یازدهمین کنگره ملی سراسری
فناوریهای نوین در حوزه توسعه پایدار ایران
11th National Congress of
the New Technologies in Sustainable Development of Iran
senaconf.ir

سیستم‌های توصیه گر فیلم: بررسی راهکارها، نقاط قوت و ضعف

علیرضا شفیعی فرد[1]

۱- دانشجوی دکتری فن‌آوری اطلاعات، گرایش تجارت الکترونیک، دانشگاه آزاد اسلامی واحد سبزوار، *Shafiee.fard@gmail.com*

چکیده

در حال حاضر ، اینترنت حاوی مقدار زیادی اطلاعات است که باید برای تعیین مناسب بودن برای برخی کاربران فیلتر شود. سیستم های پیشنهادی ابزاری بسیار مناسب برای این منظور هستند. سیستم‌های پیشنهادی ذکرشده در این مقاله برای پیشنهاد فیلم های مناسب است. این سیستمها با ژانرهای مورد علاقه و نامطبوع کاربر کار می کند . وقتی صحبت از سیستم توصیه فیلم می شود ، توصیه بر اساس شباهت بین کاربران (فیلترینگ مشارکتی۱) یا با در نظر گرفتن فعالیت خاص کاربر (فیلترینگ محتوا۲) که وی می خواهد با آن درگیر شود ، انجام می شود. برای غلبه بر محدودیت های فیلترینگ مشارکتی و مبتنی بر محتوا، به طور کلی ، ترکیبی از فیلترهای مشارکتی و مبتنی بر محتوا استفاده می شود تا سیستم پیشنهادی بهتری ایجاد شود. در این مقاله ، روش های پیشرفته فیلترینگ محتوا ، فیلترینگ مشارکتی ، فیلترینگ ترکیبی۳ و روشهای مبتنی بر یادگیری عمیق۴ برای توصیه فیلم مورد ارزیابی قرار میگیرد.

واژه‌های کلیدی

سیستم توصیه گر، سیستم توصیه مبتنی بر محتوا، فیلترینگ مشارکتی، یادگیری عمیق، فیلم ها

---

۱ Collaborative Filtering

۲ Content Based Filtering

۳ Hybrid Filtering

۴ Deep Learning

معرفی

در حال حاضر ، سیستم های مختلف توصیه شده در حال افزایش هستند و هدف اصلی آنها توصیه محتوای مناسب به کاربر بر اساس پارامترهای مختلف است. سیستم پیشنهادی یک سیستم اطلاعاتی است که برای پشتیبانی از تصمیم گیری کاربر استفاده می شود و محصولات ، اطلاعات و خدمات مناسب را در زمینه فروشگاه های الکترونیکی ، خدمات پخش جریانی ، خدمات همسریابی اینترنتی و بسیاری از زمینه های دیگر توصیه می کند [1])

سیستم های پیشنهادی از تجزیه و تحلیل نوع خاصی از داده ها برای پیش بینی رتبه بندی کاربران برای آیتم‌های مشخص استفاده می کنند. پس از آن ، (بر اساس این تجزیه و تحلیل) ، آنها توصیه هایی را ایجاد می کنند و محتوای صفحه نمایش داده شده را طوری اصلاح می کنند که تا حد ممکن با تمایلات کاربر مطابقت داشته باشد ([1]) این یکی از دلایلی است که طیف وسیعی از شرکت ها و برنامه های تحت وب اخیرا سیستم هایی را برای تجزیه و تحلیل رفتار کاربران با توجه به توصیه مناسب ترین محصولات، خدمات یا اطلاعات پیاده سازی کرده اند. هدف البته افزایش فروش و سودآوری این شرکت ها است. این در درجه اول حوزه سیستمهای پخش جریانی و اجاره آنلاین فیلم است (به طور کلی همه سرویس های پخش آثار بصری) که در آن توصیه محصولات مربوط به یک کاربر نقش مهمی دارد. هدف اصلی همه این سرویس ها این است که هر نمایش محصول توسط کاربر منجر به اقدام کاربرجهت خرید دسترسی به تماشای فیلم ، ثبت نام مشترک ، و غیره گردد. علاوه بر این، این خدمات همچنین سعی می کنند بهترین تجربه خریدار ممکن را پیشنهاد دهند، مشتریانی که با محصولات نمایش داده شده مطابق با سلیقه خود مواجه می‌گردند ، به طور بالقوه مجدد بازمی گردند و دوباره چیزی را خریداری می کنند. برجسته ترین بازیگران در میان پیشگامان سیستم های توصیه گر، آمازون و Netflix هستند.

آمازون نسخه اول سیستم توصیه کننده خود را از سال ۲۰۰۴ ثبت اختراع کرد. این سیستم محصولات خود را با ۲۹٪ افزایش به ۱۲,۸۳ میلیارد دلار افزایش داد. مقایسه دلار در سه ماهه دوم (در مقایسه با سود سال گذشته ۹,۹ میلیارد دلار) ([2]).

Netflix برای کاهش تعداد اشتراک های لغو شده و همچنین افزایش متوسط زمان تعامل کاربر با برنامه (به عنوان مثال ، تعداد ساعت های پخش جریانی) یک سیستم پیشنهاد دهنده در برنامه خود پیاده سازی کرده است. این شرکت پیش بینی می کند که ترکیبی از توصیه ها (اکنون با روند اخیر ، ادامه تماشا ، بدلیل اینکه شما تماشا کرده اید و غیره) و با شخصی سازی ، حداکثر ۱ میلیارد دلار در سال صرفه جویی می شود که در غیراینصورت این مبلغ بایستی برای بدست‌آوردن مشتری های جدید از طرف کسانی که لغو می شوند سرمایه گذاری می شود ([3])

۱,۱ برنامه های کاربردی

سیستم توصیه حوزه وسیعی است که در همه زمینه ها و در همه جاها استفاده می شود. مردم از توصیه ها استفاده می کنند زیرا باعث صرفه جویی در وقت می شود ، بنابراین این برنامه‌های کاربردی در حیطه های مختلف نقشی حیاتی دارد. در زندگی واقعی از این برنامه‌ها همانند سرگرمی ، تجارت الکترونیکی ،خدمات، رسانه های اجتماعی و غیره استفاده می شود.[۴] در حوزه سرگرمی از سیستم توصیه به طور گسترده‌ای در تماشای فیلم یا گوش دادن به موسیقی یا هر برنامه تلویزیونی استفاده می شود. وقتی ما در آیتم زمینه تجارت الکترونیک صحبت می کنیم ، آمازون بزرگترین سایت خرید در جهان است. برخی از آن برای خرید کتاب ، خرید هر کالای خانگی یا هر کالایی، برخی دیگر از آن برای لباس استفاده می کنند. بنابراین از این طریق کل جهان برای این کار یا کار دیگر به این سایت های تجارت الکترونیکی وابسته است. برخی دیگر از سایت های تجارت الکترونیکی مانند cart Flip ، yba E ، Myntra ، clues Shop و غیره نیز توصیه هایی را ارائه می دهند. برخی دیگر از برنامه های سیستم توصیه در زیر ذکر شده است:

پیشنهاد فیلم: Netflix از الگوریتمی برای توصیه فیلم ها با توجه به علاقه آنها استفاده می کند. سایر سیستم عامل هایی که توصیه هایی را ارائه می دهند شامل hotstar ، nyLIVso ، voot ، ALTBalaji و غیره هستند [۵].

توصیه موسیقی: پاندورا یک ایستگاه رادیویی تولید می کند. از ویژگی های آهنگ ها برای توصیه آهنگ های دیگر استفاده می کند. رسانه های دیگر در این زمینه که توصیه موسیقی را پیشنهاد می کنند ، ySpotif ، JioSavan ، Gaana و غیره هستند [۶].

اخبار: برنامه های مختلفی که توصیه های خبری را ارائه می دهند می توانند News Google ، News Apple (یکپارچه در IOS و macOS) ، صفحه Flip ، Feedly ، Deck Tweet ، Pocket ، Mix ، Zig ، ۳۶۰News باشند. همه اینها اخبار ، مقالات ، پست وبلاگ ، مطالب از ناشران برتر و غیره را نشان می دهد [۷].

مد: افراد می توانند اقلام مختلف لباس را به انتخاب خودشان خریداری کنند. این بخش شامل سایتهای مختلف خرید مانند Myntra ، Amazon ، Factory Club ، SHEIN ، Road Lime ، rtca Flip و سایر موارد است [۸].

خدمات مسافرتی: توصیه به شما کمک می کند تا سایت های مختلف مسافرتی را برای سفر امن پیشنهاد دهید. این شامل مسافران جاده ای است که باعث می شود هر سفر جاده ای را با خیال راحت برنامه ریزی کنید. با استفاده از Hooper ، کاربران می توانند برنامه های سفر خود را وارد کنند ، و برنامه به آنها خواهد گفت که چه زمانی بهترین زمان برای رزرو پرواز شان است[9].

۱,۲ چالش ها

سیستم پیشنهادی با چالشهای مختلفی روبرو است. این چالش ها مشکل شروع سرد۵، خلوت بودن داده‌ها۶ و مقیاس پذیری۷ است.

مشکل شروع سرد: سیستم های توصیه گر به تعداد کافی کاربر جهت مطابقت نیاز دارند. به عنوان مثال ، اگر می خواهیم کاربر مشابه یا آیتم مشابهی پیدا کنیم ، آنها را با مجموعه کاربران یا آیتم‌های موجود مطابقت می دهیم. در مرحله اولیه برای یک کاربر جدید ، پروفایل وی خالی است زیرا وی به هیچ آیتمی امتیاز نداده است و سیستم از علاقه او اطلاعی ندارد ، بنابراین ارائه یک توصیه به وی در آیتم هر آیتم برای سیستم دشوار می شود. همین آیتم می تواند با آیتم جدید نیز اتفاق بیفتد، زیرا توسط هیچ کاربری رتبه بندی نشده است زیرا برای کاربر جدید است. با اجرای تکنیک های ترکیبی می توان هر دو این مشکل را برطرف کرد [۱۰].

پراکندگی داده ها: کاربر یا ماتریس رتبه بندی بسیار پراکنده است. یافتن کاربرانی که به آیتم‌های مشابه امتیاز داده اند بسیار دشوار است زیرا بیشتر کاربران به آیتم آیتم نظر امتیاز نمی دهند. بنابراین یافتن مجموعه ای از کاربران که به آیتم امتیاز می دهند دشوار است. توصیه وقتی که اطلاعات کمتری در آیتم هر کاربری وجود داشته باشد ، بسیار دشوار است [۱۱].

مقیاس پذیری: در فیلتر کردن مشارکتی ، از داده های عظیم برای ایجاد قابلیت اطمینان بهتر که به تعداد بیشتری از منابع نیاز دارند، استفاده می شود. با رشد تصاعدی اطلاعات ، نتیجه پردازش پرهزینه و نادقیق می شود [۱۲].

۲ انواع سیستم های توصیه گر:

سه روش سیستم توصیه وجود دارد: فیلتر کردن مشارکتی، فیلتر مبتنی بر محتوا و فیلتر ترکیبی. در سیستم پیشنهادی مبتنی بر محتوا، کاربر داده ها را به طور صریح (با رتبه بندی) یا به طور ضمنی (با کلیک کردن روی پیوند) تهیه می‌نماید. سیستم این داده ها را گرفته و پروفایل کاربری را برای هر کاربر ایجاد می کند. با استفاده از پروفایل کاربر، توصیه ایجاد می شود. در فیلتر کردن محتوا، فقط با مشاهده پروفایل کاربر تنها توصیه صورت می پذیرد. سیستم سعی می کند براساس فعالیت گذشته کاربران، آیتمی مشابه آن آیتم را توصیه کند. برخلاف محتوا، فیلتر کردن مشارکتی آن دسته از کاربرانی را پیدا می کند که علاقه آنها به یک کاربر مشخص شبیه است. سپس با در نظر گرفتن اینکه کاربر داده شده آیتمی را که کاربران دیگر دوست دارند بخاطر اینکه علاقه آنها شبیه است ، آیتم یا هر کالایی را نیز پیشنهاد می کند[13]. هر دو این تکنیک ها قدرت و ضعف خاص خود را دارند ، بنابراین برای غلبه بر این ، تکنیک ترکیبی به تصویر کشیده شد که ترکیبی از هر دو این تکنیک ها است. از فیلترهای ترکیبی می توان در انواع مختلف استفاده کرد. ما می توانیم ابتدا از

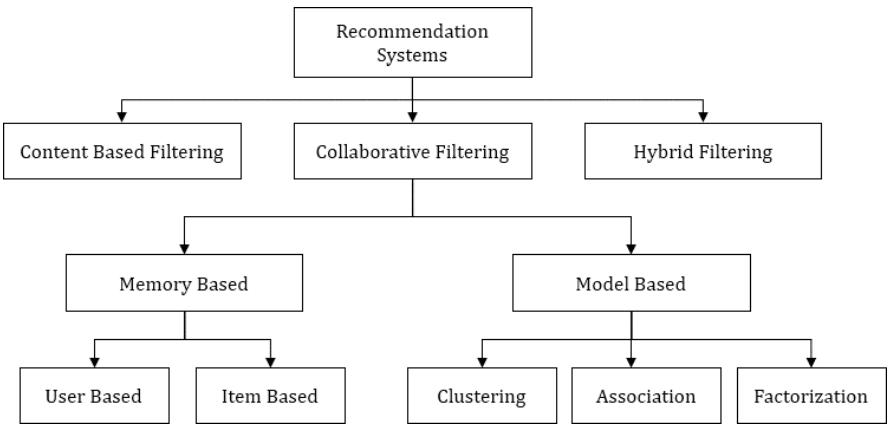
---

۵ Cold Start

۶ Data Sparsity

۷ Scalibility

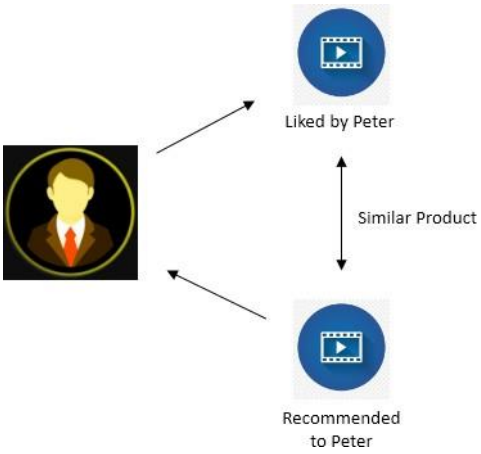
فیلتر مبتنی بر محتوا استفاده کنیم و سپس این نتایج را به فیلترینگ مشارکتی (و بالعکس) یا با تلفیق هر دو فیلتر در یک مدل برای تولید نتیجه منتقل کنیم. این نوع اصلاحات همچنین برای مقابله با شروع سرد ، کمبود یا خلوت بودن داده و مقیاس پذیری داده ها استفاده می شود. طبقه بندی سیستم پیشنهادی در شکل ۱ نشان داده شده است.
در قسمت زیر انواع سیستم های مختلف توصیه آیتم بررسی قرارگرفته است.



شکل ۱: طبقه بندی سیستم توصیه ها

۲,۱ فیلترینگ مبتنی بر محتوا

فیلتر کردن مبتنی بر محتوا به عنوان فیلتر شناختی نیز شناخته می شود [14]. این فیلتر با توجه به تجربه گذشته کاربر، آیتم را به وی توصیه می کند. به عنوان مثال ، اگر کاربری فقط فیلم های اکشن را دوست داشته باشد ، سیستم فقط فیلم های اکشن مشابه آن را که امتیاز بالایی دارد ، پیش بینی می‌کند. توضیح بیشتر می تواند این باشد که کاربر فقط محتوای مربوط به سیاست را دوست دارد، بنابراین سیستم وب سایت ها ، وبلاگ ها یا اخبار مشابه آن را پیشنهاد می کند. بر خلاف فیلترهای مشارکتی ، فیلترینگ مبتنی بر محتوا با مشکل کاربران جدید روبرو نیستند. تعامل با کاربر دیگری در آن وجود ندارد. این فقط با علاقه خاص کاربر سر و کار دارد. فیلتر کردن محتوا ابتدا علاقه و اولویت های کاربر را بررسی می کند و سپس به او فیلم ها یا هر محصول دیگری به او پیشنهاد می کند. این فقط تنها بر ایده ها و افکار کاربر تمرکز می کند و بر اساس علاقه او پیش بینی می کند. بنابراین اگر در آیتم فیلم صحبت می کنیم، آنگاه روش فیلتر کردن محتوا، رتبه بندی داده شده توسط کاربر را بررسی می کند. این روش فیلم های ارائه شده را به این طریق بررسی می کند که



شکل ۲: فیلتر مبتنی بر همکاری

با بررسی دسته های ژانر در پروفایل کاربر ، فیلم‌های رتبه های بالا توسط کاربر مشخص میگردد. پس از تجزیه و تحلیل پروفایل کاربر، این روش فیلم ها را با توجه به علاقه کاربر به او توصیه می کند. شکل ۲ نحوه کار فیلتر کردن مبتنی بر محتوا را نشان می دهد.

همانطور که در شکل ۲ نشان داده شده است ، با ارائه مثالی از اشکال هندسی فیلتر کل فرآیند مبتنی بر محتوا نشان داده شده است [15]. در اینجا در شکل ، ابتدا یک پروفایل بر اساس علاقه کاربر ساخته شده است. در اینجا کاربر دایره و مثلث رنگ آبی را دوست دارد. اکنون بر اساس پروفایل آیتم، پروفایل کاربر ساخته می شود. این پروفایل کاربر با دریافت داده ها از پروفایل آیتم ایجاد می شود. همانطور که در پروفایل آیتم مشاهده می کنیم ، کاربر دایره و مثلث رنگ آبی را دوست دارد بنابراین پروفایل کاربر نیز دارای رنگ دایره ، مثلث و آبی است. اکنون ما این پروفایل کاربری را با مجموعه ای از اشکال مختلف موجود مطابقت می دهیم. در مجموعه اشکال ، ما پنج ضلعی رنگ آبی و سپس دایره ای از رنگ زرد و دو مربع رنگ زرد داریم. بنابراین در اینجا سیستم می یابد که کدام یک از این اشکال با مشخصات کاربر مطابقت دارد.

جیون سون و سئونگ بوم کیم فیلتر کردن مبتنی بر محتوا را برای سیستم های پیشنهادی با استفاده از شبکه های چند منظوره که حاوی اطلاعات ویژگی در آیتم اقلام [16] هستند، پیشنهاد دادند. با استفاده از تمام ویژگی های مبتنی بر تجزیه و تحلیل شبکه آیتم‌های مختلفی توصیه می شود و مشکل فوق تخصصی حل می شود. نتایج نشان می دهد که مشکلاتی مانند پراکندگی و مقیاس پذیری نیز در مقایسه با فیلتراسیون محتوای خالص، داده های باز پیوند داده شده و وزن‌دهی ویژگی آیتم بررسی قرار می گیرند. برای انجام آزمایشات از مجموعه داده های Movie lens استفاده می شود، جایی که به طور تصادفی ۱۰۰ کاربر را برای انجام آزمایش انتخاب می‌کنند و در مقایسه با روش های ذکر شده دقت نیز بهبود می یابد.

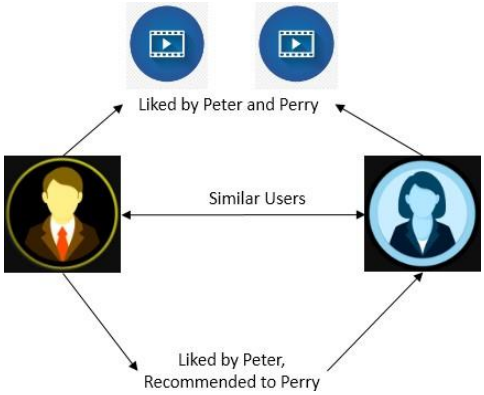
در مقاله ، "تهیه سرگرمی با استفاده از فیلتر مبتنی بر محتوا و استدلال معنایی در سیستم های توصیه گر هوشمند" توسط Yolanda Fernandez-Blanco و همکاران. مشکل فوق تخصصی را حل کرد [17]. برای این منظور ، ارتباط معنایی پنهان بین کاربر و محصول شناخته شده است و سپس از فناوری فعال سازی گسترش برای شناسایی گره‌ای که به شدت متصل است استفاده می شود.

در آیتم داده های کاربر دیگر نیازی به توصیه نیست. برای رویکرد مبتنی بر محتوا آسان است که آیتم‌های جدید را توصیه کنید. آن بر اساس سلیقه منحصر به فرد به کاربر به او توصیه می کند. اولین مشکلی برای ارزیابی وجود ندارد. همچنین ویژگی محتوایی را ارائه می دهد که به ما کمک می کند دلیل یک آیتم توصیه شده را توضیح دهیم.

به عمل پیدا کردن ویژگی مشخص همانند تصاویر یا فیلم های ژانر خاص، گاهی اوقات به عنوان یک مساله در نظر گرفته می‌شود. به طور کلی از آن به عنوان مساله یا مشکل فوق تخصصی۸ نیز یاد می شود. هرگز به کاربر هیچ چیز خارج از پروفایل کاربر توصیه نمی شود. به راحتی می توان توصیه آیتم به کاربر را از دست داد زیرا اطلاعات کافی در آیتم آن آیتم وجود ندارد.

۲,۲ فیلتر کردن مشارکتی

مفهوم فیلترینگ مشارکتی برای اولین بار در سال ۱۹۹۱ توسط گلدبرگ و همکاران مطرح شد. [18] سیستم Tapestry فقط برای گروههای کاربر کوچکتر (مثلاً یک واحد) اعمال می شود و درخواست‌های زیادی از کاربر دارد. به عنوان یک سیستم فیلترینگ مشارکتی



شکل ۳: فیلتر مبتنی بر همکاری

تاپستری نوع جدیدی از سیستم پیشنهادی فیلتر کردن مشارکتی را ارائه می دهد، اما کمبودهای فنی زیادی وجود دارد. از آن زمان به بعد ، یک سیستم توصیه مبتنی بر فیلتر کردن مبتنی بر امتیازدهی وجود دارد، همانند Grouplens که اخبار و فیلم ها را توصیه می کند. در حال حاضر بسیاری از سایت های تجارت الکترونیکی از سیستم پیشنهادی مانند آمازون، CDNow، Drugstoو Moviefinder و غیره استفاده می کنند. داده های زیادی در دسترس است. همانطور که همه ما می دانیم امروز در این زندگی شلوغ هیچ کس وقت ندارد که صدها هزار آیتم را جستجو کند و آیتمی را که شبیه سلیقه آنها است انتخاب کند. بنابراین فیلتر کردن مشارکتی یکی از راههای فیلتر کردن داده ها و ارائه اطلاعات مربوط به کاربر است که به آن علاقه مند است. فیلتر کردن مشارکتی یکی از شناخته شده ترین تکنیک ها برای توصیه آیتم است. این روش بر اساس انتخاب همسایه، آیتم مربوطه را به کاربر پیشنهاد می کند. ابتدا به شباهت بین کاربر و همسایه خود پی می برد و سپس آیتم‌ها را پیش بینی می کند. تعداد کاربران می تواند تعداد n باشد. این روش کاربر مشابه را از لیست کاربر پیدا می کند. اما بر اساس امتیازاتی که کاربران به آیتم خاص داده اند ، به شباهت بین کاربران پی برده می شود. به این ترتیب روش ادامه پیدا می کند و نتیجه مطلوب ایجاد می شود. این استراتژی رتبه بندی های داده شده توسط کاربر برای هر آیتم را از فهرست بزرگ فهرست آیتم های رتبه بندی داده شده توسط کاربر می گیرد. این کاتالوگ بزرگ به عنوان ماتریس آیتم - کاربر معرفی می شود [۱۹]. این روش کاربر مشابه را از لیست کاربران پیدا می کند. اما بر اساس امتیازاتی که کاربران به آیتم خاص داده اند ، به شباهت بین کاربران پی برده شده است. به این ترتیب روش ادامه می یابد و نتیجه مطلوب ایجاد می شود. این استراتژی رتبه بندی های داده شده توسط کاربر برای هر آیتم را از فهرست بزرگ فهرست آیتم های رتبه بندی داده شده توسط کاربر می گیرد. این کاتالوگ

---

۸ Overspecialization Problem.

بزرگ به عنوان ماتریس آیتم-کاربر شناخته می شود [۱۹]. این روش کاربر مشابه را از لیست کاربر پیدا می کند. اما بر اساس امتیازاتی که کاربران به آیتم خاص داده اند، به شباهت بین کاربران پی برده شده است. به این ترتیب روش ادامه می یابد و نتیجه مطلوب ایجاد می شود. این استراتژی رتبه بندی های داده شده توسط کاربر برای هر آیتم را از فهرست بزرگ فهرست آیتم های رتبه بندی داده شده توسط کاربر می گیرد. این کاتالوگ بزرگ به عنوان ماتریس آیتم-کاربر معرفی می شود [۱۹].

شکل ۳ فیلتر کردن مشارکتی را با مثالی از پیشنهاد فیلم به کاربر توضیح می دهد. همانطور که در نمودار به وضوح نشان داده شده است که کاربر ۱، کاربر ۲ و کاربر ۳ فیلم ها را با توجه به علاقه خود رتبه بندی کرده اند. بر اساس آن ماتریس فیلم های کاربر ایجاد می شود و هر مدل شباهت برای یافتن شباهت بین آنها اعمال می شود تا بتوان به کاربر ۳ توصیه کرد.

چینگ-Wu She و همکاران از فیلترهای مشارکتی هر دو رویکرد یعنی فیلتر کاربر بر اساس و آیتم استفاده کرده اند [۲۰]. نویسندگان برای یافتن شباهت بین کاربران از شباهت همبستگی پیرسون استفاده کرده اند. یک الگوریتم دیگر به نام User N Nearest odNeighborho برای بدست آوردن N کاربر مشابه تا بتوان آنها را در یک گروه قرار داد. برای یافتن شباهت آیتم ، الگوریتمی معروف به Logicely Log شباهت. نتایج بدست آمده از پیشنهاد مبتنی بر آیتم ، در سیستم پرونده توزیع شده Hadoop (HDFS) ذخیره می شود. مجموعه داده آیتم استفاده در اینجا از Database Scope Web hResearc Yahoo بدست آمده است.

مهربخش نیلاشی روش پیشنهادی مبتنی بر فیلتر کردن مشارکتی را با استفاده از روش هستی شناسی و کاهش ابعاد برای بهبود پراکندگی و مقیاس پذیری در فیلترینگ مشارکتی پیشنهاد کرد [۲۱]. در اینجا آزمایشاتی نیز ارائه می شود که دقت پیش بینی و توان سیستم توصیه فیلم را بهبود می بخشد. تیانکی ژو و همکاران از مدل برنامه نویسی Hadoop برای پیاده سازی الگوریتم توصیه بر اساس فیلترینگ مشارکتی استفاده کرد. مجموعه داده M۱۰-Movielens استفاده می شود که شامل ۱۰۰۰ ردیف درجه بندی است [۲۲].

فیلترینگ مشارکتی به دو نوع تقسیم می شود. آنها روشهای مبتنی بر حافظه و مدل هستند [۲۳] ، [۲۴].

۲,۲,۱ *روش مبتنی بر حافظه*

این روش به عنوان رویکرد همسایگی نیز شناخته می شود [۲۵]. روش مبتنی بر حافظه از معیارهای تشابه محاسبه شده از رتبه بندی صریح کاربر برای یافتن همسایه و تولید پیش بینی استفاده می کند [۲۶] ، [۲۷]. این نوع روش علاقه کاربر را برای هر آیتم مشاهده می کند. پس از تجزیه و تحلیل دیدگاه کاربر برای آیتمی ، کاربر مشابهی را که دارای علاقه مشابه آن کاربر است نیز بررسی می کند. بنابراین یافتن کاربران مشابه با مطالعه ماتریس ابزار انجام می شود. بنابراین این نوع روش اساساً مبتنی بر حافظه سیستم برای بدست آوردن پیش بینی کاربر مشابه است. بنابراین در اینجا رتبه بندی ناشناخته هر کاربر را می توان با استفاده از ماتریس رتبه بندی آیتم کاربر (ماتریس سودمند) در صورت یافتن کاربر مشابه ، ایجاد کرد. در آخر می توان توصیه کرد [۲۸].

رویکرد مبتنی بر حافظه بیشتر به دو نوع طبقه بندی می شود. رویکرد مبتنی بر کاربر و رویکرد آیتم محور [۲۹] ، [۳۰].

رویکرد مبتنی بر کاربر این روش به فیلتر کردن کاربر به کاربر نیز معروف است. در این روش ، یک ماتریس رتبه بندی از n کاربر و m آیتم ایجاد می شود. برای یافتن توصیه برای کاربر جدید ، این روش نزدیکترین همسایه را با استفاده از رتبه قبلی همسایه پیدا می کند و برای یک آیتم پیش بینی ایجاد می کند. به عبارت دیگر ، با بررسی اینکه کدام کاربر ذائقه مشابهی دارد ، پیشنهاد داده می شود [۳۱]. شباهت بین کاربران با استفاده از اقدامات مختلف شباهت یا ایجاد خوشه ها پیدا می شود.

حمیدرضا کوهی و کوروش کیانی خوشه بندی C به معنای فازی را برای CF مبتنی بر کاربر پیشنهاد دادند [۳۲]. ماتریس رتبه بندی به پنج مجموعه آموزش مختلف تقسیم می شود. سپس تکنیک های خوشه بندی: means-K ، SOM و خوشه بندی فازی برای ایجاد خوشه هایی برای یافتن نزدیکترین همسایه برای کاربر جدید برای پیش بینی رتبه بندی وی و ارائه توصیه استفاده می شود. با آزمایش های انجام شده ، نشان داده شده است که c-means فازی از نظر دقت عملکرد بهتری نسبت به means-K و SOM دارد. در این مقاله همچنین مشاهده می شود که با افزایش تعداد خوشه ، دقت نیز کاهش می یابد. برای آزمایش ، ۸۰٪ داده برای آموزش و ۱۰٪ برای آزمایش استفاده می شود.

نینگ نینگ یی و همکاران یک سیستم توصیه فیلم با استفاده از پایگاه داده گراف پیشنهاد داده است [۳۳]. برای یافتن شباهت، از CF مبتنی بر کاربر استفاده می شود. از آنجا که داده ها کم است ، ماتریس رتبه بندی آیتم کاربر از قبل پر می شود. جدول موجود در پایگاه داده آیتم استفاده عبارتند از user.u (شناسه کاربر ، سن ، جنس ، شغل) ، item.u (شناسه فیلم ، نام ، تاریخ انتشار ، وب سایت) ، data.u (شناسه کاربر ، شناسه فیلم ، امتیاز ، مهر زمان). برای تشخیص فیلم ها از رنگ های مختلف استفاده می شود. با انجام آزمایشات ، مشاهده می شود که برای فیلمهای بسیار توصیه شده ، گره لبه زرد اضافه می کند و ضخامت لبه پیشنهاد فیلم است. مشاهده می شود که با افزایش شعاع گره ها و ضخیم شدن لبه ها ، امتیاز فیلم ها افزایش می یابد. برای پیاده سازی ، از neo۱py استفاده شده است که یک کتابخانه فعال با j۴Neo و مجموعه داده k۱۰۰movielens است.

رویکرد مبتنی بر آیتم این روش همچنین به عنوان فیلتر کردن آیتم به آیتم شناخته می شود. این برای توصیه هر آیتم بر اساس رتبه بندی آیتم مرتبط استفاده می شود. با تجزیه و تحلیل رتبه بندی، کاربرانی که امتیاز آنها برای آیتم‌های مختلف مشابه است ، فقط آن آیتم‌ها برای توصیه در دسترس هستند [۳۴]. به طور گسترده توسط همه غول های وب از جمله Netflix ، YouTube و غیره استفاده شده است.

گیلبرت بادارو و دیگران [۳۵] یک روش ترکیبی را ارائه داد که ترکیبی وزنی از فیلترینگ کاربر و فیلترینگ آیتم برای یافتن رتبه های ناشناخته برای یک آیتم است ، به طوری که آیتم با رتبه برتر قابل توصیه است. امتیازات ubcf و ibcf با وزنی که به آنها اختصاص داده می شود برای پیش بینی رتبه بندی یک آیتم ضرب می شوند. نویسنده نشان داد که این ترکیب همچنین به غلبه بر مساله کاربر و آیتم شروع سرد مورد استفاده قرار می‌گیرد. این رویکرد همچنین آدرس می دهد که پراکندگی داده ها کاهش می یابد و دقت آنها بهبود می یابد. برای هدف آزمایش از مجموعه داده های movie lens استفاده می شود.

۲,۲,۲ *روش مبتنی بر مدل*

در روش مبتنی بر مدل ، یک مدل کاربر با استفاده از رتبه بندی هر کاربر برای ارزیابی مقدار آیتم انتظار از آیتم‌های بدون رتبه توسعه می دهد [۳۶] ، [۳۷]. این روش به طور کلی از الگوریتم یادگیری ماشین یا داده کاوی برای ایجاد یک مدل استفاده می کند. این مدل با استفاده از ماتریس سودمندی ساخته می شود که با استفاده از رتبه بندی داده شده توسط کاربر برای هر آیتم ساخته می شود. این مدل با گرفتن اطلاعات از ماتریس ابزار آموزش داده می شود. اکنون این مدل با استفاده از داده های داده شده برای تولید پیش بینی برای کاربران آموزش دیده است [۳۸] ، [۳۹]. رویکرد مبتنی بر مدل بیشتر در دسته های مختلف طبقه بندی می شود. آنها به عنوان استخراج قوانین انجمنی، درخت تصمیم ، خوشه بندی ، شبکه عصبی مصنوعی ، رگرسیون و غیره هستند. نمونه های مختلفی وجود دارد که روی روش مبتنی بر مدل کار می کنند. بعضی از آنها نهفته هستند

روشهای معنایی مانند Latent Semantic Analysis و Latent Semantic Indexing و تکنیک‌های Dimensionality Reduction مانند تجزیه ارزش واحد (SVD) ، فاکتوراسیون ماتریس و غیره می باشند. تکنیک‌های مبتنی بر مدل برای حل مساله پراکندگی که در سیستم‌های توصیه گر رخ می دهد استفاده می‌شوند.

(الف) فاکتورسازی ماتریس [۴۰] ، [۴۱] قدرتمندترین مدلی است که به طور کلی کاربران و فیلم ها را به شکل ماتریس استفاده می کنند. این ردیف ها را به عنوان کاربر یا فیلم ها را به عنوان ستون نشان می دهد. راه دیگر نیز استفاده می شود. همچنین می تواند فیلم ها را به عنوان سطر و ستون به عنوان کاربر نشان دهد. به طور کلی ، همه کاربران به هر آیتم امتیاز نمی دهند. بنابراین هنگامی که ما یک ماتریس ایجاد می کنیم ، آن ماتریس به ماتریس پراکنده معروف است. کاربر آشکار تحت این دسته قرار می گیرد زیرا هر کاربری امتیاز نمی دهد. وقتی در آیتم رتبه بندی ضمنی صحبت می کنیم ، رتبه بندی هر آیتم با دیدن آنچه کاربر در گذشته خریداری کرده محاسبه می شود. براساس آن رتبه بندی تاریخچه خریداری شده داده می شود. فاکتوراسیون ماتریس با استفاده از فرمول زیر عمل می کند.

$$r_{ui} = {q_i}^T \times p_u \qquad (۱)$$

در اینجا q_i به ما در مورد آیتم iام می گوید که کاربر آن را دوست دارد یا نه. با اختصاص دادن ، بازخورد مثبت یا منفی می دهد ارزش مثبت یا منفی p_u درباره کاربری که به آیتم خاصی علاقه دارد به ما می گوید ، r_{ui} درآیتم رتبه بندی که کاربر u به آیتم i می دهد به ما می گوید.

اکنون ابتدا باید فاکتورهای q_i و p_u را پیدا کنیم. ما باید p_uو q_i را فاکتور بندی کنیم تا بتوانیم ماتریس کاربر به آیتم را ایجاد کنیم. ما باید فاکتورهایی را پیدا کنیم که بسیار نزدیک به مقدار واقعی باشند. بنابراین احتمال خطا بسیار کمتر است.

فرمول یافتن حداقل مقادیر برای q_i و p_u در زیر آورده شده است.

$$min \sum_{(u,i\epsilon k)} (r_{ui} - q_i^T \times p_u)^2 + \delta\left(||q_i||^2 + ||p_u||^2\right) \quad (۲)$$

در اینجا k یک جفت (u, i) است. با استفاده از این فرمول می توان ماتریس کاربر به آیتم را پیدا کرد و ما باید δ را در معادله (۲-۱) به حداقل برسانیم

تا خطا کاهش یابد. برای به حداقل رساندن مقدار ، روشهای شناخته شده Stochastic Gradient Descent و Alternating Least Square استفاده می شود. از این دو روش می توان برای به حداقل رساندن معادله فوق استفاده کرد.

در مقاله ای که به دست آمده است ، بهنوش عبدالهی و اولفا نصراوی برای محاسبه توصیه های برتر [۴۲] ، فاکتورسازی ماتریس قابل توضیح برای فیلتر کردن مشارکتی را پیشنهاد کرده اند. نویسندگان نشان دادند که اگر j آیتم برای کاربر i قابل توضیح باشد ، نمایندگی آنها در فضای نهفته باید نزدیک به یکدیگر باشد. از شباهت کسینوس برای یافتن نزدیکترین همسایگان استفاده می شود. آزمایش این ایده را ارائه می دهد که این روش عملکرد خوبی برای آن دارد

نسل برتر k توصیه.

نویسندگان Bokde Dheeraj و همکاران مدل فاکتورسازی ماتریس در الگوریتم فیلترهای مشارکتی بررسی شده است. مدل های فاکتوراسیون ماتریس مانند SVD ، PCA ، PMF و NMF بحث شده است. آزمایش ها نتیجه گرفتند که SVD تصادفی دقت و درستی CF مبتنی بر کاربر و آیتم را افزایش می دهد. همچنین هزینه محاسبه هر دو الگوریتم CF را کاهش می دهد [۴۳].

(ب) تجزیه مقدار منفرد تجزیه ارزش واحد از مدل های فاکتورسازی ماتریس است [۴۴] ، [۴۵]. این روش A را به عنوان ماتریس داده ورودی می گیرد و آن را در محصول سه ماتریس مختلف که به شرح زیر شرح داده شده ، توصیف می کند.

$$A_{[m\times n]} = U_{[m\times r]} S_{[r\times r]} (V_{[r\times n]})^T \quad (۳)$$

A ماتریس داده ورودی با اندازه $m \times n$به عنوان مثال m کاربر و n فیلم است. U ماتریس با اندازه $m \times r$ است که بردارهای منفرد سمت چپ را ذخیره می کند. بردارهای منفرد S حاوی مقادیر واحدی از ماتریس مورب اندازه $r \times r$ است که جز قطر اصلی خود در همه جا صفر دارد. بنابراین تمام این مقادیر غیر صفر مقادیر واحدی هستند و به ترتیب مرتب می شوند تا بزرگترین مقدار در اولویت قرار گیرد. V یک بردار مجزا راست از اندازه $r \times n$ است.

معادله ۳ را می توان به صورت فشرده به شرح زیر نشان داد:

$$A_k = U_k S_k (V_k)^T \quad (۴)$$

A_k نشان دهنده نزدیکترین تقریب خطی ماتریس اصلی A با کاهش رتبه k است. پس از اتمام این تبدیل ، می توان کاربران و آیتم‌ها را به عنوان نقاطی در فضای k-بعدی تصور کرد. خاصیت SVD بیان می کند که U ، S ، V منحصر به فرد هستند که در آنها U و V ستون متعارف هستند یعنی U = ۱ ، VTV = ۱ و S مورب است که ورودی های آن مقادیر منفرد مثبتی هستند و به ترتیب نزولی مرتب می شوند.

Vozalis MG و Margaritis GK از SVD با داده های دموگرافیک استفاده کردند [۴۶] و آن را در فیلترینگ مشارکتی مبتنی بر کاربر و بر اساس آیتم‌ها اعمال کردند. نتیجه انجام شده در مقاله به این نتیجه رسید که CF مبتنی بر کاربر نتیجه بهتری ندارند و CF مبتنی بر آیتم میزان خطای کمتری را ارائه می دهد. همچنین دیده می شود که استفاده از SVD در سیستم پیشنهادی از عهده مقیاس پذیری و مشکل پراکندگی داده ها و همچنین بهبود دقت برمی‌آید.

برای هیچ آیتم خاصی نیازی به انتخاب ویژگی نیست. همانطور که با هر نوع آیتمی کار می کند. این بزرگترین مزیت فیلترینگ مشارکتی است.

این معمولاً با اولین مشکل ارزیابی کننده روبرو می شود ، زیرا نمی تواند هر آیتم بدون رتبه را توصیه کند. همچنین با مشکل سوگیری محبوبیت روبرو است. و آن در زمانی است که حالتی پیش می‌آید که کاربر ممکن است آیتمی را دوست نداشته باشد اما به دلیل محبوبیت آن آیتم به کاربر توصیه می شود.

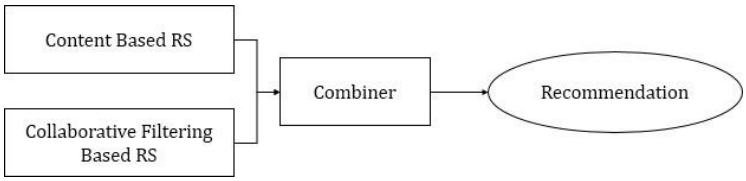
۲,۳ فیلتر ترکیبی

این فیلتر یک سیستم فیلتر کردن اطلاعات است که از فیلم ها به عنوان ورودی از کاربران رتبه بندی می کند و سپس فیلتر کردن مبتنی بر محتوا و مشارکت را اعمال می کند و لیست توصیه ها را ایجاد می کند [۴۷]. این ترکیبی از دو تکنیک است یعنی فیلتر کردن مشارکتی و فیلتر کردن محتوا. وقتی فقط یک روش واحد یعنی فیلتر کردن مشارکتی یا فیلتر کردن محتوا به تنهایی نمی تواند مشکل را حل کند ، مفهوم فیلتر ترکیبی به تصویر کشیده می شود. با استفاده از فیلتر ترکیبی می توان بسیاری از مشکلات فیلترینگ مشارکتی و فیلترینگ محتوا را برطرف کرد. مشکلی مانند مشکل شروع سرد در فیلترینگ مشارکتی، یک چالش اساسی در آن است. بنابراین اگر از فیلتر کردن محتوا استفاده کنیم و سپس از فیلتر مشترک استفاده کنیم ، می تواند راه حلی برای آن باشد. بنابراین ساخت آن بصورت ترکیبی می تواند مشکل را برطرف کند.

Ali.M.S و همکاران مدل ترکیبی پیشنهادی، که از برچسب های ژنومی فیلم ها با مفهوم فیلتر مبتنی بر محتوا استفاده می کند [۴۸]. برای کاهش برچسبهای اضافی از تجزیه و تحلیل مولفه اصلی با تکنیک همبستگی پیرسون استفاده می کند ، بنابراین می توان از پیچیدگی محاسبات کاسته شود. در اینجا از مجموعه داده های Movie lense استفاده شده است که در ۱۸ اکتبر ۲۰۱۶ منتشر شد.

Deng.F و دیگران روشی را برای محاسبه ترجیحات بالقوه کاربر بر اساس ویژگیهای ترکیبی مانند ویژگیهای تولید شده توسط کاربر ، ویژگیهای بصری تصویر و تبدیل رتبه بندی آیتم کاربر به رتبه بندی ویژگیهای ترکیبی تعریف کرد [۴۹]. آزمایشات انجام شده در اینجا نشان می دهد که کار پیشنهادی نتیجه بهتری در مجموعه داده های پراکنده دارد و کارایی بالاتری در مجموعه داده های بزرگ دارد. در [۵۰] سی. یانگ و همکاران. یک رویکرد ترکیبی را بر اساس شباهت اجتماعی و ویژگی آیتم نظر ارائه دهید. نویسنده از روش فیلتر مشارکتی همراه با شباهت های اجتماعی و ژانرهای فیلم استفاده کرده است. وی برای حل مسئله پراکندگی از مدل MF-BPR استفاده کرد. روش پیشنهادی در دو مرحله کار می کند. ابتدا از مدل MF-BPR برای بدست آوردن مجموعه کاندیدها با استفاده از رتبه بندی با آموزش مجموعه داده های آموزشی استفاده می شود. پس از یافتن مجموعه کاندیدها ، رتبه های ناشناخته با استفاده از رتبه بندی های موجود پیش بینی می شود. سپس رتبه بندی ها مرتب شده و مجموعه کاندیدهای نهایی برای هر کاربر بدست می آید. هر مجموعه دارای چندین آیتم برتر است. در مرحله دوم ، فیلم ها با استفاده از روش انتخاب ویژگی IDF-TF برای یافتن شباهت بین کاربران و مجموعه داده های movie lens در اینجا استفاده می شود. نتیجه نشان می دهد که با استفاده از MF-BPR نتیجه دقیق تری نسبت به فیلترینگ مشارکتی نشان می دهد. در [۵۱] ایده استفاده شده Valdiviezo Priscila و J. Bobadilla این است که آنها رتبه بندی های مختلفی از کاربر و اطلاعات جمعیتی مانند سن ، جنسیت و شغل را بدست آورده و آنها را در یک مدل ماتریسی ترکیب کردند. سپس برای کشف رتبه های از دست رفته از فیلتر مشترک استفاده می شود. ایده اصلی آیتم استفاده در اینجا بهبود پیش بینی رتبه بندی کلی است. سه مجموعه داده مختلف مانند Lens Movie (k۱۰۰) ، Crossing Book (BX) و Trust Film. در اینجا MAE (میانگین خطای مطلق) برای اندازه گیری عملکرد روش پیشنهادی استفاده می شود. مشکل پراکندگی داده ها نیز با استفاده از ویژگی

دموگرافیک کاربر و آیتم حل می شود. در [۵۲] Bharti .R و Gupta .D سیستمی را پیشنهاد داده اند که از فیلتر مبتنی بر محتوا برای کاربران جدید و فیلتر مشارکتی برای کاربر قدیمی استفاده می کند و برای یافتن شباهت نویسنده از شباهت کسینوس و پیرسون استفاده کرده و در آخر از کندو برای ذخیره سازی استفاده می شود کاربر و جزئیات فیلم ها در پایگاه داده. برای فیلم آزمایشی از مجموعه داده لنز استفاده می شود. گوپتا سیستمی را پیشنهاد داده است که از فیلتر محتوا برای کاربران جدید و فیلترینگ مشارکتی برای کاربر قدیمی استفاده می کند و برای یافتن شباهت نویسنده مقاله از شباهت کسینوسی و پیرسون استفاده کرده و در نهایت مکانی برای ذخیره اطلاعات کاربر و فیلم ها در پایگاه داده استفاده می شود. برای فیلم آزمایشی از مجموعه داده لنز استفاده می شود.



شکل ۴: فیلتر ترکیبی

فیلترهای ترکیبی دسته های مختلفی دارد. فیلتر ترکیبی با درک شکل زیر قابل درک است. همانطور که در شکل نشان داده شده است ، ابتدا ورودی به فیلتر مشترک و فیلتر مبتنی بر محتوا اعمال می شود ، و سپس نتیجه نهایی آنها برای بدست آوردن فیلتر ترکیبی ترکیب می شود.

۲,۳,۱ *ترکیبی وزن دار*

در این روش ترکیبی ، در ابتدا نتایج آیتم‌هایی که توصیه می شوند از لیست تکنیک های توصیه شده که مورد استفاده در سیستم خاص می باشند تولید می شوند. سیستمی با نام Tango-P معروف به Tango شخصی سازی شده [۵۳] از این مفهوم ترکیبی استفاده می کند. این سیستم Tango-P از frontend ، back end و پایگاه داده تشکیل شده است. کاربر از طریق مرورگر وب به front end دسترسی پیدا می کند ، back end مقاله را بارگیری می کند و پیش بینی می کند. به طور کلی فیلتر و ترکیبی وزنی مبتنی بر همکاری و محتوا استفاده می شود. به طور خلاصه ، ترکیبی وزنی فیلترینگ مشارکتی و مبتنی بر محتوا را به صورت جداگانه اجرا می کند و سپس سرانجام پیش بینی های آنها را ترکیب می کند.

۲,۳,۲ *سوئیچینگ هیبرید/سیون*

همانطور که از نام آن مشخص است ، تغییر یا تغییر روش توصیه انجام می شود. این سوئیچینگ بر اساس وضعیت فعلی سیستم انجام می شود. بنابراین معیارها توسط سیستم تصمیم می گیرند تا بین دو سیستم پیشنهادی جابجا شوند. این روش سوئیچینگ به طور کلی برای جلوگیری از مشکل افزایش سطح استفاده می شود. اما فیلتر کردن مشارکتی و محتوایی هر دو با مشکل کاربر جدید روبرو است. این تکنیک توسط سیستمی به نام Learner Daily اجرا شده است که در ابتدا از فیلتر کردن محتوا استفاده می کند و سپس از فیلترینگ مشارکتی استفاده می کند. این روش ترکیبی توسط [۵۴] برای مقابله با مشکلات مختلف شروع به کار سرد استفاده می شود. نویسنده از CC-CAMF مبتنی بر محتوا و CC-CAMF برمبنای جمعیت استفاده کرده است تا راهی برای کنار آمدن با مشکل شروع سرد ایجاد کند.

۲,۳,۳ ترکیبی آبشار

در این تکنیک توصیه توسط یک تکنیک تولید می شود و نتایج پیشنهادی برای بهبود سیستم پیشنهادی مجدداً اصلاح یا فیلتر می شود. سیستم توصیه موسیقی ارائه شده توسط [۵۵]. مانند یک سیستم میان افزار عمل می کند که امکان کتابخانه های صوتی / موسیقی دیجیتال را فراهم می کند. برای ارائه توصیه ها از فیلتر مبتنی بر محتوا و سپس از فیلتر مشارکتی استفاده می کند. این سیستم با توجه به درخواست کاربر، موسیقی از همان ژانر را توصیه می کند و همچنین با فیلترینگ مشارکتی، اولویت کاربر قبلی و همچنین اولویت کاربر دیگر را در نظر می گیرد. توصیه کننده رستوران EntreeC ، همچنین یک توصیه مبتنی بر دانش و همکاری آبشاری است.

۲,۳,۴ ترکیبی میکس شده

همانطور که از نام آن پیداست ، همزمان بیش از یک توصیه کننده را ارائه می دهد تا تعداد زیادی از توصیه ها را با هم ارائه دهد. وقتی کاربران بخواهند به طور همزمان تعداد بیشماری از توصیه ها را دریافت کنند ، از این پیشنهاد مختلط استفاده می شود. یک مثال که نوع مخلوط را نشان می دهد ، سیستم توصیه کننده ProfBuilder [۵۶] است که به عنوان یک سیستم پیشنهادی مبتنی بر عامل برای یک وب سایت کار می کند. ProbBuilder ابتدا اطلاعات استفاده از سایت را برای یافتن علاقه کاربر و پایه ای برای فیلتر کردن مشترک جمع آوری می کند. سپس به کاربر کمک می کند تا با استفاده از هر دو تکنیک فیلترینگی مبتنی بر محتوا و فیلترینگ مشارکتی ، صفحه مربوط به انتخاب خود را پیدا کند.

۲,۳,۵ ترکیب ویژگی

در اینجا در ابتدا برخی از ویژگی ها توسط یک تکنیک توصیه تولید می شوند و سپس آنها بر روی تکنیک های پیشنهادی دیگر اعمال می شوند. بنابراین اساساً این روش ادغامی را برای فیلترینگ مشارکتی و مبتنی بر محتوا فراهم می کند. به این ترتیب ویژگی های موتور پیشنهادی مختلف با هم ترکیب می شوند و در یک الگوریتم پیشنهاد واحد قرار می گیرند [۵۷].

۲,۳,۶ تقویت ویژگی

در اینجا ، خروجی یک موتور پیشنهاد دهنده به عنوان ورودی به سیستم پیشنهاد دهنده دیگر استفاده می شود. سیستم توصیه کننده اول از رتبه بندی و اطلاعات اضافی استفاده می کند که به عنوان ورودی به سیستم دوم خورانده می شود. اکنون سیستم دوم این اطلاعات را می گیرد و همچنین برخی عملکردهای اضافی را برای تولید توصیه ها اضافه می کند [۵۷]. تقویت ویژگی بهتر از ویژگی ترکیبی است زیرا ویژگی های بیشتری به موتور توصیه کننده قبلی اضافه می کند.

۲,۳,۷ سطح متا

این نیز یکی از راه های ترکیب دو سیستم و ارائه توصیه ها است. مدل داخلی یا پایه توسط موتور توصیه کننده تولید می شود و سپس آن مدل به طور کلی برای سیستم دیگر اعمال می شود. این می تواند یک تصور غلط باشد که سطح متا و تقویت ویژگی ها یکسان هستند درحالی که کاملاً با یکدیگر متفاوت هستند. در تقویت ویژگی ، برخی اطلاعات اضافی با داده های موجود به سیستم داده می شوند در حالی که در سطح متا مدل تولید شده توسط یک سیستم خود منبعی است که برای سیستم دیگر اعمال می شود [۵۸]. در اینجا نیازی به ارائه داده های اضافی با مدل نیست.

برای مقابله با مشکل شروع سرد ، مشکل پراکندگی و مشکل گوسفند خاکستری استفاده می شود. هزینه اجرا افزایش می یابد و افزایش پیچیدگی نیز یک فاکتور است.

۲,۴ رویکردهای مبتنی بر یادگیری عمیق

جفری لوند و ییو-کای نگ [۵۹] رویکرد مبتنی بر یادگیری عمیق را اتخاذ کردند که در آن نویسندگان از سیستم فیلترینگ مشارکتی مبتنی بر خود رمزگذار استفاده کرده اند. خطای پیش بینی با استفاده از قاعده مند سازی به حداقل رسیده است.

تکنیک های سنتی مبتنی بر محتوای سنتی و فیلترهای مشترک به اطلاعات قبلی مانند سابقه کاربر و عادت ها نیاز دارند. عملکرد چنین سیستم هایی بستگی زیادی به دانش اولیه درباره کاربر دارد. Kumar .S و همکاران [۶۰] روش ترکیبی مبتنی بر تجزیه و تحلیل احساسات را برای توصیه فیلم پیشنهاد کرد. از وب سایت های کوچک ، توئیت هایی در آیتم فیلم جمع آوری شده است و احساسات کاربران مورد تجزیه و تحلیل قرار گرفته و مشکل شروع سرد به خوبی برطرف شده است. جی کی لونگ و همکاران تأثیر خلق و خوی کاربر بر سیستم توصیه را مطالعه کرده اند. نویسندگان با تشخیص خودکار صفات عاطفی ، برچسب های احساسی را به فیلم ها اختصاص داده اند. مدل تشخیص عاطفه مبتنی بر متن با استفاده از متن توئیت شده آموزش داده شده است و از مدل آموزش‌داده شده به منظور شناسایی ویژگی های عاطفی استفاده می شود [۶۱].

ت سینگ و همکاران [۶۲] برای بهبود عملکرد سیستم توصیه از توییت های چند زبانه در زمان واقعی استفاده کرده اند. تجزیه و تحلیل احساسات در توییت ها اعمال می شود ، و توییت ها با استفاده از RNN طبقه بندی می شوند. این سیستم در توییت های زمان واقعی نتایج امیدوار کننده ای کسب می کند.

رویکرد مبتنی بر یادگیری عمیق توسط Bobadilla .J و همکاران ارائه شده است. [۶۳] برای بهبود عملکرد فیلتر کردن مشترک. رابطه غیر خطی بین پیش بینی ها، قابلیت اطمینان و توصیه دقیق با استفاده از یادگیری عمیق استخراج شده است. معماری پیشنهادی از طریق سه لایه انتزاع به عملکرد برتر دست می یابد: خطاهای واقعی پیش بینی ، خطاهای پیش بینی و رتبه بندی های پیش بینی شده.

۲,۵ مقایسه روشهای پیشرفته

تکنیک های مختلف کار بر روی فیلترینگ مشارکتی و فیلترینگ محتوا در بررسی انواع سیستم های توصیه گر بحث شده است. در این بخش خاص ، مقایسه برخی از تکنیک های اخیر مورد استفاده برای توصیه ، همانطور که در مطالعه تطبیقی نشان داده شده است ، بررسی می شود.

کارایی	ابزاهای مشخص کردن شباهت	روش	دیتا ست	CF/CB/H	مرجع
Cutoff = ۰,۵ برچسب های مفیدی بر جای می گذارد Cutoff = ۰,۳ مقداری از دست می دهد	تشابه کسینوسی	PCA	Movie lens	H	[۵۰]

برچسب های مهم					
تغییر رتبه از ۱۰ به ۸۰ KNN-HFB ، KNN-CFB عملکرد بهتری دارد		HFB-KNN ، CFB-KNN ، کاربر KNN	Movielens-1k	H	[۵۱]
AAD = ۰. دقت پیش بینی = ۱۰۰٪	همبستگی پیرسون	NNUN, Log Likely hood	از ، تحقیقات یاهو پایگاه داده دامنه وب	CF	[۲۲]
میزان پراکندگی مجموعه داده فیلم لنز ۹۳,۷٪ سطح پراکندگی است از یاهو! دامنه وب دامنه 4R ۹۹,۸٪ است		PNN ، SVD	Movie lens	CF	[۲۳]
HDFS عملکرد را بهتر می کند.		کاربر مبتنی بر CF ، براساس آیتمCF	Movielens-10M	CF	[۲۴]
MAE = ۰,۸۱۷ ، RMSE = ۱,۰۳۷	شباهت کسینوس	MF-BPR ، IDF-TF	Movielens	H	[۵۲]
برای movie lens ، دقت = ۰,۸۷۰۴ فراخوان = ۰,۱۰۷ برای فیلم اعتماد ، پیش cision = ۰,۷۳۹۱ فراخوان = ۰,۰۰۵۱		Biased Matrix Factorization	Movielens 100k,Film Trust	H	[۵۳]
توصیه می شود با استفاده از Map Hadoop Framework Reduce سریع انجام می شود اگرچه مجموعه داده بزرگ است.	تشابه کسینوسی ، تشابه کسینوسی مرکز قرارگرفته شده	Map Reduce	Movie lens	H	[۵۴]
فوق تخصص و پراکندگی داده ها بهبود یافته است		سیستم MN-CB ، FW	Movie lens -1M	CB	[۱۸]
مشکل تخصصی بزرگ حل شده و مشاهده می شود که وقتی هیچ کدام از آنها وجود ندارد خوشه دقت را کاهش می دهد.		SA	IMDB and BBC web server	CB	[۱۹]
Fuzzy-C و حداکثر میانگین دقت -۸۰,۴۴ و C Fuzzy-C و حداکثر دقت پیرسون -۸۱,۱	همبستگی پیرسون	K-means, SOM, Fuzzy clustring	Movie lens with 6.3% رتبه بندی موجود است	کاربر مبتنی بر CF	[۳۴]
شعاع بیشتر گره ها و لبه ضخیم تر و سپس امتیاز فیلم بالا است	فاصله ی اقلیدسی	Neo4j is NoSQL graph databas	Movielens-100K	کاربر مبتنی بر	[۳۵]

	CF				
[۳۷]	H	Movie lens	UBCF و IBCF	همبستگی پیرسون	بهبود ۱۳٪ نسبت به UBCF و ۱۶٪ بیش از IBCF
[۴۴]	CF	Movie lens	EMF، NMF و PMF	کسینوس	میانگین RMSE برای EMF مقدار 1.3411 میباشد
[۴۸]	CF	Movie lens	SVD با UBCF و IBCF	همبستگی جمعیتی	IBCF نسبت به UBCF مقدار خطای کمتری تولید می کند

۲,۶ معیارهای ارزیابی برای سیستم توصیه ها

اقدامات مختلف شباهت برای کشف شباهت بین کاربر و آیتم وجود دارد. برخی از آنها در اینجا آیتم بحث قرار گرفته است.

۲,۶,۱ *تشابه ژاکارد*

این نسبت اقلام معمولی است که توسط کاربر رتبه بندی می شود به تعداد کل مواردی که توسط هر دو کاربر رتبه بندی می شود. فرمول محاسبه شباهت ژاکارد در زیر آمده است [۶۶]

$$sim(u,v)^{jaccard} = \frac{|I_u \cap I_v|}{|I_u \cup I_v|} \quad (۵)$$

در اینجا I_u و I_v به ترتیب مجموعه آیتم‌هایی هستند که به ترتیب توسط کاربر u و v ربته بندی می شوند.

۲,۶,۲ *تشابه کسینوسی*

شباهت کسینوسی [۶۵] ، [۶۶] اندازه گیری تشابه بین دو بردار غیر صفر در فضای ضرب داخلی است. زاویه بین این دو بردار اندازه‌گیری می گردد. کسینوس از دو بردار غیر صفر را می توان با استفاده از ضرب نقطه ای این دو بردار محاسبه کرد:

$$u.v = ||u||.||v||.\cos\theta \quad (۶)$$

شباهت کسینوس به ویژه در فضای مثبت که در آن نتیجه به طور کارآمد در [۰ ، ۱] محدود می شود ، استفاده می شود. بدین ترتیب برای دو بردار داده شده u و v ، شباهت کسینوس ، θcos را می توان به عنوان ترکیبی از محصول نقطه و مقدار بردارها ، cosine ، کرد.

$$sim(\vec{u},\vec{v})^{cosine} = \cos\theta = \frac{\sum uv}{||u||.||v||} \quad (۷)$$

۲,۶,۳ *تشابه همبستگی پیرسون*

برای تعیین شباهت بین کاربران از ضریب همبستگی پیرسون [۶۷] ، [۶۸] استفاده می شود. هرچه ضریب این دو کاربر بیشتر باشد به هم نزدیک تر هستند. فرمول محاسبه ضریب همبستگی پیرسون در زیر آورده شده است ،

$$r = \frac{\sum((u-\bar{u})(v-\bar{v}))}{\sqrt{(u-\bar{u})^2.(v-\bar{v})^2}} \quad (۸)$$

۲,۶,۴ *فاصله مربع میانگین*

فاصله مربع میانگین [۶۸] با نسبت مجموع مربع اختلاف آیتم‌های رتبه بندی شده توسط کاربر به آیتم‌های مشترک که توسط هر دو کاربر رتبه بندی شده است محاسبه می شود. سپس میانگین شباهت مربع با کسر میانگین فاصله مربع توسط ۱ محاسبه می شود. فرمول محاسبه میانگین فاصله مربع در زیر آورده شده است.

$$sim(u,v)^{MSD} = 1 - \frac{\sum i \in I_{(u,v)}(R(u,i)-R(v,i))^2}{|I_{(u,v)}|} \quad (۹)$$

$R(u,i)$ و $R(v,i)$ به ترتیب امتیازاتی هستند که توسط کاربر u و v به آیتم i داده می شود. $I_{(u,v)}$ نشانگر امتیازات مشترک کاربران u و v.

۲,۶,۵ *فاصله میانگین مربع Jaccard (JMSD)*

این امر با ضرب دو معیار شباهت ، یعنی تشابه ژاکارد و معیار تشابه فاصله مربع [۶۸] حاصل می شود.

$$sim(u,v)^{JMSD} = sim(u,v)^{jacard}.sim(u,v)^{MSD} \quad (۱۰)$$

$$sim(u,v)^{JMSD} = \frac{|I_u \cap I_v|}{|I_u \cup I_v|}.1 - \frac{\sum i \in I_{(u,v)}(R(u,i) - R(v,i))^2}{|I_{(u,v)}|} \quad (۱۱)$$

۲,۶,۶ *شباهت مربوط به Jaccard (RJaccard)*

$$sim(u,v)^{RJaccard} = \frac{1}{1 + \frac{1}{|I_u \cap I_v|} + \frac{|\bar{I}_u|}{1+|I_u|} + \frac{1}{1+|\bar{I}_v|}} \quad (۱۲)$$

بطوریکه اگر $|I_u \cap I_v| = 0$ آنگاه $sim(u,v)^{RJaccard} = 0$ میگردد.

۲,۶,۷ *فاصله مربع مربوط به Jaccard (RJMSD)*

میانگین فاصله مربع Jaccard مربوطه با ضرب jaccard مربوطه و فاصله مربع متوسط بدست می آید. فرمول محاسبه RJMSD در زیر آورده شده است [۶۸].

$$sim(u,v)^{RJMSD} = sim(u,v)^{Rjacard}.sim(u,v)^{MSD} \quad (۱۳)$$

$$sim(u,v)^{RJMSD} = \frac{1}{1 + \frac{1}{|I_u \cap I_v|} + \frac{|\bar{I}_u|}{1+|I_u|} + \frac{1}{1+|\bar{I}_v|}}.1 - \frac{\sum i \in I_{(u,v)}(R(u,i) - R(v,i))^2}{|I_{(u,v)}|} \quad (۱۴)$$

بطوریکه اگر $|I_u \cap I_v| = 0$ آنگاه $sim(u,v)^{RJMSD} = 0$ میگردد.

۲,۷ معیارهای ارزیابی برای سیستم توصیه

اقدامات مختلفی برای ارزیابی سیستم پیشنهاد دهنده وجود دارد. این معیارها راهی برای پی بردن به دقت سیستم پیشنهادی است. برای کارکرد هر سیستم پیشنهادی دقت لازم است. عملکرد هر مدل با توجه به این معیارها ارزیابی می شود. معیارهای ارزیابی عبارتند از: خطای مطلق میانگین (MAE) ، خطای میانگین مربع ریشه (RMSE) ، دقت ، یادآوری ، اندازه گیری ۱F ، تنوع جمع.

۲,۷,۱ *خطای مطلق میانگین*

$$MAE = \frac{1}{N}\sum_{(u,i)} P_{(u,i)} - R_{(u,i)} \quad (۱۵)$$

در اینجا $P_{(u,i)}$ امتیاز پیش بینی شده برای کاربر u در آیتم i است ، $R_{(u,i)}$ امتیاز واقعی و N کل تعداد رتبه بندی مجموعه آیتم است. هرچه MAE کمتر باشد ، موتور توصیه با دقت بیشتری کاربر را پیش بینی می کند

رتبه بندی [۶۹].

۲,۷,۲ *خطای میانگین مربع ریشه*

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(P_{(u,i)} - R_{(u,i)})^2}{N}} \quad (۱۶)$$

خطای میانگین مربعات ریشه (RMSE) [۷۰] بیشتر بر خطای مطلق بزرگتر تأکید می کند و هنگامی که RMSE کم است آنگاه دقت سیستم پیشنهادی بهتر است.

۲,۷,۳ *دقت*

$$precision = \frac{currently\ recommended\ items}{Total\ recommended\ items} \quad (۱۷)$$

دقت [۷۱] نتیجه ای را نشان می دهد که مرتبط است، یعنی آیتم‌هایی که سیستم به درستی توصیه کرده است.

۲,۷,۴ *فراخوانی(Recall)*

$$Recall = \frac{currently\ recommended\ items}{Total\ useful\ recommended\ items} \quad (۱۸)$$

یادآوری [۷۱] نتایجی را نشان می دهد که با موفقیت توسط سیستم توصیه می شود.

۲,۷,۵ *اندازه گیری F1*

$$f1\ score = 2 \times \frac{Precision\ \times Recall}{Precision\ +\ Recall} \quad (۱۹)$$

برای یافتن F۱-Measure [۷۱] ابتدا باید دقت و فراخوانی را بیابیم.

۲,۷,۶ *تنوع*

$$Diversity\ = U_{u \epsilon U} L_n(u) \quad (۲۰)$$

در جایی که شما کاربر خاصی هستید ، U کل کاربر موجود در مجموعه داده و $L_n(u)$ لیست آیتم‌های مربوطه است که به کاربر u [۷۲] ، [۷۳] توصیه می شود.

۳ توضیحات مربوط DATASET

مجموعه داده واحدی برای سنجش اطلاعات منتشر شده در یک مخزن عمومی داده آزاد است. محققان از این مجموعه داده برای انجام آزمایشات استفاده می کنند. برای این منظور ، آنها مجموعه داده را به مجموعه آموزش و آزمایش تقسیم می کنند تا نتیجه مطلوب را بدست آورند. یک مجموعه داده مربوط به یک یا چند جدول پایگاه داده است که در آن هر ستون از یک جدول متغیر خاصی را نشان می دهد و هر سطر مربوط به یک رکورد داده شده است. مدل با استفاده از مجموعه آموزش اجرا می شود و سپس با نتیجه هدف مقایسه می شود. سپس داده های آزمایش برای ارائه برآورد مدل نهایی استفاده می شود. مجموعه فیلم های مختلفی برای فیلم ها در دسترس است که به طور گسترده در دسترس هستند. از مجموعه های داده مانند k۱۰۰Movielens ، M۱۰-Movielens ، k̄۱۰-Movielens و غیره استفاده می شود. مجموعه داده های دیگر عبارتند از IMDB ، مجموعه داده Netflix ، مجموعه داده Film Trust و غیره. مجموعه داده فیلم شامل فیلدهایی مانند شناسه کاربر ، شناسه آیتم ، رتبه بندی و غیره می‌باشد.

نام	فیلم ها	رتبه بندی	کاربران	برچسب ها	آدرس اینترنتی
[۵۰] k۱۰۰-Movielens	۹۱۱۵	۱،۰۰،۰۰۵	۶۷۱	۱۱۹۷	/movielens/datasets/org.grouplens//:http /k۱۰۰
Movielens [۵۴]	۲۷،۲۷۸	۹،۹۹۹،۹۹۹	۶۹۱۳۹	–	al/movielens/datasets/org.rouplensg//:http /test
[۷۴] M۱-Movielens	۴۰۰۰۰	۱۴ م	۱،۶۰،۰۰۰	۶۷۰،۰۰۰	/movielens/datasets/org.grouplens//:https
Yahoo Web scope (R4) [22]	۱۱،۹۱۵	۱،۱۱،۱۳۱	۷،۶۴۱	–	com.yahoo.xsandbo.webscope//:http
[۵۱] k۱۰۰-Movielens	۱۶۸۲	۱۰۰۰۰۰۰	۹۴۳	–	org.movielens//:http
M۱۵-Movielens	۱۷۰۰۰	۱۰ میلیون	۱،۳۸،۰۰۰	۴،۶۵،۰۰۰	/movielens/datasets/org.grouplens//:https /m۱۵
Movie lens Full	۵۸۰۰۰	۱،۸۰،۰۰۰	۶۰۰	۱،۱۰۰،۰۰۰	l/movielens/datasets/org.grouplens//:https /atest
k۱۰۰-Movielens	۱۷۰۰	۱۰۰۰۰۰۰	۱۰۰۰	–	/movielens/datasets/org.grouplens//:http /k۱۰۰
M۱۰-Movielens	۱۰۰۰۰	۱۰ میلیون	۷۱۰۰۰	۱۰۰۰۰۰۰	/movielens/datasets/org.grouplens//:http /m۱۰

۴ نتیجه گیری و اهداف آینده

در این مقاله انواع مختلف تکنیک های فیلتر کردن شرح داده شده است. موارد مختلف ، مزایا ، معایب نیز آیتم بحث قرار گرفته است. برای ساختن یک سیستم توصیه گر کارآمد ، ترکیبی از روشهای مختلف توصیه باید باشد. نتیجه گرفته می شود که با استفاده از ترکیبی از اندازه گیری شباهت می توان شباهت کاربری بهتری نسبت به استفاده از اندازه گیری شباهت منفرد ایجاد کرد و کارایی سیستم نیز افزایش می یابد. یکی از این واقعیت این است که اندازه گیری شباهت مانند RJMSD توسط نویسنده تکامل یافته است و تاکنون فقط در توصیه فیلم استفاده می شود. نویسنده همچنین نشان داد که این اندازه گیری شباهت از نظر پارامترهای کارآیی بهتر از هر اندازه دیگر است. اگر ویژگی های اضافی به فیلم اضافه کنیم ، دقت هر سیستم پیشنهادی می تواند بهبود یابد. بطور کلی، بیشتر مقالات ترکیبی از فیلترینگ مشترک و فیلترینگ مبتنی بر محتوا را نشان داده اند. با ترکیب روشها ، مشکلات مربوط به دو روش برطرف می شود. بنابراین فیلترینگ هیبریدی شناخته‌شده‌ترین تکنیک در هر سیستم پیشنهادی است. زیرا استفاده از این امر به ایجاد یک سیستم توصیه موثر کمک می کند.

زمینه های مختلفی از سیستم توصیه وجود دارد که قبلاً بحث شد. تکنیک های مختلف نیز مورد بحث قرار گرفته است که در توصیه به کاربرده می شود بنابراین ، دامنه هر سیستم توصیه کننده ساختن یک مدل به گونه ای است که کاربران توصیه مناسب را بدست آورد و کارایی سیستم نیز حفظ شود.منابع

[1] *Falk, K. (2019). Practical recommender systems (1st ed.). Shelter Island: Manning ISBN 1-61729-270-2 .*

[2] *Fortune, Amazon's recommendation secret, (2012), https://fortune.com/2012/07/30/ amazons- recommendation- secret/ .*

[3] *Gomez-Uribe, C. A. , & Hunt, N. (2015). The netflix recommender system: algorithms, business value, and innovation. ACM Transactions on Management Information Systems (TMIS), 6 (4), 1–19 .*

[4] P. N. V. Kumar and V. R. Reddy, "A Survey on Recommender Systems (RSS) and Its Applications", International Journal of Innovative Research in Computer and Communication Engineering 2(8):5254- 5260, 2014.

[5] C. A. Gomez-Uribe and N. Hunt, "The Netflix Recommender System: Algorithms, Business Value, and Innovation", ACM Transactions on Management Information Systems 6(4):1-19, 2015. DOI: http://dx.doi.org/10.1145/2843948

[6] Ò. Celma, Music Recommendation, Springer, Berlin, Heidelberg, 43-85, 2010. DOI: http://dx.doi.org/10.1007/978-3-642-13287-2 3

[7] F. Zhang, Z. Zheng, Y. Xiang, N. J. Yuan, X. Xie, and Z. Li "DRN: A Deep Reinforcement Learning Framework for News Recommendation", Proceedings of the World Wide Web Conference, 167-176, 2018. DOI: http://dx.doi.org/10.1145/3178876.3185994

[8] W.-C. Kang, C. Fang, Z. Wang, and J. McAuley, "Visually-aware fashion recommendation and design with generative image models", IEEE International Conference on Data Mining, 207-216, 2017.

[9] V. C. S. Heung, H. Qu, and R. Chu, "The relationship between vacation factors and socio-demographic and travelling characteristics: The case of Japanese leisure travellers", Tourism Management 22(3):259- 269, 2001.

[10] *J. Bobadilla, F. Ortega, A. Hernando, and A. Gutiérrez, "Recommender systems survey", Knowledge-*Based Systems 46(2):109-132, 2013. DOI: http://dx.doi.org/10.1016/j.knosys.2013.03.012

[11] M. Sharma and S. Mann, "A Survey of Recommender Systems: Approaches and Limitations", Interna- tional Journal of Innovations in Engineering and Technology 2(2):8-14, 2013.

[12] F. Mansur, V. Patel, and M. Patel, "A review on recommender systems", IEEE International Con- ference on Innovations in Information, Embedded and Communication Systems, 1-6, 2017. DOI: http://dx.doi.org/10.1109/ICIIECS.2017.8276182

[13] J. Salter and N. Antonopoulos, "CinemaScreen Recommender Agent:Combining Collaborative and Content-Based Filtering", IEEE Intelligent Systems 21(1):35-41, 2006.

[14] H. Li, F. Cai, and Z. Liao, "Content-based filtering recommendation algorithm using HMM", IEEE Fourth International Conference on Computational and Information Sciences, 275-277, 2012.

[15] A. Patel, A. Thakkar, N. Bhatt, and P. Prajapati, Survey and Evolution Study Focusing Comparative Analy- sis and Future Research Direction in the Field of Recommendation System Specific to Collaborative Filter- ing Approach, Springer, Singapore, 155-163, 2019. DOI: http://dx.doi.org/10.1007/978-981-13-1742-2 16

[16] J. Son and S. B. Kim, "Content-based filtering for recommendation systems using multiattribute net- works", Expert Systems with Applications 89:404-412, 2017.

[17] Y. B. Fernandez, J. J. P. Aris, A. G. Dolla, M. R. Cabrer, and M. L. Nores, "Providing Entertainment by Content-based Filtering and Semantic Reasoning in Intelligent Recommender Systems", IEEE Transac- tions on Consumer Electronics 54(2):727-735, 2008.

[18] D. Goldberg, D. Nichols, B. M. Oki, and D. Terry, "Using collaborative filtering to Weave an Information tapestry", Communications of the ACM 35(12):61-70, 1992. DOI: http://dx.doi.org/10.1145/138859.138867

[19] S. Khatwani and D. M. B. Chandak, "Building Personalized and Non Personalized Recommendation Systems", IEEE International Conference on Automatic Control and Dynamic Optimization Techniques , 623-628, 2016. DOI: http://dx.doi.org/10.1109/ICACDOT.2016.7877661

[20] C. M. Wu, D. Garg, and U. Bhandary, "Movie Recommendation System Using Collaborative Filtering", IEEE 9th International Conference on Software Engineering and Service Science , 11-15, 2018.

[21] M. Nilashi, O. Ibrahim, and K. Bagherifard, "A recommender system based on collaborative filtering using ontology and dimensionality reduction techniques", Expert Systems with Applications 92:507-520, 2018. DOI: http://dx.doi.org/10.1016/j.eswa.2017.09.058

[22] T. Zhou, L. Chen, and J. Shen, "Movie Recommendation System Employing the User-Based CF in Cloud Computing", IEEE International Conference on Computational Science and Engineering and IEEE Inter- national Conference on Embedded and Ubiquitous Computing , 2:46-50, 2017.

[23] F. O. Isinkaye, Y. O. Folajimi, and B. A. Ojokoh, "Recommendation systems: Prin- ciples, methods and evaluation", Egyptian Informatics Journal 16(3):261-273, 2015. DOI: http://dx.doi.org/10.1016/j.eij.2015.06.005

[24] T. Bai, J. R. Wen, J. Zhang, and W. X. Zhao, "A neural collaborative filtering model with interaction- based neighborhood", Proceedings of ACM on Conference on Information and Knowledge Management ,1979-1982, 2017.

[25] L. T. Ponnam, S. D. Punyasamudram, S. N. Nallagulla, and S. Yellamati, "Movie recommender system using item based collaborative filtering technique", IEEE International Conference on Emerging Trends in Engineering, Technology and Science , 1-5, 2016.

[26] J. Bobadilla, F. Serradilla, and A. Hernando, "Collaborative filtering adapted to rec- ommender systems of e-learning", Knowledge-Based Systems 22(4):261-265, 2009. DOI: http://dx.doi.org/10.1016/j.knosys.2009.01.008

[27] Xiaoyuan Su and T. M. Khoshgoftaar, "A Survey of Collaborative Filtering Techniques", Advances in artificial intelligence 2009:1-9, 2009.

[28] Delgado, Joaquin, and N. Ishii, "Memory-Based Weighted-Majority Prediction for Recommender Sys- tems", Res. Dev. Inf. Retr 1999.

[29] J. Pareek, M. Jhaveri, A. Kapasi, and M. Trivedi, "SNetRS: Social networking in Recommendation sys- tem", Springer, Berlin, Heidelberg, 195-206, 2013. DOI: http://dx.doi.org/10.1007/978-3-642-31552-7 21

[30] M. Deshpande and G. Karypis, "Item-based Top-N Recommendation Algorithms", ACM Transactions on Information Systems 22(1):143-177, 2004. DOI: http://dx.doi.org/10.1145/963770.963776

[31] D. Almazro, G. Shahatah, L. Albdulkarim, M. Kherees, R. Martinez, and W. Nzoukou, "A Survey Paper on Recommender Systems", arXiv preprint arXiv:1006.5278 2010.

[32] H. Koohi and K. Kiani, "User based Collaborative Filtering using fuzzy C-means", Measurement: Journal of the International Measurement Confederation 91:134-139, 2016.

[33] N. Yi, C. Li, M. Shi, and X. Feng, "Design and Implementation of Movie Recommender System Based on Graph Database", IEEE 14th Web Information Systems and Applications Conference , 132-135, 2017.

[34] C.-H. Piao, J. Zhao, and L. J. Zheng, "Research on entropy-based collaborative filtering algorithm and personalized recommendation in e-commerce", Service Oriented Computing and Applications 3(2):147- 157, 2009.

[35] G. Badaro, H. Hajj, W. El-Hajj, and L. Nachman, "A Hybrid Approach with Collaborative Filtering for Recommender Systems", IEEE 9th International Wireless Communications and Mobile Computing Conference , 349-354, 2013. DOI: http://dx.doi.org/10.1109/IWCMC.2013.6583584

[36] H. Liu, Z. Hu, A. Mian, H. Tian, and X. Zhu, "A new user similarity model to improve the accuracy of collaborative filtering", Knowledge-Based Systems 56:156-166, 2014.

[37] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Item-based collaborative filtering recommendation al- gorithms", Proceedings of the 10th International Conference on World Wide Web , 285-295, 2001. DOI: http://dx.doi.org/10.1145/371920.372071

[38] *M. Liphoto, C. Du, and S. Ngwira, "A Survey on Recommender Systems", IEEE International Conference on Advances in Computing and Communication Engineering , 276-280, 2016. DOI: http://dx.doi.org/10.1109/ICACCE.2016.8073761*

[39] *A. K. S. C. Pradhan and B. S. P. Mishra, "SVD based Privacy Preserving Recommendation Model using Optimized Hybrid Item-based Collaborative Filtering", IEEE International Conference on Communication and Signal Processing , 0294-0298, 2019.*

[40] *Y. Koren, R. Bell, and C. Volinsky, "Matrix Factorization Techniques For Recommender Systems", Com- puter 42(8):30-37, 2009. DOI: http://dx.doi.org/10.1109/MC.2009.263*

[41] *G. Takacs, I. Pilaszy, B. Nemeth, and D. Tikk, "Major components of the Gravity Recommendation System", Acm Sigkdd Explorations Newsletter 9(2):80-83, 2007.*

[42] *B. Abdollahi and O. Nasraoui, "Explainable Matrix Factorization for Collaborative Filtering", Pro- ceedings of the 25th International Conference Companion on World Wide Web , 5-6, 2016. DOI: http://dx.doi.org/10.1145/2872518.2889405*

[43] *D. Bokde, S. Girase, and D. Mukhopadhyay, "Matrix Factorization model in Collabora- tive Filtering algorithms: A survey", Procedia Computer Science 49:136-146, 2015. DOI: http://dx.doi.org/10.1016/j.procs.2015.04.237*

[44] *Y. Azar, A. Fiat, A. Karlin, F. McSherry, and J. Saia, "Spectral analysis of data", Confer- ence Proceedings of the Annual ACM Symposium on Theory of Computing , 619-626, 2001. DOI: http://dx.doi.org/10.1145/380752.380859*

[45] *M. E. Wall, A. Rechtsteiner, and L. M. Rocha, "Singular Value Decomposition and Principal Component Analysis", Springer, Boston, MA, 91-109, 2005.*

[46] *M. G. Vozalis and K. G. Margaritis, "Using SVD and demographic data for the enhance- ment of generalized Collaborative Filtering", Information Sciences 177(15):3017-3037, 2007. DOI: http://dx.doi.org/10.1016/j.ins.2007.02.036*

[47] *K. N. Jain, V. Kumar, P. Kumar, and T. Choudhury, "Movie recommendation system: Hybrid information Filtering System", Springer, Singapore, 677-686, 2018.*

[48] *S. M. Ali, G. K. Nayak, R. K. Lenka, and R. K. Barik, "Movie Recommendation System Using Genome Tags and Content-Based Filtering", Springer, Singapore, 85-94, 2018.*

[49] *F. Deng, P. Ren, Z. Qin, G. Huang, and Z. Qin, "Leveraging Image Visual Features in Content-Based Rec- ommender System", Scientific Programming 2018, 2018. DOI: http://dx.doi.org/10.1155/2018/5497070*

[50] *C. Yang, X. Chen, L. Liu, T. Liu, and S. Geng, "A Hybrid Movie Recommendation Method Based on Social Similarity and Item Attributes", International Conference on Sensing and Imaging. Springer, Cham* , 275-285, 2018.

[51] *Priscila Valdiviezo D'ıaz and J. Bobadilla, "A Hybrid Approach of Recommendation via Extended Matrix Based on Collaborative Filtering with Demographics Information", International Conference on Technol- ogy Trends. Springer, Cham , 384-398, 2018.*

[52] *R. Bharti and D. Gupta, "Recommending Top N Movies Using Content-Based Filtering and Col- laborative Filtering with Hadoop and Hive Framework", Springer, Singapore, 109-118, 2019. DOI: http://dx.doi.org/10.1007/978-981-13-1280-9 10*

[53] *M. Claypool, A. Gokhale, T. Miranda, P. Murnikov, D. Netes, and M. Sartin, "Combining Content-Based and collaborative filters in an Online Newspaper", Proceedings of ACM SIGIR Workshop on Recommender Systems , 1999.*

[54] *M. Braunhofer, V. Codina, and F. Ricci, "Switching hybrid for cold-starting context-aware recommender systems", Proceedings of the 8th ACM Conference on Recommender systems , 349-352, 2014. DOI: http://dx.doi.org/10.1145/2645710.2645757*

[55] *A. S. Lampropoulos, P. S. Lampropoulou, and G. A. Tsihrintzis, "A Cascade-Hybrid Music Recommender System for mobile services based on musical genre classification and personality diagnosis", Multimedia Tools and Applications 59(1):241-258, 2012. DOI: http://dx.doi.org/10.1007/s11042-011-0742-0*

[56] *A. M and A. Wasfi, "Collecting user access patterns for building user profiles and collaborative filtering",* Proceedings of the 4th international conference on Intelligent user interfaces , 57-64, 1998.

[57] *E. Cano and M. Morisio, "Hybrid Recommender Systems: A Systematic Literature Review", Intelligent Data Analysis 21(6):1487-1524, 2017.*

[58] *R. Burke, "Hybrid Systems for Personalized Recommendations", Springer, Berlin, Heidelberg, 133-152, 2003. DOI: http://dx.doi.org/10.1007/11577935 7*

[59] *J. Lund and Y. Ng, "Movie Recommendations Using the Deep Learning Approach", IEEE Inter- national Conference on Information Reuse and Integration, 47-54, Salt Lake City, UT, 2018. DOI: http://dx.doi.org/10.1109/IRI.2018.00015*

[60] *Kumar, Sudhanshu, Kanjar De, and Partha Pratim Roy, "Movie recommendation system using sentiment analysis from microblogging data", IEEE Transactions on Computational Social Systems 2020.*

[61] *Leung, John Kalung, Igor Griva, and William G. Kennedy, "Making Use of Affective Features from Media Content Metadata for Better Movie Recommendation Making", arXiv preprint arXiv:2007.00636 (2020).*

[62] *Singh, Tarana, Anand Nayyar, and Arun Solanki, "Multilingual Opinion Mining Movie Recommendation System Using RNN", First International Conference on Computing, Communications, and Cyber-Security, Springer, Singapore, 2020.*

[63] *Bobadilla, Jesus, Santiago Alonso, and Antonio Hernando, "Deep Learning Architecture for Collaborative Filtering Recommender Systems", Applied Sciences 10(7):24-41, 2020. DOI: http://dx.doi.org/10.3390/app10072441*

[64] *M. Ayub, M. A. Ghazanfar, M. Maqsood, and A. Saleem, "A Jaccard base similarity measure to im- prove performance of CF based recommender systems", IEEE International Conference on Information Networking , 1-6, 2018.*

[65] *T. Q. Lee, Y. Park, and Y. T. Park, "A similarity measure for collaborative filtering with implicit feedback",* International Conference on Intelligent Computing. Springer, Berlin, Heidelberg , 385-397, 2007.

[66] *S. Bansal, C. Gupta, and A. Arora, "User tweets based genre prediction and movie recommendation using LSI and SVD", IEEE Ninth International Conference on Contemporary Computing , 1-6, 2016. DOI: http://dx.doi.org/10.1109/IC3.2016.7880220*

[67] *S. M. Choi, S. K. Ko, and Y.S. Han, "A movie recommendation algorithm based on genrecorrelations",* Expert Systems with Applications 39(9):8079-8085, 2012.

[68] *S. Bag, S. K. Kumar, and M. K. Tiwari, "An efficient recommendation generation using relevant Jaccard similarity", Information Sciences 483:53-64, 2019.*

[69] *G. Shani and A. Gunawardana, "Evaluating Recommendation Systems", Springer, Boston, MA, 257-297, 2011. DOI: http://dx.doi.org/10.1007/978-0-387-85820-3 8*

[70] *D. Parra and S. Sahebi, "Recommender systems: Sources of knowledge and evaluation metrics", Springer,* Berlin, Heidelberg, 149-175, 2013. DOI: http://dx.doi.org/10.1007/978-3-642-33326-2 7

[71] *G. Carullo, A. Castiglione, A. De Santis, and F. Palmieri, "A triadic closure and homophily-based recommendation system for online social networks", World Wide Web 18(6):1579-1601, 2015. DOI: http://dx.doi.org/10.1007/s11280-015-0333-5*

[72] *G. Adomavicius and Y. Kwon, "Improving Aggregate Recommendation Diversity Using Ranking- Based Techniques", IEEE Transactions on Knowledge and Data Engineering 24(5):896-911, 2011. DOI: http://dx.doi.org/10.1109/TKDE.2011.15*

[73] *G. Adomavicius and Y. Kwon, "Optimization-based approaches for maximizing aggregate recommendation diversity", INFORMS Journal on Computing 26(2):351-369, 2014. DOI: http://dx.doi.org/10.1287/ijoc.2013.0570*

[74] *M. F. Aljunid and D. H. Manjaiah, Movie Recommender System Based on Collaborative Filtering Using Apache Spark, Springer, Singapore, 283-295, 2019. DOI: http://dx.doi.org/10.1007/978-981-13-1274-8 22*