



## ارائه یک سیستم توصیه گر مبتنی بر زمینه بر اساس رویکرد پالایش مشارکتی در فیلم

مهران صارمی نیا<sup>(1)</sup> - محمد رضا مجمع<sup>(2)</sup> - لاله طلائى<sup>(3)</sup>

(1) گروه مهندسی فناوری اطلاعات - دانشگاه آزاد اسلامی، واحد تهران جنوب

mehransareminia@gmail.com

(2) گروه مهندسی کامپیوتر، واحد پردیس، دانشگاه آزاد اسلامی، پردیس

m\_majma@pardisiau.ac.ir

(3) گروه مهندسی کامپیوتر - دانشگاه آزاد اسلامی واحد اهواز

laletalaie@gmail.com

**خلاصه:** زمان و مکان شخص در انتخاب یک فیلم عامل‌های بسیار مهمی هستند. علاوه بر آن در نظر گرفتن تغییرات جریان علاقه کاربر با گذشت زمان تأثیر به سزایی روی دقت توصیه فیلم می‌گذارد. شناسایی موقعیت جغرافیایی کاربران امکان بهره‌گیری از رویکردهای پالایش مبتنی بر زمینه را فراهم می‌نماید. مدل پیشنهادی با استفاده از اطلاعات زمینه مانند مکان کاربران در حین امتیازدهی به فیلم و در نظر گرفتن جریان علاقه کاربر به این معنی که فیلم‌هایی که به تازگی امتیازدهی شده‌اند و جدیدتر هستند، با اولویت بالاتری به کاربر فیلم توصیه می‌کند. این کار را روی مجموعه داده‌ی Movielens انجام دادیم و مکان کاربران خوشه‌بندی کردیم و جهت پیدا کردن کاربران همسایه با کاربر هدف از معیار شباهت کسینوسی استفاده کردیم. بدین ترتیب فیلم‌هایی که امتیاز بالاتری دارند و مطابق با سلیقه کاربر باشد در لیست توصیه قرار خواهند گرفت. نتایج ارزیابی‌های انجام شده نشان می‌دهد که رویکرد پیشنهادی در مقایسه با دو تکنیک توصیه گر ABC-KM و CAMR، به ترتیب دارای بهبود 3/1 درصدی و 2/25 درصدی از لحاظ معیار دقت پیش‌بینی می‌باشد.

**کلمات کلیدی:** سیستم توصیه گر، پالایش مشارکتی، زمینه آگاه، سیستم توصیه گر فیلم، زمینه زمان و مکان، جریان علاقه کاربر

### 1 - مقدمه

شخصی‌سازی می‌تواند بر پایه پروفایل کاربر، علایق کاربر یا نرخ آیت‌هایی که کاربر قبلاً مشاهده کرده باشد [2].

یکی از مهم‌ترین و پرستفاده‌ترین روش‌ها در سیستم‌های توصیه گر، پالایش مشارکتی می‌باشد که پایه و مبنای کار در بسیاری از راهکارهای دیگر می‌باشد. از مهم‌ترین موضوعات در سیستم‌های توصیه گر تولید پیش‌بینی‌های دقیق می‌باشد. سیستم‌های توصیه گر پالایش مشارکتی مبتنی بر آیت‌م پیش‌بینی را از طریق سابقه همسایگان آیت‌ها محاسبه می‌کنند و انتخاب صحیح نزدیک‌ترین همسایگان که بیشترین شباهت را با آیت‌م مورد نظر داشته باشند تأثیر زیادی بر افزایش دقت پیش‌بینی‌ها خواهند داشت [3].

مشکلات اساسی و مهم در این سیستم‌ها است که با عنوان شروع سرد شناخته می‌شود. البته این سیستم‌ها از مشکل دیگری نیز رنج می‌برند که پراکندگی داده‌ها می‌باشد. بدین معنی که اطلاعات در سیستم وجود دارد اما پراکنده هستند و نمی‌توان به درستی و با قطعیت

تعداد فیلم‌های آنلاین به سرعت در حال افزایش است و تعداد زیادی از وب‌سایت‌های فیلم ساخته شده، کاربران را با طیف وسیعی از فیلم‌ها با موضوعات مختلف مواجه کرده‌اند. چگونگی ارائه فیلم به کاربران با توجه به علاقه آن‌ها یک چالش در بهبود کیفیت خدمات وب‌سایت‌ها و کیفیت تجربه کاربران است. سیستم‌های توصیه گر یک روش مؤثر برای مقابله با این چالش است [1].

سیستم توصیه گر سیستمی است که پیشنهادهای خود را بر اساس آنچه کاربران علاقه‌مند هستند و آنچه در گذشته علاقه داشتند و نیز سلاقی که در پروفایل خود وارد کردند ارائه می‌دهد. هدف سیستم توصیه گر صرفه‌جویی در وقت و افزایش رضایت کاربر است. سیستم توصیه گر به کاربر کمک می‌کند اطلاعات صحیح را پیدا کند این

گفت که چه آیمتی قابل قبول تر است. اساس کار در سیستم‌های توصیه گر مبتنی بر پالایش مشارکتی تعیین میزان شباهت بین کاربران و آیمت‌ها می‌باشد تا بر اساس آن بتوان همسایگی را به دست آورد. روش پالایش مشارکتی به جمع‌آوری نظرات کاربران یعنی رتبه‌دهی آن‌ها به آیمت‌ها می‌پردازد. این نوع سیستم برای ارائه توصیه‌های مناسب به یک کاربر، ابتدا کاربران مشابه را تعیین کرده و سپس آیمت‌های مورد علاقه افراد مشابه را به کاربر پیشنهاد می‌دهد. اولین مزیت این روش نسبت به روش‌های مبتنی بر محتوا، این است که این الگوریتم نیازی به ارائه آیمت‌ها بر حسب ویژگی‌هایشان ندارد [4].

در سیستم توصیه گر زمینه آگاه، تأکید این سیستم‌ها بر اطلاعاتی از قبیل زمان، مکان، اطلاعات امنیتی و نیز پارامترهای روان‌شناختی، اجتماعی، فیزیکی و غیره می‌باشد. این اطلاعات را می‌توان به صورت صریح و یا با استفاده از روش‌های داده‌کاوی کسب نمود. بسیاری از رویکردهای موجود در سیستم‌های توصیه گر، بر روی توصیه‌ی مرتبط‌ترین عناصر به کاربران تمرکز می‌کنند و اهمیتی برای اطلاعات زمینه‌ای مانند زمان، مکان و همکاری یا شرکت کردن افراد دیگر در کار مورد نظر، قائل نمی‌شوند. با این حال، در بسیاری از برنامه‌های کاربردی توجه به اطلاعات زمینه‌ای در ارائه‌ی پیشنهاد حیاتی می‌باشد. برای مثال، با استفاده از زمینه زمان و مکان کاربر می‌تواند دقت و درستی توصیه‌ها را بالا ببرد بر طبق اکثر کاربران که در همان موقعیت جغرافیایی مشابه به هم واقع شده‌اند ممکن است علاقه مشابهی در مورد انتخاب فیلم داشته باشند. کاربرانی که در فلوریدا هستند بیشتر به فیلم‌های فانتزی و انیمیشن علاقه دارند اما کاربران از ایالت مینه سوتا بیشتر علاقه مند به دیدن ژانر جنگ و درام هستند. همچنین فراهم شدن اطلاعات زمینه همچون زمان و مکان برای رأی‌ها، اقلام و کاربران، سیستم‌های جدیدی شکل گرفتند که یا صرفاً از این اطلاعات برای توصیه استفاده می‌کنند یا این اطلاعات را برای ارتقای سیستم‌های پالایش مشارکتی با آن‌ها ترکیب می‌کنند [5]. سیستمی که به کاربر، فیلم‌های نزدیک‌تر و مناسب را توصیه می‌کند، اساساً یک سیستم توصیه گر مبتنی بر مکان خواهد بود. یا اگر یک شبکه‌ی خبری را در نظر بگیریم، این شبکه‌ی خبری هم بر مبنای مکان فرد سعی می‌کند اخبار مناسب را نمایش دهد و هم زمان را در نظر می‌گیرد تا تازگی خبر بر ارزش خبری آن برای مخاطب بیفزاید. نکته‌ی قابل توجه دیگر هم این است که سلیقه‌ی کاربر در طول زمان تغییر می‌کند و در نظر گرفتن زمان سیستم‌های توصیه گر را بهبود می‌بخشد.

سیستم‌های توصیه گر ترکیبی سعی دارند انواع روش‌های گزینش را به گونه‌ای با هم ترکیب کرده و استفاده کنند و مزیت‌های گزینش را به گونه‌ای با هم ترکیب کرده و استفاده می‌کنند که مزیت‌های هر کدام از آن‌ها را داشته و بر کاستی‌هایشان فائق آیند؛ در نتیجه معمولاً این نوع سامانه‌ها بهترین عملکرد را نسبت به انواع دیگر دارا می‌باشند، روش‌های مختلفی برای این ترکیب وجود دارد [6].

## 2- مروری بر ادبیات پیشین

در این بخش، به برخی تحقیقات انجام شده در زمینه‌ی سیستم‌های

توصیه گر پرداخته شده و در ادامه با معرفی مجموعه‌ای از سیستم‌های توصیه گر پالایش مشارکتی و سیستم‌های توصیه گر ترکیبی و زمینه آگاه مبتنی بر برجسب، زمینه مکان و زمان را بررسی کردیم.

**1-2 سیستم توصیه گر پالایش مشارکتی بر اساس پالایش اولیه**  
راهول کاتاریا در [7] بر روی سیستم‌های توصیه فیلم تمرکز کرد که هدف اصلی آن پیشنهاد دادن یک سیستم پیشنهاددهنده از طریق خوشه‌بندی داده‌ها و هوش محاسباتی بود.

او یک سیستم پیشنهادی جدید را مورد بحث قرار داد که با استفاده از الگوریتم بهینه‌سازی جستجوی کوکاکو از خوشه‌بندی k-means استفاده می‌کند که در مجموعه داده MovieLens کاربرد دارد. او در [8] یک سیستم پیشنهاددهنده ترکیبی پیشنهاد داد که با استفاده از الگوریتم خوشه‌بندی k-means با روش بهینه‌سازی بذر زنبورعسل مصنوعی (ABC) به داده‌های MovieLens اعمال شده است و به صورت سیستماتیک توصیف شده است.

مانجو کمار و همکاران در [9] یک سیستم پیشنهاددهنده برای پیشنهاد فیلم با نام MovieREC معرفی کردند. سیستم این اجازه می‌دهد تا یک کاربر انتخاب‌های خود را از مجموعه‌ای از ویژگی‌های مشخص داشته باشد و سپس یک لیست فیلم بر اساس وزن تجمعی از ویژگی‌های مختلف و با استفاده از الگوریتم K-means توصیه می‌شود. بر اساس ارزیابی‌های غیررسمی که آن‌ها بر روی یک مجموعه کوچک از کاربران انجام دادند، نتایج مثبت گرفتند. آن‌ها به عنوان کارهای آینده می‌خواهند روش خود را روی یک مجموعه داده بزرگ‌تر انجام دهند که نتایج معنی‌دارتری را با استفاده از سیستم آن‌ها ایجاد می‌کند. علاوه بر این، آن‌ها می‌خواهند الگوریتم‌های مختلف ماشین‌های یادگیری و خوشه‌بندی را در نظر بگیرند و نتایج تطبیقی را مطالعه کنند. در نهایت آن‌ها می‌خواهند یک رابط کاربری مبتنی بر وب را که دارای پایگاه داده کاربر است را اجرا کنند و مدل یادگیری را برای هر کاربر طراحی کنند.

## 2-2 سیستم توصیه گر زمینه آگاه:

اوتبالکو و ماریا آندرا در [10] سیستم شخصی توصیه گر زمینه آگاه توصیه‌ی فیلم که از رویکرد ترکیبی مبتنی بر اطلاعات زمینه و محتوا محور و پالایش مشارکتی ارائه داد. از نظرات و امتیازهای همسایه‌های کاربر هدف برای ساخت ماتریس کاربر-آیمت و ی کاهش اثر شروع سرد استفاده کرد. سیستم ابتدا زمینه جاری کاربر هدف را تعیین می‌کند، سپس کاربرانی با زمینه مشابه را برای کاربر هدف پیدا می‌کند و پس از آن امتیاز کاربران را بر اساس اطلاعات زمینه محاسبه می‌کند و در نهایت بر اساس امتیازهایی که کاربران همسایه داده‌اند توصیه را ارائه می‌کند. برای مدل کردن پروفایل کاربر و دریافت اطلاعات ضمنی از دو رابطه وزن دهی تعداد دفعات مصرف و طول عمر استفاده کردند به این ترتیب آیمت‌هایی که بیشترین استفاده را داشته‌اند و آیمتی که جدیداً امتیازدهی شده است، اهمیت بیشتری می‌یابد. زمینه‌ها را به دو صورت عددی یا مقوله‌ای تقسیم کردند و یک ماتریس شباهت بر اساس وزن زمینه‌های کاربران و کاربر هدف ایجاد نمودند سپس Top N کاربر مشابه را پیدا کردند و بر اساس آن نمره‌ی پیش‌بینی را محاسبه کردند.

## 3-2 سیستم توصیه گر مبتنی بر زمان

و سن و همکاران در [11] به بهینه‌سازی معیار شباهت و وزن دهی زمان پرداختند. سیستم‌های توصیه گر پالایش مشارکتی عواملی مانند تغییرات جریان علاقه کاربر در طول زمان را در نظر نمی‌گیرند. نویسندگان نشان دادند که جریان علاقه کاربر دقت توصیه‌ها را بالاتر می‌برد. با استفاده از تئوری روانشناسی منحنی فراموشی ایبینگ هاوس نشان دادند که فرایند فراموشی کاربر در طی زمان، معمولی و نامتعادل است. با رشد سریع تجارت الکترونیک بسیاری از برنامه‌های کاربردی برای مدت طولانی استفاده می‌گردد و نیز بسیاری از وب‌سایت‌ها بعد از گذشت زمان اعتبار اطلاعات و امتیازهای خود را از دست می‌دهند، یعنی اطلاعات آن‌ها قدیمی می‌شوند و برای برطرف کردن این مشکل و در نظر گرفتن جریان علاقه و سلیقه کاربران از معیار شباهت ترکیبی پیرسون با در نظر گرفتن میانگین وزن امتیازهای مشترک بین کاربر هدف و کاربر همسایه استفاده نمودند.

#### 4-2 سیستم توصیه گر مبتنی بر مکان

سوها المومن محامد در [12] یک سیستم توصیه گر زمینه محور برای شخصی‌سازی مکان در موبایل ارائه دادند که از الگوریتم ژنتیک به عنوان یک الگوریتم بهینه و نیز از پارامترهایی مانند (موقعیت کاربر فعلی، خلق کاربر) و شرایط آب‌وهوا و زمان روز برای توصیه استفاده نمودند. نویسندگان با جمع‌آوری امتیازها در مکان‌های مختلف و سپس خوشه‌بندی امتیاز مکان‌ها با الگوریتم DBSCAN و سپس تعیین مکان‌های محبوب در هر خوشه و نیز گرفتن آب‌وهوا و زمان در هر نقطه اقامت و دریافت فرکانس هوا و زمان روز در هر خوشه به پیش‌بینی فیلم می‌پردازد.

این سیستم پارامترهای کاربر را مانند (آب‌وهوا، زمان روز و خلق کاربر) می‌گیرد، به عنوان مثال، اگر خلق‌وخوی کاربر (ناراحت) باشد، کاربر ممکن است نیاز به رفتن به سینما برای تماشای یک فیلم کمدی داشته باشد، اگر آب‌وهوای آفتابی باشد، کاربر ممکن است نیاز به رفتن به یک محل باز داشته باشد یا در صورتی که آب‌وهوا بارانی باشد، کاربر نیاز به مکان بسته دارد. در نهایت، سیستم برای هر مکان و با توجه به هر کاربری و هر حالت زمینه، مکان را ارائه می‌دهد. روش‌های زیادی برای تعیین نرخ وجود دارد (الگوریتم ژنتیک و فاکتور ماتریس) نویسندگان از الگوریتم ژنتیک استفاده کردند. داده‌ها در سه بعد (کاربر، مکان، متن) ذخیره شدند الگوریتم ژنتیک برای مقادیر امتیاز مکان ناشناس است و مقادیر آن را پیش‌بینی می‌کند. الگوریتم ژنتیک بر اساس تابع گاما در 166 مکان، 182 کاربر و 12 حالت مختلف زمینه بکار گرفتند.

#### 4-3 مدل پیشنهادی

در این کار یک سیستم توصیه گر از ترکیب پالایش مشارکتی، زمینه آگاه و محتوا محور ارائه شده است. این سیستم با توجه به موقعیت جغرافیایی و سلیقه کاربر فیلم متناسب با علاقه کاربر توصیه می‌کند. رویکرد پالایش مشارکتی با استفاده از امتیازها به تنهایی نمی‌تواند نتایج دقیقی ارائه کند، علاوه بر آن از مشکلات سیستم‌های توصیه گر سنتی مانند پراکندگی داده‌ها و شروع سرد رنج می‌برد. در این کار ما با در نظر

گرفتن اطلاعات زمینه مانند مکان و زمان و پیش پالایش مکان کاربران و ترکیب رویکرد پالایش مشارکتی با زمینه آگاه تا حدی این مشکلات را برطرف کردیم. ساخت سیستم پیشنهادی و مراحل آن به شرح زیر است:

- پیش پالایش مکان کاربران با استفاده از خوشه‌بندی DBSCAN
- شناسایی خوشه هدف با توجه به مکان کاربر هدف با استفاده از رابطه (هاورسین)
- ساخت ماتریس کاربر - آیتم
- محاسبه شباهت کاربر هدف و پیدا کردن کاربران همسایه
- وزن دهی زمان امتیازدهی با استفاده از رابطه تابع‌نمایی فراموشی (بررسی جریان علاقه کاربر)
- محاسبه نمره پیش‌بینی
- مرتب‌سازی نزولی فیلم‌ها
- توصیه فیلم

#### 4-3-1 بهره‌گیری از پالایش اولیه مبتنی بر زمینه مکان

مکان کاربران مجموعه داده‌های Movielenz بر اساس طول و عرض جغرافیایی با الگوریتم DBSCAN خوشه‌بندی گردید. بدین ترتیب ما با خوشه‌بندی فیلم‌ها محاسبات را به صورت برون خط ۱ انجام دادیم و زمان محاسباتی سیستم را کاهش دادیم. الگوریتم DBSCAN یک الگوریتم خوشه‌بندی مبتنی بر تراکم، با استفاده از رویکرد بالا به پایین است و بر این اصل استوار است که خوشه‌ها، ناحیه‌هایی از فضای داده با چگالی زیاد هستند که توسط نواحی با چگالی کمتر از یکدیگر جدا شده‌اند.

#### 4-3-2 شناسایی خوشه مناسب جهت کاربر هدف

فاصله مکان کاربر هدف با مراکز خوشه‌ها با استفاده از رابطه هاورسین برای شناسایی نزدیک‌ترین خوشه محاسبه گردید. رابطه هاورسین در مطالعه هندسه کروی، به خصوص در محاسبه فاصله بین دو نقطه توزیع شده در خارج از کره کاربرد دارد.

#### 4-3-3 محاسبه شباهت و پیدا کردن کاربران همسایه

محاسبه شباهت با هدف پیدا کردن کاربران همسایه با کاربر هدف انجام می‌شود. معیارهای زیادی برای محاسبه شباهت بین کاربران معیارهای مختلفی وجود دارد؛ در اینجا از شباهت کسینوسی به علت عملکرد بهتر در ارائه نتایج استفاده شده است. محاسبه ماتریس شباهت برای کاربران نزدیک  $u$  و  $v$  مطابق با تاریخچه بازدیدها با فرض اینکه احتمال صفر شدن شباهت بین کاربران وجود ندارد از رابطه (3) به دست آمد. با این رابطه  $k$  نزدیک‌ترین همسایه را برای کاربر فعال پیدا کردیم. در رابطه (1) مقدار امتیاز کاربر  $u$  به فیلم  $i$  و امتیاز کاربر همسایه به آیتم  $i$  و نیز  $I$  مجموعه امتیازهای مشترک بین دو کاربر می‌باشد. بردارهای کاربر هدف و کاربر همسایه نیز به ترتیب با  $\vec{u}$  و  $\vec{v}$  نشان داده شده است.

$$\text{sim}(u, v) = \frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\| \cdot \|\vec{v}\|} = \frac{\sum_{i \in I} (r_{u,i}) \cdot (r_{v,i})}{\sqrt{\sum_{i \in I} (r_{u,i})^2 \cdot \sum_{i \in I} (r_{v,i})^2}} \quad (1)$$

### 3-4 بهره‌گیری از زمینه زمان

جریان علاقه از روزنه زمان یا توابع فراموشی برای یادگیری و پیگیری تغییر رفتار کاربر با گذشت زمان استفاده می‌کند. توجه به اینکه آخرین امتیاز داده شده به فیلم قدیم یا جدید می‌باشد تأثیر بسزایی در توصیه دارد. چه بسا کاربر فیلمی را در سال‌ها پیش به دفعات بازدید و امتیازدهی نموده باشد، اما با گذشت زمان و بر سر کار آمدن فیلم‌هایی باکیفیت‌تر و محبوب‌تر تمایل کاربر نسبت به بازدید از فیلم قبلی کاهش یابد. این ویژگی با یک تابع جدید به عنوان پارامتر وزن دهی زمان لحاظ گردید. تابع فراموشی نمایی  $Tweight(u, i)$  با رابطه (2) برای وزن دهی زمان امتیازهای کاربر  $u$  روی آیت  $i$  استفاده گردید که در آن ضریب  $\alpha$  تعیین کننده سرعت فراموشی است و به صورت پویا مقادیری بین صفر و یک می‌گیرد و مقدار بالای آن نشان‌دهنده سرعت بالای فراموشی کاربر است. مقدار  $T_{u,i}$  اختلاف زمانی که کاربر آیت  $i$  را نشانه‌گذاری کرده است با اولین زمان نشانه‌گذاری در سیستم و  $T_u$  مدت زمان استفاده کاربر از سیستم توصیه گر را تعیین می‌کند.

$$Tweight(u, i) = e^{-\alpha \frac{T_{u,i}}{T_u}} \quad (2)$$

### 3-5 تعیین مقدار آستانه امتیاز و مرتب‌سازی آیت‌ها

جهت پیش‌بینی امتیاز و توصیه فیلم به کاربر هدف، نظرات کاربران همسایه و نزدیک‌ترین مکان به کاربر لحاظ گردید. سرانجام نمره پیش‌بینی که کاربر هدف  $u$  به فیلم  $i$  با توجه به مدل‌سازی‌های صورت گرفته از رفتار کاربران سیستم می‌دهد با رابطه (3) محاسبه گردید. وزن زمان تابع فراموشی با  $Tweight(v, i)$  برای کاربر  $v$  و  $sim(u, v)$  شباهت کاربر هدف با کاربر همسایه  $v$  نشان داده شده است. نمره پیش‌بینی  $score(u, i)$  بین صفر و یک می‌باشد.

$$score(u, i) = \frac{\sum_{v \in Neighbor(u)} Tweight(v, i) \times sim(u, v)}{\sum_{v \in Neighbor(u)} sim(u, v)} \quad (3)$$

در انتهای این بخش و پس از پیش‌بینی امتیازهای موجود در لیست، با توجه به علاقه و ترجیحات کاربر هدف، آیت‌هایی که از مقدار آستانه بالاتر باشند جهت توصیه در لیست به صورت نزولی مرتب می‌گردند. آستانه امتیاز برای تشخیص بین آیت‌های مورد علاقه استفاده می‌شود. آستانه خاص برای تجزیه و تحلیل مربوطه برای انتخاب اندازه امتیاز برای یک استقرار خاص مورد نیاز است. فیلم‌هایی با نمره امتیاز بالاتر و نیز بیشتر از آستانه 0/75 در بالای لیست پیشنهادی قرار گرفتند و به کاربر هدف توصیه گردید.

### 4- ارزیابی و مقایسه کارایی سیستم پیشنهادی

عملکرد سیستم پیشنهادی با عملکرد دو مدل توصیه معرفی شده در مقاله (ABC-KM) [8] و (CAMR) [10] مقایسه شده است. ایده اصلی بکار رفته در مقاله [8] در نظر گرفتن رویکرد پالایش مشارکتی و خوشه‌بندی kmeans جهت پیش‌بینی فیلم و بهینه‌سازی زنبور عسل می‌باشد. در مقابل، اساس کار مقاله [10] وزن زمینه‌های مختلف با استفاده از رویکرد زمینه آگاه را بررسی کردند. برای ارزیابی سیستم

پیشنهادی از معیارهای پیش‌بینی دقت و فراخوانی استفاده کردیم.

دقت: نشان دهنده این است که چگونه انتخاب آیت پیش‌بینی شده موفق بوده است. به بیان ساده‌تر نسبت پیشنهادهای درست به کل داده‌های پیشنهاد شده توسط سیستم است که در رابطه (4) آمده است.

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

فراخوانی: توانایی سیستم پیشنهاددهنده برای تولید تمام پیشنهادهایی که توسط کاربر دیده می‌شود در رابطه (5) را فراخوانی گویند. در تعریف دیگر فراخوانی نسبت پیشنهادهای درست تشخیص داده شده به تعداد کل داده‌های مورد انتظار برای پیش‌بینی می‌باشد.

$$recall = \frac{TP}{TP+FN} \quad (5)$$

Fmeasure: به علت موازی عمل کردن این معیارها، برای جمع آن‌ها باید از معیار F رابطه (6) استفاده کرد. این معیار میانگین هارمونیک دقت و فراخوانی است.

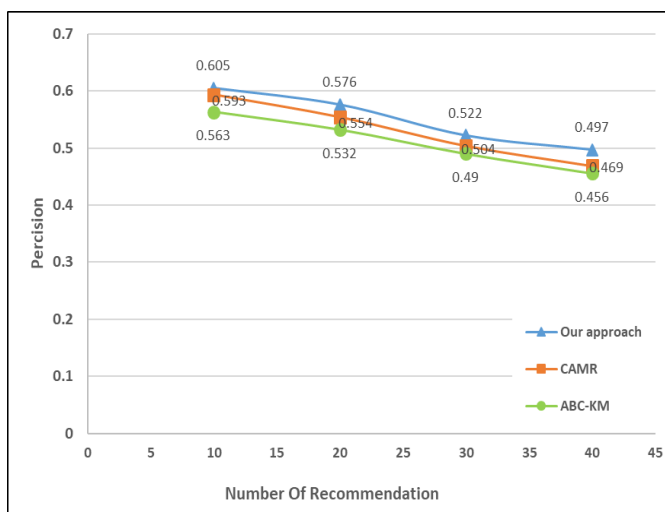
$$F_{measure} = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (6)$$

مثبت واقعی TP: بیانگر تعداد رکوردهایی است که دسته واقعی آن‌ها مثبت بوده و الگوریتم دسته‌بندی نیز دسته آن‌ها را به درستی مثبت تشخیص داده است.

منفی‌های واقعی TN: تعداد رکوردهایی است که دسته واقعی آن‌ها منفی بوده و الگوریتم دسته‌بندی نیز دسته آن‌ها را به درستی منفی تشخیص داده است.

مثبت کاذب FP: بیانگر تعداد رکوردهایی است که دسته واقعی آن‌ها منفی بوده و الگوریتم دسته‌بندی دسته آن‌ها را به اشتباه مثبت تشخیص داده است.

منفی‌های کاذب FN: بیانگر تعداد رکوردهایی است که دسته واقعی آن‌ها مثبت بوده و الگوریتم دسته‌بندی دسته آن‌ها را به اشتباه منفی تشخیص داده است.



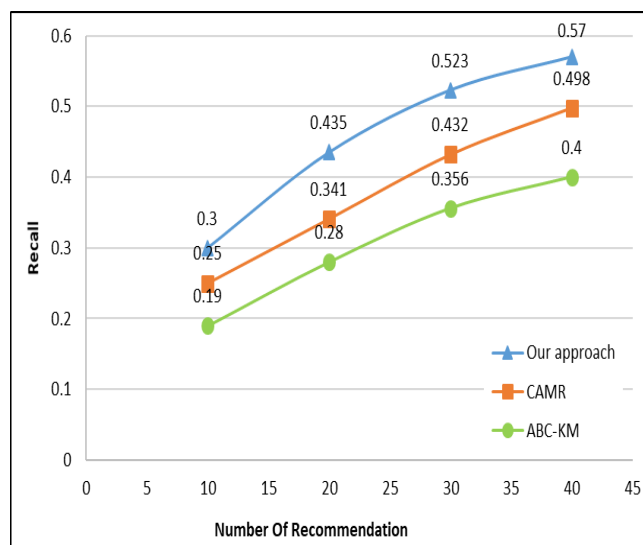
شکل (1): دقت برای Top N توصیه

در شکل (3) مدل پیشنهادی با معیار F که میانگین هارمونیک دقت و فراخوانی است نیز مقایسه شده است.

## 5- نتیجه‌گیری

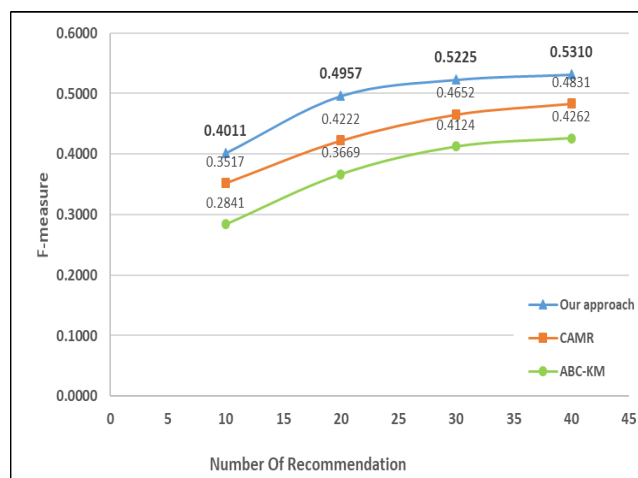
در این مقاله، سیستم توصیه گر ترکیبی از رویکرد پالایش مشارکتی زمینه آگاه و محتوا محور جهت توصیه فیلم با استفاده از مجموعه داده MovieLense ارائه کردیم. با پالایش اطلاعات زمینه مکان، موقعیت جغرافیایی کاربران را در نظر گرفتیم و امتیازهای کاربران نزدیک به کاربر هدف را در نظر گرفتیم. جهت شناسایی مکان‌های نزدیک به کاربر هدف و کاهش زمان پاسخگویی سیستم از خوشه‌بندی DBSCAN استفاده شده است. سیستم پیشنهادی، فیلم‌های مطابق با سلیقه و علاقه کاربر در طول مدت استفاده از سیستم و بر اساس نظرات مشابه کاربران همسایه به کاربر هدف توصیه می‌کند. مهم‌ترین دستاورد این پژوهش ارائه یک روش جدید جهت شخصی‌سازی وب و تولید توصیه‌های فیلم منطبق با علایق و شرایط زمینه‌ای درخواست از سوی کاربران و همچنین در نظر گرفتن جریان علاقه کاربر در طول زمان استفاده از سیستم، میزان مؤثر بودن آن لحاظ معیارهای متداول اندازه‌گیری کارایی سیستم‌های توصیه گر، یعنی معیار دقت و فراخوانی اثبات گردیده است. به منظور انجام کارهای آینده موارد زیر پیشنهاد می‌گردد از رویکردهای دیگر توصیه گر مانند جمعیت شناختی و دانش‌محور به منظور حل مشکل شروع سرد استفاده کرد.

شکل (1) نتایج عملکرد سیستم پیشنهادی را نسبت به دو تکنیک توصیه گر (ABC-KM) [8] و (CAMR) [10] بر حسب معیار اندازه‌گیری دقت نشان می‌دهد. همان‌طور که در شکل (1) ملاحظه می‌گردد، مطابق با انتظار با افزایش تعداد پیشنهادهای ارائه شده از سوی سیستم، دقت پیش‌بینی در تمامی روش‌های مورد بررسی کاهش یافته است. با این حال، روش پیشنهادی در اکثر موارد نسبت به سایر روش‌ها، عملکرد بهتری از خود نشان داده است. از لحاظ معیار دقت پیش‌بینی، الگوریتم پیشنهادی این پژوهش، توانسته است کارایی سیستم توصیه گر را به میزان 3/1 درصد نسبت به ABC\_KM [8] و 2/25 درصد نسبت به روش CAMR [10] بهبود بخشد.



شکل (2): فراخوانی برای Top N توصیه

شکل (2) نتایج عملکرد سیستم پیشنهادی را بر حسب معیار اندازه‌گیری فراخوانی نشان داده شده است. در این نمودار نیز مطابق انتظار، با افزایش تعداد پیشنهادهای ارائه شده از سوی سیستم، معیار فراخوانی در تمام الگوریتم‌های مورد بررسی افزایش یافته است.



شکل (3): معیار F برای Top N توصیه

- [1] Li, J. Xu, W. Wan, W. and Sun, J, "Movie recommendation based on bridging movie feature and user interest", *Journal of Computational Science*, vol. 26, no.3,2018, pp.128-134.
- [2] M. Ashrafa, M. Lalitha, D. Radha, and R. Jayaparvathy, "Analysis on Increasing Customer Sales by the Use of Restaurant Recommender System", *Analysis*, vol.139, 2016, pp.100-128.
- [3] F. Isinkaye, Y. Folajimi, and B. Ojokoh, "Recommendation systems: Principles, methods and evaluation", *Egyptian Informatics Journal*, vol.16, no.3, 2015, pp.261-275.
- [4] Cover, T. & Hart, P, "Nearest Neighbor Pattern Classification", *IEEE Transactions on Information Theory*, Vol.13, No.1, 1967, PP.21–27.
- [5] A.K.Tvete and M. Pettersen, "A Hybrid Recommender System for Context-Aware Recommendations of Restaurants", master's Thesis, NTNU, 2016, PP.101 pages.
- [6] LE, Tan Quach, "A Context Aware Recommendation System for movies", Master's Thesis. NTNU, 2017, pp.90 pages.
- [7] Katarya, R. and Verma, O.P, An effective collaborative movie recommender system with cuckoo search. *Egyptian Informatics Journal*, vol.18, no.2, 2017, pp.105-112.
- [8] Katarya, R, "Movie recommender system with metaheuristic artificial bee", *Neural Computing and Applications*, 2018, pp.1-8.
- [9] Kumar, M. Yadav, D.K. Singh, A. and Gupta, V.K. A movie recommender system: Movrec. *International Journal of Computer Applications*, vol.124, no.3, 2015, pp.1-10.
- [10] A. M. Otebolaku and M. T. Andrade, "Context-aware personalization using neighborhood-based context similarity", *Wireless Personal Communications*, vol.94, no.3, 2017, pp.1595-1618.
- [11] W. Sen, Z. Xiaonan, and D. Yannan, "A collaborative filtering recommender system integrated with interest drift based on forgetting function", *Int. J. u-and e-Serv. Sci. Technol*, vol.8, no.4, 2015, pp.247-264.
- [12] S.A. Mohamed, T. H.A. Soliman, and A. A. Sewisy, 2016, "A context-aware recommender system for personalized places in mobile applications", *Int. J. Adv. Comput. Sci.Appl*, vol. 7, no.3, pp.442-448.