR	0.6519	0.5310	0.1415
EWVMBR	0.5970	0.4636	0.1292
EWVMBR-G	**0.5893**	**0.4612**	**0.1258**

جدول (۲) زمان اجرای روش‌های ارزیابی‌شده را با بهترین مقدار در نظر گرفته‌شده برای پارامتر δ نشان می‌دهد. همان‌طور که در این جدول به‌وضوح نشان داده شده است، برای هر سه مجموعه‌داده، VMBR-I یا VMBR-U کوتاه‌ترین زمان اجرا را داشته‌اند. روش مبتنی بر کاربر VMBR-U کمترین زمان اجرا را برای مجموعه‌داده‌های MovieLens100K و DouBan داشته و روش مبتنی بر آیتم VMBR-I کوتاه‌ترین زمان اجرا را برای مجموعه‌داده EachMovie نشان داده است. این تفاوت در زمان اجرا را می‌توان به تفاوت این مجموعه داده‌ها از نظر نسبت تعداد کاربران به تعداد آیتم‌ها نسبت داد. در مجموعه داده‌های MovieLens100K و DouBan، کاربران کمتر از آیتم‌ها هستند، اما در مجموعه‌داده EachMovie ، تعداد کاربران بیشتر از آیتم‌های موجود است. از این‌رو، برای مجموعه داده EachMovie (با ۷۲۹۱۶ کاربر و ۱۶۲۸ آیتم)، یافتن نزدیک‌ترین آیتم‌های همسایه بسیار سریع‌تر از نزدیک‌ترین کاربران همسایه است، زیرا برای تعیین نزدیک‌ترین همسایگان با VMBR-I، فاصله VM بین آیتم هدف از سایر آیتم‌ها باید محاسبه شود، اما برای VMBR-U، فاصله VM بین کاربر هدف و سایر کاربران باید تعیین شود. تنها تفاوت بین MBR و VMBR-I استفاده از واریانس رتبه‌بندی آیتم‌ها علاوه‌بر میانگین آن‌ها در روش اخیر است. بنابراین، انتظار می‌رود که VMBR-I مدت زمان بیشتری نسبت به MBR داشته

[DOI: 10.52547/jsdp.19.3.147]

باشد، اما جدول (۲) عکس این را نشان می‌دهد، زیرا در روش VMBR-I از مقدار مطلوب δ حین اجرا استفاده شده است. مقدار مطلوب δ برای VMBR-I، 0.001 بود اما برای MBR، 0.002 بود. بنابراین، آیتم همسایه کمتری در VMBR-I نسبت به MBR در پیش‌بینی شرکت داشته‌اند، که منجر به زمان اجرای کوتاه‌تر VMBR-I نسبت به MBR می‌شود.

(جدول-۲): زمان اجرای روش‌های ارزیابی شده (ثانیه)

(Table-2): Runtime of the evaluated methods (in seconds)

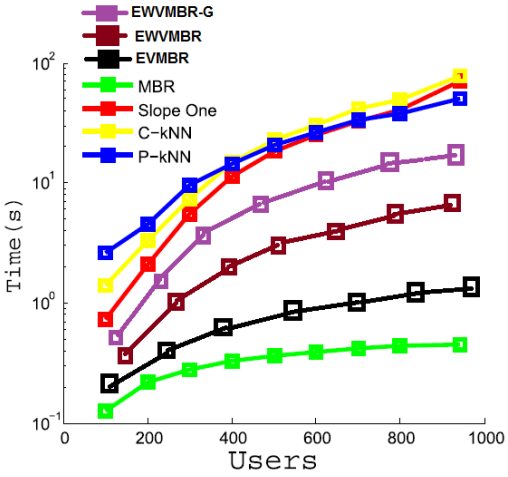
Method	MovieLens100K	DouBan	EachMovie
P-kNN	410.2365	120765.6565	21087.7639
C-kNN	399.6544	118654.7629	20546.6532
Slope-one	397.2097	117233.7626	20351.6334
[35]	4.3344	9765.2024	3232.5632
[36]	4.8934	9234.3250	8113.7501
[37]	3.3250	-	-
[38]	4.5634	-	-
MBR	2.3885	6543.0187	1072.4803
VMBR-I	0.5260	4040.3133	**106.7081**
VMBR-U	**0.2961**	**66.5703**	584.4321
WMBR-I	3.1222	8569.7917	1271.8774
WMBR-U	3.0339	835.3166	7145.0752
WVMBR-I	2.9670	4755.3064	1232.4631
WVMBR-U	3.2888	814.72806	7113.4829
EVMBR	0.6877	4093.5993	709.9855
EWVMBR	6.5287	9444.0830	8374.5360
EWVMBR-G	6.8901	9546.0932	8560.5500

شکل (۵) حساسیت MAE از WVMBR-I به مقدار δ را برای مجموعه‌داده Movielens100k نشان می‌دهد. در این ارزیابی، δ به 10^{-4} تنظیم شد. مشاهده می‌شود که با افزایش δ، MAE کاهش می‌یابد. با این حال، مقادیر δ بالاتر نیز مربوط به زمان‌های طولانی‌تر است (شکل ۶)، زیرا منجر به مشارکت بیشتر همسایگان در پیش‌بینی می‌شود. در روش‌های غیروزنی مانند MBR، VMBR-I یا VMBR-U، آستانه δ تنها متغیری است که تعداد همسایگان را در پیش‌بینی تعیین می‌کند، اما در WVMBR-I، که از روش‌های وزنی استفاده می‌کند، σ گاوسی وزنی، همچنین می‌تواند اثر همسایگان دور را تغییر داده و از مشارکت آن‌ها در پیش‌بینی جلوگیری کند. با کاهش σ، تأثیر همسایگان دورتر نیز کاهش می‌یابد. توجه داشته باشید که این ارزیابی با تنظیم σ به 10^{-4} انجام شده است، که اطمینان حاصل می‌کند همسایگان دورتر، وزن بسیار کمتری (گاهی اوقات صفر) دارند، نسبت به همسایگان که در فاصله نزدیک‌تر قرار دارند. ازاین‌رو، به‌نوعی، پارامتر σ می‌تواند تعداد همسایگان مشارکت‌کننده را نیز کنترل کند. شکل (۷) حساسیت MAE از WVMBR-I به مقدار σ برای مجموعه‌داده Movielens100k را نشان می‌دهد. در این ارزیابی، مقدار آستانه σ به ۰.۱ تنظیم شد. همان‌طور که مشاهده می‌شود، با کاهش σ، MAE نیز کاهش می‌یابد، اما این در مقادیر σ کمتر از ۰.۰۰۰۱ اعمال نمی‌شود، زیرا با σ بسیار کوچک، همسایگان بسیار کمی (گاهی‌اوقات صفر) در پیش‌بینی نقش دارند.

شکل (۸) سه روش پیشنهادی نهایی را از نظر زمان اجرا با روش‌های رقیب مقایسه کرده است. همان‌گونه که در شکل (۸) مشخص می‌باشد، زمان اجرای سه روش پیشنهادی نهایی نسبت به روش MBR تا حدودی بیشتر بوده ولی نسبت به سایر روش‌ها زمان اجرای قابل قبولی دارند.

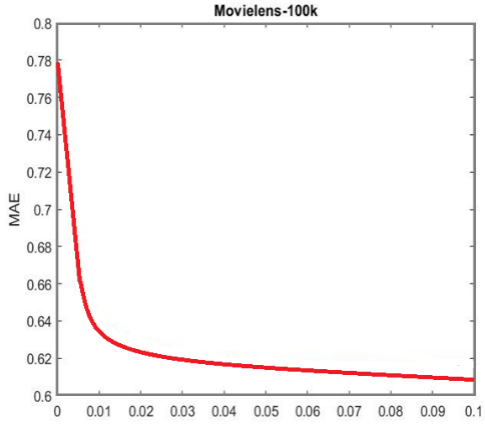
در شرایط شروع سرد جزئی، شباهت شی یادگیری محاسبه شده و در منطق توصیه اعمال می‌شود. برای محاسبه دقت در شرایط شروع سرد روش پیشنهادی با آستانه‌های مختلف مورد آزمایش قرار گرفته است. در این بخش فرعی، آزمایش‌های انجام‌شده برای یافتن اینکه چگونه دقت توصیه‌ها با تعداد مواد آموزشی رتبه‌بندی شده با آستانه‌های مختلف (آستانه δ) متفاوت است، ارائه می‌شوند. شکل‌های (۹ تا ۱۱)، دقت توصیه را برای روش‌های پیشنهادی با آستانه‌های مختلف نشان می‌دهند. همان‌طور که از شکل‌های (۹ تا ۱۱) درک می‌شود، وقتی مقدار آستانه 10^{-4} است، دقت پیش‌بینی‌شده به‌تدریج با افزایش تعداد کاربران افزایش می‌یابد.

[DOI: 10.52547/jsdp.19.3.147]



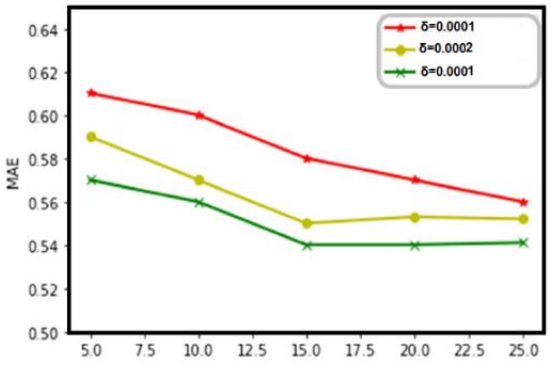
(شکل-۸): مقایسه سه روش پیشنهادی نهایی از نظر زمان اجرا با روش‌های رقیب

(Figure-8): Comparison of the final three proposed methods in terms of execution time with competing methods



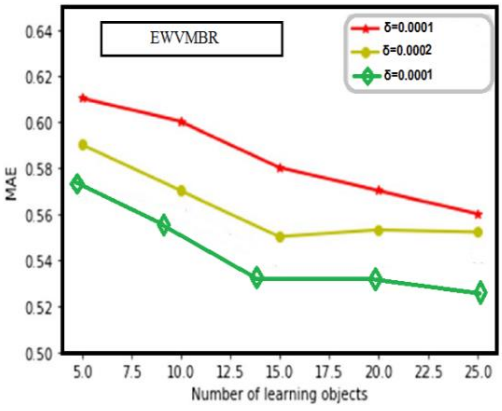
(شکل-۵): حساسیت MAE از WVMBR-I به مقدار δ برای مجموعه‌داده Movielens100k

(Figure-5): Sensitivity of MAE of the WVMBR-I to the δ-value for the Movielens100k dataset



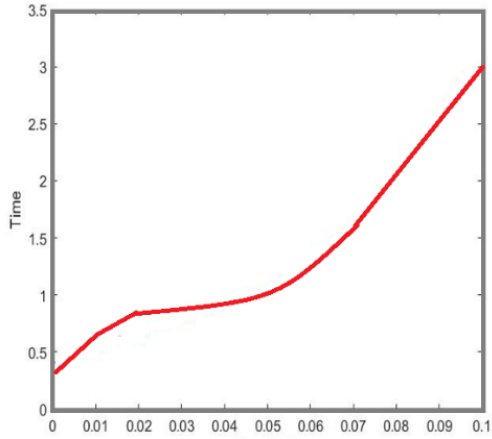
(شکل-۹): دقت روش EVMBR در حالت شروع سرد کاربر با آستانه‌های مختلف

(Figure-9): Accuracy of EVMBR method in user cold start mode with different thresholds



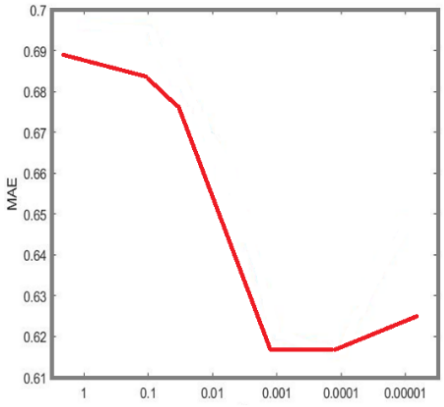
(شکل-۱۰): دقت روش EWVMBR در حالت شروع سرد کاربر با آستانه‌های مختلف

(Figure-10): Accuracy of EWVMBR method in user cold start mode with different thresholds



(شکل-۶): حساسیت زمان اجرا WVMBR-I به مقدار δ برای مجموعه‌داده Movielens100k

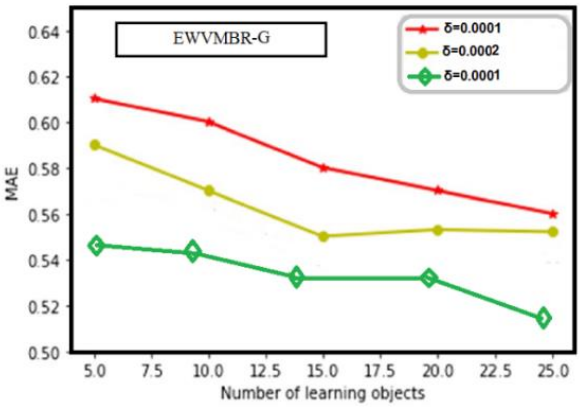
(Figure-6): Sensitivity of runtime of the WVMBR-I to the δ-value for the Movielens100k dataset



(شکل ۷): حساسیت MAE از WVMBR-I به مقدار σ برای مجموعه‌داده Movielens100k

(Figure-7): Sensitivity of MAE of the WVMBR-I to the σ-value for the Movielens100k dataset

[DOI: 10.52547/jsdp.19.3.147]



(شکل-۱۱): دقت روش EWVMBR-G در حالت شروع سرد کاربر با آستانه‌های مختلف

(Figure-11): Accuracy of EWVMBR-G method in user cold start mode with different thresholds

۵- نتیجه‌گیری

در این مقاله، ما هشت سیستم پیشنهادی مبتنی بر MBR را پیشنهاد کردیم که با ترکیب رویکردهای زیر توسعه یافته‌اند:

- استفاده از میانگین و واریانس رتبه‌بندی آیتم‌ها به‌عنوان ویژگی‌های آیتم برای یافتن آیتم‌های مشابه در رویکرد نزدیکترین آیتم‌های همسایه
- استفاده از میانگین و واریانس رتبه‌بندی کاربران به‌عنوان ویژگی‌های کاربر برای یافتن کاربران مشابه در رویکرد نزدیکترین کاربران همسایه
- استفاده از فرمول میانگین وزنی برای تلفیق رتبه‌بندی کاربران/ آیتم‌های همسایه
- استفاده از یادگیر جمعی

تحولات انجام‌شده با سه مجموعه‌داده واقعی نشان داد که روش‌های پیشنهادی EVMBR، EWVMBR و EWVMBR-G، که از یادگیر جمعی استفاده می‌کنند، دقیق‌ترین روش‌ها در بین روش‌های ارزیابی‌شده هستند. دقیق‌ترین روش پیشنهادی، EWVMBR-G، ۲۰ تا ۳۰ درصد میانگین خطای مطلق کمتری نسبت به MBR اصلی نشان داده است. از نظر زمان اجرا، روش‌های پیشنهادی قابل مقایسه با MBR و بسیار سریع‌تر از روش‌های P-kNN، C-kNN و slope-one بودند. برای مجموعه‌داده‌هایی که تعداد آیتم‌های کمتری نسبت به کاربران دارند، روش‌های پیشنهادی مبتنی بر آیتم نسبت به روش‌های مبتنی بر کاربر مدت زمان کمتری دارند.

در کارهای آینده، می‌توان با استفاده از ترکیبی از میانگین و واریانس با سایر خصوصیات آماری، مانند میانگین هر چارک رتبه‌بندی آیتم (یا کاربر)، به‌عنوان ویژگی‌های آیتم (یا کاربر)، دقت روش‌های مبتنی بر MBR را بهبود بخشید.

6- References ۶- مراجع

[1] R. Logesh, V. Subramaniyaswamy, D. Malathi, N. Sivaramakrishnan, V. J. N. C. Vijayakumar, “and Applications, Enhancing recommendation stability of collaborative filtering recommender system through bio-inspired clustering ensemble method,” vol. 32, no. 7, pp. 2141–2164, 2020.

[2] C. Qin et al., ‘‘A survey on knowledge graph-based recommender systems.’’ *Sci. Sin.Inf.* vol. 50(7), pp. 937–956, 2020.

[3] T. V. Yadalam, V. M. Gowda, V. S. Kumar, D. Girish, and M. Namratha, ”Career recommendation systems using content based filtering,” *in 2020 5th International Conference on Communication and Electronics Systems (ICCES*), 2020, pp. 660–665: IEEE.

[4] R. Van Meteren and M. Van Someren, “Using content-based filtering for recommendation,” *in Proceedings of the Machine Learning in the New Information Age: MLnet/ECML2000 Workshop*, Barcelona, 2000, pp. 47–56.

[5] J. Salter, N. Antonopoulos, “CinemaScreen recommender agent: combining collaborative and content-based filtering”, *IEEE Intell. Syst*. Vol. 21(1), pp. 35–41, 2006.

[6] F. Ortega, R. Lara-Cabrera, A´. Gonza´lez-Prieto, J.J.I.S. Bobadilla, “Providing reliability in recommender systems through Bernoulli matrix factorization”, *Inf. Sci*, vol. 553, pp. 110–128, 2021.

[7] M. Mohammadian, Y. Forghani, M.N. Torshiz, “An initialization method to improve the training time of matrix factorization algorithm for fast recommendation”, *Soft Comput*, vol. 25(5), pp. 3975–3987, 2021.

[8] C. C. Aggarwal, Recommender Systems. Springer, 2016.

[9] W. Yue, Z. Wang, W. Liu, B. Tian, S. Lauria, X.J.N. Liu, “An optimally weighted user-and item-based collaborative filtering approach to predicting baseline data for Friedreich’s Ataxia patients,” *Neurocomputing*, vol. 419, pp. 287–294, 2021.

[10] C. Desrosiers, G. Karypis, “A comprehensive survey of neighborhood- based recommendation methods, in Recommender Systems Handbook”, *ed. by F. Ricci, L. Rokach, B. Shapira, B.K. Paul (Springer, Boston, 2011*, pp. 107–144, 2011.

[11] X. Ning, C. Desrosiers, G. Karypis, “A comprehensive survey of neighborhood-based

[DOI: 10.52547/jsdp.19.3.147]

IEEE Transactions on Knowledge and Data Engineering, vol. 28, pp. 2363-2375, 2016.

[23] J. Salter and N. Antonopoulos, "CinemaScreen recommender agent: combining collaborative and content-based filtering," *IEEE Intelligent Systems*, vol. 21, pp. 35-41, 2006.

[24] J. Salter, N. Antonopoulos, "CinemaScreen recommender agent: combining collaborative and content-based filtering", *IEEE Intell. Syst*. vol. 21(1), pp. 35–41, 2006.

[25] J. Qi, T. Qian, and L. Wei, "The connections between three-way and classical concept lattices," *Knowledge-Based Systems*, vol. 91, pp. 143-151, 2016.

[26] Y. Yao, "Rough sets and three-way decisions," in *International Conference on Rough Sets and Knowledge Technology*, 2015, pp. 62-73.

[27] Y. Yao, "Three-way decisions with probabilistic rough sets," *Information Sciences*, vol. 180, pp. 341-353, 2010.

[28] M. K. Condli, D. D. Lewis, D. Madigan, and C. Posse, "Bayesian mixed-E ects models for recommender systems," in *ACM SIGIR*, 1999.

[29] Y. Yuan, X. Luo, and M.-S. Shang, "Effects of preprocessing and training biases in latent factor models for recommender systems," *Neurocomputing*, vol. 275, pp. 2019-2030, 2018.

[30] L. Ren and W. Wang, "An SVM-based collaborative filtering approach for Top-N web services recommendation," *Future Generation Computer Systems*, vol. 78, pp. 531-543, 2018.

[31] M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt, and B. Scholkopf, "Support vector machines," *IEEE Intelligent Systems and their applications*, vol. 13, pp. 18-28, 1998.

[32] H. Wu, Z. Zhang, K. Yue, B. Zhang, J. He, and L. Sun, "Dual-regularized matrix factorization with deep neural networks for recommender systems," *Knowledge-Based Systems*, vol. 145, pp. 46-58, 2018.

[33] K. Bagherifard, M. Rahmani, M. Nilashi, V. Rafe. "Performance improvement for recommender systems using ontology", *Telematics Informatics*, vol. 34(8), pp. 1772-1792, 2017.

[34] K. Bagherifard, M. Rahmani, M. Nilashi, V. Rafe, M. Nilashi, "A Recommendation Method Based on Semantic Similarity and Complementarity Using Weighted Taxonomy: A Case on Construction Materials Dataset", *J. Inf. Knowl. Manag*. vol. 17(1), pp. 1-26, 2018.

[35] M. Nilashi, O. Ibrahim, K. Bagherifard, "A recommender system based on collaborative filtering using ontology and dimensionality recommendation methods, in Recommender Systems Handbookm", *ed. by F. Ricci, L. Rokach, B. Shapira (Springer, Boston, 2015*), pp. 37–76, 2015.

[12] G. Jain, T. Mahara, K.N. Tripathi, "A survey of similarity measures for collaborative filtering-based recommender system*", in Soft Computing: Theories and Applications. ed. by M. Pant, K.T. Sharma, O.P. Verma, R. Singla, A. Sikander (Springer, Singapore, 2020)*, pp. 343–352, 2020.

[13] H. Liu, Z. Hu, A. Mian, H. Tian, X.J.K.B.S. Zhu, "A new user similarity model to improve the accuracy of collaborative filtering", *Knowl. Based Syst*, vol. 56, pp. 156–166, 2014.

[14] B. M. Sarwar, G. Karypis, J. A. Konstan, and J. Riedl, "Item-based collaborative filtering recommendation algorithms," *Www*, vol. 1, pp. 285-295, 2001.

[15] 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.

[16] Y. Park, S. Park, W. Jung, and S.-g. Lee, "Reversed CF: A fast collaborative filtering algorithm using a k-nearest neighbor graph," *Expert Systems with Applications*, vol. 42, pp. 4022-4028, 2015.

[17] J. Wang, A. P. De Vries, and M. J. Reinders, "Unifying user-based and item-based collaborative filtering approaches by similarity fusion," in *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*, 2006, pp. 501-508.

[18] M. Zheng, F. Min, H.-R. Zhang, and W.-B. Chen, "Fast recommendations with the m-distance," *IEEE Access*, vol. 4, pp. 1464-1468, 2016.

[19] D. Lemire and A. Maclachlan, "Slope one predictors for online rating-based collaborative filtering," in *Proceedings of the 2005 SIAM International Conference on Data Mining*, 2005, pp. 471-475.

[20] J. Li, L. Sun, and J. Wang, "A slope one collaborative filtering recommendation algorithm using uncertain neighbors optimizing," in *International Conference on Web-Age Information Management*, 2011, pp. 160-166.

[21] Q.-X. Wang, X. Luo, Y. Li, X.-Y. Shi, L. Gu, and M.-S. Shang, "Incremental Slope-one recommenders," *Neurocomputing*, vol. 272, pp. 606-618, 2018.

[22] J. Bu, X. Shen, B. Xu, C. Chen, X. He, and D. Cai, "Improving collaborative recommendation via user-item subgroups,"

[DOI: 10.52547/jsdp.19.3.147]

reduction techniques", *Expert Syst. Appl.* vol. 92, pp.507-520, 2018.

[36] H. Wang, N. Wang, "Collaborative deep learning for recommender systems", *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2015, pp 1235–1244. https://doi.org/10.1145/2783258.2783273.

[37] J. Jeevamol, V.G. Renumol, "An ontology-based hybrid e-learning content recommender system for alleviating the cold-start problem", *Educ. Inf. Technol*, pp. 1–30, 2021.

[38] J. Joy, N.S. Raj, V.G. Renumol, "Ontology-based E-learning content recommender system for addressing the pure cold-start problem", *ACM J. Data Inf. Qual*, vol. 13 (3), pp. 1–27, 2021.

پیام بحرانی دانش‌آموخته دوره دکترای تخصصی دانشگاه آزاد اسلامی واحد علوم و تحقیقات تهران در رشته مهندسی کامپیوتر گرایش سیستم‌های نرم‌افزاری می‌باشد.

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

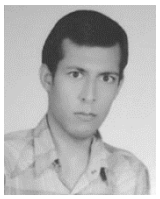
نشانی رایانامه ایشان عبارت است از:

bahranipayam@gmail.com

بهروز مینایی بیدگلی دانش آموخته دانشگاه ایالتی میشیگان آمریکا در رشته علوم و مهندسی کامپیوتر با تخصص هوش مصنوعی و داده‌کاوی است. او در حال حاضر عضو هیأت علمی و دانشیار دانشکده مهندسی کامپیوتر دانشگاه علم و صنعت و رئیس دانشکده مهندسی کامپیوتر است. او 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 کرد و پس از آن به عضویت هیأت علمی دانشگاه آزاد اسلامی واحد نورآباد ممسنی درآمدند. وی هم‌اکنون در چندین واحد دانشگاهی در رشته کامپیوتر 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 کامپیوتر 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

[DOI: 10.52547/jsdp.19.3.147]