Journal of Information and Communication Technology

Volume 17, Issue 66-67, Autumn and Winter 2025, pp. 198-237

A Review of Multimedia Recommender Systems

Saeedeh Momtazi[1*], Zahra Pourbahman[2], Mohammad Reza Azizi[3], Meysam Bagheri[4]

[1] Amirkabir Uiversity of Technology, Tehran, Iran
[2] Shahed University, Tehran, Iran
[3] Amirkabir Uiversity of Technology, Tehran, Iran
[4] Apasay Data System, Tehran, Iran

Received: 28 January 2026, Revised: 11 February 2026, Accepted: 12 February 2026
Paper type: Review

Abstract

The purpose of this study is to comprehensively review multimedia recommender systems with a focus on content-based approaches. The diversity and abundance of data on the web has prompted experts to research and develop recommender systems for automatically predicting the user's favorite content. In the upcoming research, various types of recommender systems were introduced, which are collaborative, content-based, and hybrid. Then, a review of 47 articles centered on content-based recommender systems in the three areas of movie, music, and television program recommendations was undertaken. In addition, the approaches used in recommender systems used in large and famous companies such as Netflix, YouTube, Facebook, Amazon, and TikTok were examined in detail. Due to the importance of the body used in recommender systems, MMTF-14K, MovieLens and Spotify Audio Features datasets were also introduced. According to the conducted research, it is clear that many practical recommender systems in large companies have benefited from the content-based approach for recommendations.

Keywords: Content-based recommender systems, industry recommender systems, datasets for content-based recommender systems

[*] Corresponding Author's email: momtazi@aut.ac.ir

دوفصلنامه
فناوری اطلاعات و ارتباطات ایران

سال هفدهم، شماره‌های ۶۶ و ۶۷، پاییز و زمستان ۱۴۰۴، صفحه ۱۹۸ الی ۲۳۷

مطالعه مروری بر سیستم‌های توصیه‌گر حوزه چندرسانه‌ای

سعیده ممتازی[1*]، زهرا پوربهمن[2]، محمدرضا عزیزی[3]، میثم باقری[4]

[1]دانشکده مهندسی کامپیوتر، دانشگاه صنعتی امیرکبیر، تهران، ایران

[2]دانشکده علوم پایه، گروه علوم کامپیوتر، دانشگاه شاهد، تهران، ایران

[3]دانشکده مهندسی کامپیوتر، دانشگاه صنعتی امیرکبیر، تهران، ایران

[4]شرکت آپاسای داده سیستم، تهران، ایران

تاریخ دریافت: ۱۴۰۴/۱۱/۰۸ تاریخ بازبینی: ۱۴۰۴/۱۱/۲۲ تاریخ پذیرش: ۱۴۰۴/۱۱/۲۳

نوع مقاله: مروری

چکیده

پژوهش حاضر یک مقاله مروری محسوب می‌شود که به‌مرور جامعی بر سیستم‌های توصیه‌گر چندرسانه‌ای با تمرکز بر رویکردهای مبتنی بر محتوا می‌پردازد. تنوع و کثرت داده‌ها در وب، متخصصان را بر آن داشته است که برای پیش‌بینی خودکار محتوای موردعلاقه کاربر به پژوهش و توسعه سیستم‌های توصیه‌گر بپردازند. در پژوهش پیش رو به معرفی انواع سیستم‌های توصیه‌گر، اعم از پالایش مشارکتی، مبتنی بر محتوا و ترکیبی، پرداخته شده است. سپس ۴۷ مقاله با محوریت سیستم‌های توصیه‌گر مبتنی بر محتوا در سه بخش توصیه فیلم، موسیقی و برنامه تلویزیونی بررسی گردیده است. به‌علاوه، رویکردهای مورداستفاده در سیستم‌های توصیه‌گر مورداستفاده در شرکت‌های بزرگ و مشهور جهان مانند نتفلیکس، یوتیوب، فیسبوک، آمازون و تیک‌تاک به‌تفصیل مورد بررسی قرار گرفته است. با توجه به اهمیت پیکره مورداستفاده در سیستم‌های توصیه‌گر به معرفی مجموعه دادگان MMTF-14K, MovieLens و Spotify Audio Features نیز پرداخته شده است. با توجه به پژوهش انجام‌شده، واضح است که اکنون بسیاری از سیستم‌های توصیه‌گر کاربردی در شرکت‌های بزرگ از رویکرد مبتنی بر محتوا برای توصیه بهره گرفته‌اند.

کلیدواژگان: سیستم‌های توصیه‌گر مبتنی بر محتوا، سیستم‌های توصیه‌گر در صنعت، پیکره دادگان جهت سیستم‌های توصیه‌گر مبتنی بر محتوا.

* رایانامهٔ نویسنده مسؤول: momtazi@aut.ac.ir

۱- مقدمه

امروزه داده‌های موجود در وب و ارائه‌دهندگان محتوا چندین نوع رسانه مختلف ازجمله متن، صدا، ویدئو و تصاویر را در اختیار کاربران قرار می‌دهند. فراوانی این نوع داده‌ها چنان گسترده است که تنها از طریق سیستم‌های توصیه‌گر چندرسانه‌ای[1] قابل‌دسترسی است [۱].

به‌غیراز کثرت تنوع، فایل‌های ویدئویی ازلحاظ صدا، محتوا و ساختار نیز گسترده شده است؛ به‌عنوان‌مثال فیلم‌های تولیدشده توسط کاربر، فیلم‌های تولیدشده توسط موسسات مختلف، کلیپ‌های ویدئویی موسیقی و غیره. پیدا کردن مطالب جدید جالب با استفاده از ابزارهای جستجوی سنتی برای کاربران دشوارتر شده است. درنتیجه، سیستم‌های توصیه‌گر که به‌طور خودکار محتوایی را که ممکن است کاربر دوست داشته باشد پیش‌بینی می‌کنند، در دهه گذشته ظهور و تکامل‌یافته‌اند [۲].

اکثر الگوریتم‌های سیستم‌های توصیه‌گر چندرسانه‌ای با استفاده از پالایش مبتنی بر محتوا[2] بر اساس داده‌های متنی، مانند فراداده در تعاملات کاربر، کار می‌کنند. بااین‌حال محتوای یک اثر چندرسانه‌ای را می‌توان به روش‌های متنوع‌تری توصیف کرد. ویژگی‌های برجسته یک فیلم شامل سبک، بازیگران و صحنه‌پردازی است که در محتوای دیداری و شنیداری آن منعکس‌شده است. برای یک قطعه موسیقی، سبک، ریتم، ساز، شعر و همچنین فاکتورهای فرهنگی توصیفات مهمی هستند. هنوز هم ویژگی‌های فراداده بیشترین استفاده را در سیستم‌های توصیه‌گر امروزی دارند [۱].

از سوی دیگر با توجه به ایجاد شبکه‌های تلویزیونی گوناگون مانند شبکه‌های خانگی یا تلویزیون‌های آی‌پی و غیره که نوع خاصی از تلویزیون هستند و خدمات متفاوتی به کاربران ارائه می‌کنند، در این شبکه‌های جدید پخش همگانی برنامه‌ها وجود ندارد و در عوض کاربران مختلف بر اساس سلیقه و زمان آزاد خود برنامه‌هایی را انتخاب و مشاهده می‌کنند. در این سیستم‌ها هر فرد یک پروفایل دارد که اطلاعاتی از خود در آن ذخیره می‌کند. به‌غیراز این اطلاعات صریح اطلاعات دیگری بر اساس فعالیت‌های کاربر بر روی این کانال‌ها استخراج می‌شود و سلیقه و تمایلات او در ارتباط با برنامه‌های تلویزیونی مشخص می‌شود. این پیشنهاد برنامه در حوزه سیستم‌های توصیه‌گر جای می‌گیرد.

با توجه به اهمیت سیستم‌های توصیه‌گر در فضای رسانه‌ای و همچنین ارزش روش‌های مبتنی بر محتوا در این حوزه، در این پژوهش مطالعه مروری جامعی بر روی سیستم‌های توصیه‌گر چندرسانه‌ای با تمرکز بر رویکردهای مبتنی بر محتوا ارائه می‌گردد.

در ادامه این مقاله، در بخش ۲، به تعریف ریاضیاتی سیستم‌های توصیه‌گر می‌پردازیم. در بخش ۳، به معرفی انواع سیستم‌های توصیه‌گر خواهیم پرداخت. مروری بر سیستم‌های توصیه‌گر موجود با تمرکز بر اهداف این پژوهش در بخش ۴ ارائه خواهد شد. در بخش ۵، به تبیین سیستم‌های توصیه‌گر مطرح در صنعت بین‌الملل می‌پردازیم. پیکره‌های آموزشی مورداستفاده در این سیستم‌ها در بخش ۶ مطرح می‌شود. بخش‌های ۷ و ۸ به ترتیب بحث و مقایسه و خلاصه و جمع‌بندی مطالب اختصاص خواهد یافت. در بخش آخر، منابع مورداستفاده در تحقیق پیش رو ارائه می‌شود.

۲- تعریف ریاضیاتی سیستم‌های توصیه‌گر

به‌طور کلی، یک سیستم توصیه‌گر با مجموعه‌ای از کاربران و مجموعه‌ای از اقلام سروکار دارد. برای هر جفت (ui, oj)، توصیه‌کننده می‌تواند امتیاز j را محاسبه کند که میزان بهره مورد انتظار کاربر ui در مورد oj (یا سودمندی مورد انتظار قلم oj برای کاربر) را اندازه‌گیری می‌کند. محاسبه امتیاز با استفاده از الگوریتم رتبه‌بندی انجام می‌شود که به‌طورکلی می‌تواند ترکیب مختلفی از ویژگی‌های زیر را در بر بگیرد: [۴]

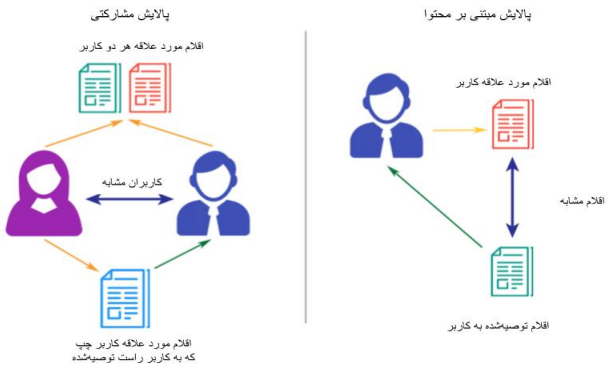
- تنظیمات کاربر و رفتار گذشته
- تنظیمات و رفتار جامعه کاربر
- ویژگی اقلام و چگونگی مطابقت با تنظیمات کاربر
- بازخورد کاربر،
- اطلاعات زمینه و چگونگی تغییر پیشنهادات همراه با محتوا.

۳- مرور انواع سیستم‌های توصیه‌گر: چالش‌ها و افق‌های آینده

به‌طور عمده سه نوع سیستم توصیه‌گر وجود دارد (شکل ۱): پالایش مشارکتی، مبتنی بر محتوا، ترکیبی. در این بخش به توضیح این رویکردها پرداخته می‌شود.

[1] Multimedia Recommender System(MMRS)

[2] Content Based Filtering (CBF)



شکل ۱. دو نوع از سیستم‌های توصیه‌گر؛ مشارکتی و محتوایی. [۱۸]

۳-۱- پالایش مشارکتی

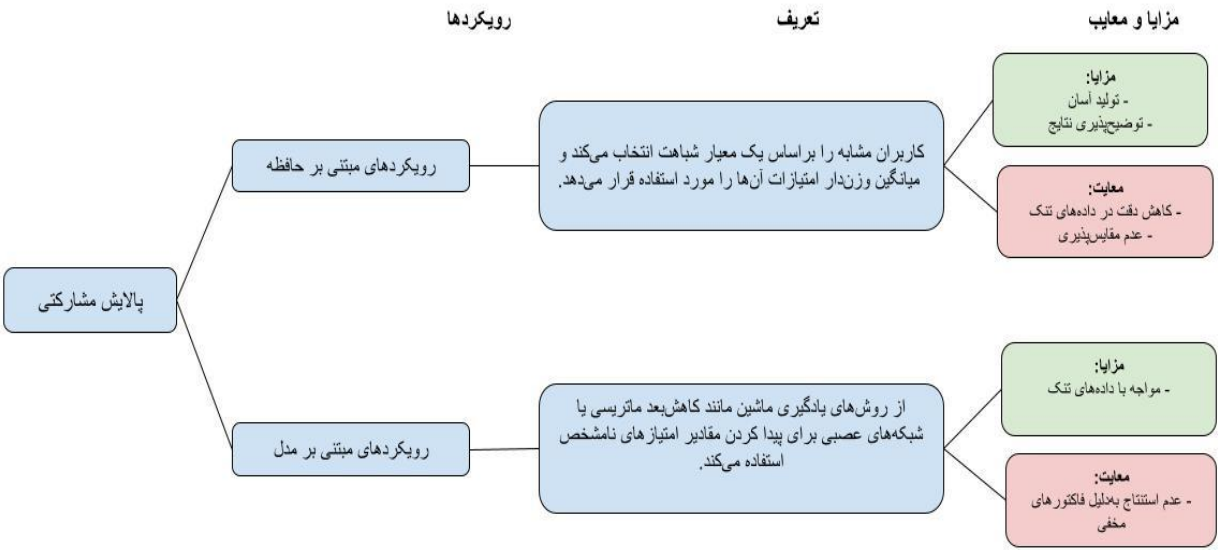
پالایش مشارکتی[1] یکی از متداول‌ترین و موفق‌ترین رویکردها در سیستم‌های توصیه‌گر است. ایده اصلی این روش بر این فرض استوار است که کاربرانی که درگذشته سلیقه مشابهی داشته‌اند، در آینده نیز به احتمال زیاد اقلام مشابهی را خواهند پسندید. سیستم‌های CF با جمع‌آوری و تحلیل رفتار گذشته کاربران (مانند امتیازدهی، خرید، یا مشاهده) الگوهای مشترک را شناسایی کرده و بر اساس این الگوها، توصیه‌های جدیدی برای کاربر فعال تولید می‌کنند.

این رویکرد معمولاً از یک ماتریس کاربر-قلم(ماتریس سودمندی) استفاده می‌کند که در آن سطرها نمایانگر کاربران و ستون‌ها نمایانگر اقلام هستند. درایه‌های این ماتریس نشان‌دهنده تعامل کاربر با قلم است که می‌تواند به‌صورت بازخورد صریح (مانند امتیاز ۱ تا ۵) یا بازخورد ضمنی (مانند تعداد دفعات کلیک، مشاهده یا خرید) باشد. مطابق با شکل ۲ دو رویکرد کلی وجود دارد که عبارت هستند از رویکرد مبتنی بر حافظه و رویکرد مبتنی بر مدل.

رویکرد مبتنی بر حافظه[2]

این رویکرد مستقیماً از کل داده‌های ماتریس کاربر-قلم برای پیش‌بینی استفاده می‌کند. دو زیرشاخه اصلی آن عبارت‌اند از:

- پالایش مشارکتی کاربر-محور :[3] ابتدا کاربران مشابه با کاربر فعال شناسایی می‌شوند و سپس اقلامی که این کاربران مشابه پسندیده‌اند (و کاربر فعال هنوز با آن‌ها تعامل نداشته) به کاربر فعال توصیه می‌شوند.
- پالایش مشارکتی قلم-محور [4]: در این روش، شباهت بین اقلام بر اساس الگوی امتیازدهی کاربران محاسبه می‌شود. سپس، برای کاربر فعال، اقلامی توصیه می‌شوند که بیشترین شباهت را با اقلامی دارند که کاربر قبلاً به آن‌ها امتیاز بالایی داده است (Sarwar et al., ۲۰۰۱). این روش اغلب به دلیل پایداری بیشتر شباهت اقلام نسبت به شباهت کاربران و مقیاس‌پذیری بهتر، ترجیح داده می‌شود.



شکل ۲. انواع رویکردهای پالایش مشارکتی

[1] Collaborative Filtering(CF)
[2] Memory-based CF
[3] User-based CF
[4] Item-based CF

چالش‌های رویکردهای مبتنی بر حافظه:

- پراکندگی داده[1]: ماتریس کاربر-قلم معمولاً بسیار پراکنده است، به این معنی که اکثر کاربران تنها با تعداد کمی از اقلام تعامل داشته‌اند. این امر محاسبه دقیق شباهت را دشوار می‌کند.
- مقیاس‌پذیری[2]: با افزایش تعداد کاربران و اقلام، محاسبات شباهت بسیار زمان‌بر و پرهزینه می‌شود.
- مشکل شروع سرد[3]: برای کاربران جدید (new users) یا اقلام جدید (new items) که هیچ تعاملی ثبت نکرده‌اند، تولید توصیه دشوار است.
- کاربران خاکستری (Gray Sheep): کاربرانی باسلیقه منحصربه‌فرد، به‌سختی در گروه‌های مشابه قرار می‌گیرند و توصیه‌های دقیقی دریافت نمی‌کنند.

رویکرد مبتنی بر مدل[4]

در این رویکرد، یک مدل از داده‌های موجود آموزش داده می‌شود تا بتواند امتیازات کاربران به اقلام را پیش‌بینی کند. برخی از تکنیک‌های برجسته عبارت‌اند از (شکل ۳):

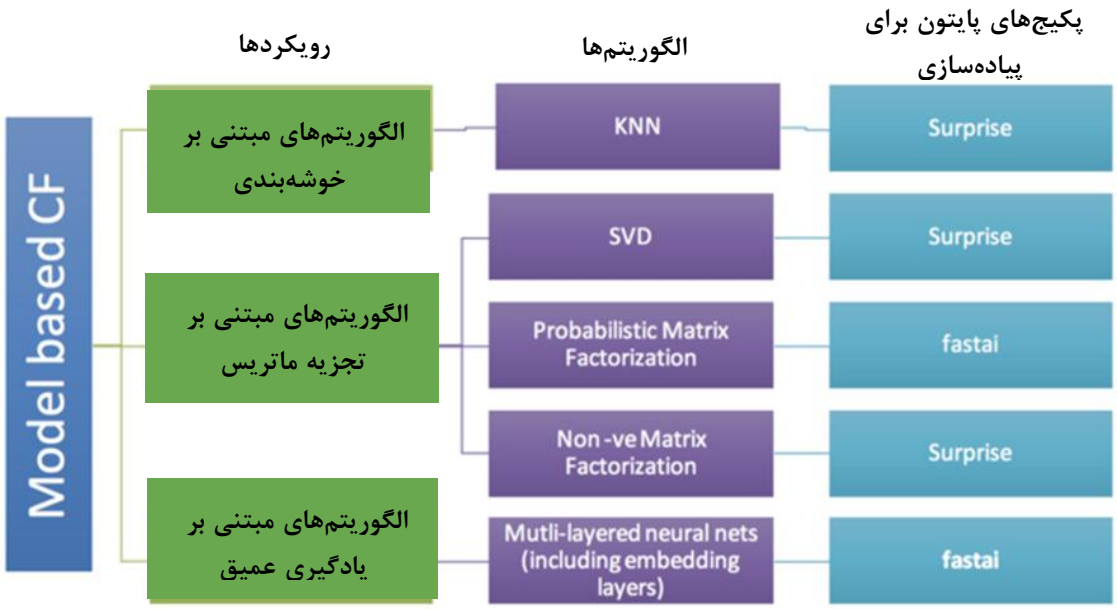
- فاکتورگیری ماتریس[5]: تجزیه ماتریس کاربر-قلم به دو ماتریس با ابعاد کوچک‌تر (مانند SVD[6]، SGD[7]، ALS[8])
- خوشه‌بندی[9]: دسته‌بندی کاربران یا اقلام بر اساس شباهت‌هایشان.
- یادگیری عمیق[10]: استفاده از مدل‌هایی مانند پالایش مشارکتی مشارکتی عصبی برای یادگیری نمایش‌های پیچیده از تعاملات کاربر-قلم.

چالش‌های رویکردهای مبتنی بر مدل:

- هزینه ساخت مدل: مدل‌های پیچیده مانند یادگیری عمیق ازنظر محاسباتی پرهزینه هستند.
- تفسیرپذیری کمتر: برخی مدل‌ها مانند جعبه سیاه عمل می‌کنند.

۳-۲- سیستم‌های توصیه‌گر مبتنی بر محتوا

سیستم‌های مبتنی بر محتوا بر اساس ویژگی‌های اقلام و پروفایل علایق کاربر عمل می‌کنند. ایده اصلی این است که اگر کاربری درگذشته از اقلامی با ویژگی‌های خاص خوشش آمده، در آینده نیز اقلامی با ویژگی‌های مشابه را خواهد پسندید.



شکل ۳. انواع رویکردها در روش مبتنی بر مدل

[1] Data Sparsity
[2] Scalability
[3] Cold Start
[4] Model-based CF
[5] Matrix Factorization(MF)
[6] Singular Value Decomposition
[7] Stochastic Gradient Descent
[8] Alternating Least Squares
[9] Clustering
[10] Deep Learning

در این رویکرد، برای هر قلم یک پروفایل (نمایه) شامل ویژگی‌های توصیفی آن (مانند سبک، کارگردان، بازیگران برای فیلم؛ یا کلمات کلیدی، نویسنده، موضوع برای مقالات) ایجاد می‌شود. به‌طور مشابه، یک پروفایل برای کاربر ساخته می‌شود که علایق او را بر اساس ویژگی‌های اقلامی که قبلاً با آن‌ها تعامل مثبت داشته (مثلاً امتیاز بالا داده) نشان می‌دهد. سپس، سیستم اقلامی را توصیه می‌کند که پروفایل آن‌ها بیشترین شباهت را با پروفایل کاربر داشته باشد. این شباهت معمولاً با استفاده از معیارهایی مانند شباهت کسینوسی بین بردارهای ویژگی محاسبه می‌شود.

مزایای سیستم‌های مبتنی بر محتوا:

- عدم نیاز به داده‌های سایر کاربران: این سیستم‌ها برای تولید توصیه برای یک کاربر خاص، تنها به داده‌های خود آن کاربر نیاز دارند.
- شفافیت و تفسیرپذیری: توصیه‌ها قابل توضیح هستند، زیرا می‌توان مشخص کرد که یک قلم بر اساس کدام ویژگی‌هایش توصیه شده است.
- مقابله با مشکل شروع سرد برای اقلام جدید: تا زمانی که ویژگی‌های یک قلم جدید مشخص باشد، می‌توان آن را به کاربران علاقه‌مند توصیه کرد، حتی اگر هیچ کاربری هنوز با آن تعامل نداشته باشد.

چالش‌های سیستم‌های مبتنی بر محتوا:

- مهندسی ویژگی[1]: استخراج و نمایش مناسب ویژگی‌های اقلام می‌تواند دشوار و زمان‌بر باشد، به‌خصوص برای اقلامی با محتوای چندرسانه‌ای یا پیچیده. کیفیت توصیه‌ها به‌شدت به کیفیت این ویژگی‌ها وابسته است.
- بیش‌تخصصی شدن[2] و عدم تنوع[3]: این سیستم‌ها تمایل دارند اقلامی را توصیه کنند که بسیار شبیه به اقلامی هستند که کاربر قبلاً دیده یا پسندیده است. این امر کشف اقلامی جدید و متنوع را محدود می‌کند.
- مشکل شروع سرد برای کاربران جدید: تا زمانی که کاربر با تعدادی قلم تعامل نکند تا پروفایل علایق او شکل بگیرد، سیستم قادر به ارائه توصیه‌های دقیق نخواهد بود.

۳-۳- سیستم‌های توصیه‌گر ترکیبی[4]

سیستم‌های ترکیبی باهدف بهره‌گیری از مزایای رویکردهای مختلف (عمدتاً پالایش مشارکتی و مبتنی بر محتوا) و پوشش دادن معایب آن‌ها ایجاد می‌شوند. ترکیب این رویکردها می‌تواند به روش‌های مختلفی انجام شود:

- وزن‌دهی[5]: نتایج حاصل از چندین سیستم توصیه‌گر با هم ترکیب‌شده و امتیاز نهایی بر اساس وزن‌های تخصیص داده شده به هر سیستم محاسبه می‌شود.
- تعویض[6]: بسته به شرایط خاص (مانند در دسترس بودن داده‌ها یا وضعیت کاربر)، سیستم بین رویکردهای مختلف جابجا می‌شود.
- ترکیب ویژگی‌ها[7]: ویژگی‌های حاصل از یک رویکرد به‌عنوان ورودی برای رویکرد دیگر استفاده می‌شود. برای مثال، ویژگی‌های محتوایی می‌توانند برای بهبود ماتریس کاربر-قلم در پالایش مشارکتی استفاده شوند.
- آبشاری[8]: ابتدا یک رویکرد برای تولید مجموعه‌ای از کاندیداها استفاده می‌شود و سپس رویکرد دیگری برای پالایش و رتبه‌بندی نهایی این کاندیداها به کار می‌رود.
- فرا سطح[9]: خروجی یک سیستم توصیه‌گر به‌عنوان ورودی برای سیستم دیگر استفاده می‌شود. مدل یاد گرفته‌شده توسط یک تکنیک، به‌عنوان ورودی برای تکنیک دیگر استفاده می‌شود.

نتفلیکس نمونه بارزی از یک سیستم توصیه‌گر ترکیبی موفق است. آن‌ها از ترکیبی از پالایش مشارکتی (مقایسه عادات تماشا و جستجوی کاربران مشابه) و پالایش‌های مبتنی بر محتوا (توصیه فیلم‌هایی با ویژگی‌های مشترک با فیلم‌هایی که کاربر امتیاز بالایی به آن‌ها داده) به همراه الگوریتم‌های پیچیده یادگیری ماشین برای ارائه توصیه‌های شخصی‌سازی شده بهره می‌برند. سیستم‌های پیشرفته مانند نتفلیکس از رویکردهای ترکیبی پیچیده‌ای بهره می‌برند که اغلب شامل چندین مدل و تکنیک است.

۳-۴- چالش‌های کلیدی و افق‌های تحقیقاتی آینده

علیرغم پیشرفت‌های قابل‌توجه، سیستم‌های توصیه‌گر همچنان با چالش‌های متعددی روبرو هستند و زمینه‌های تحقیقاتی گسترده‌ای برای بهبود آن‌ها وجود دارد:

[1] Feature Engineering
[2] Overspecialization
[3] Limited Serendipity
[4] Hybrid Recommender Systems
[5] Weighted
[6] Switching
[7] Feature Combination
[8] Cascade
[9] Meta-level

چالش‌های موجود و نیازمند توجه بیشتر

- دقت در مقابل تنوع و سرندیپیتی[1]: اکثر سیستم‌ها بر بهینه‌سازی دقت تمرکز دارند، اما توصیه‌های بسیار دقیق ممکن است منجر به یکنواختی و عدم کشف اقلام جدید و غیرمنتظره شود. ایجاد تعادل بین این معیارها یک چالش مهم است.
- تفسیرپذیری و شفافیت[2]: کاربران اغلب می‌خواهند بدانند چرا یک قلم خاص به آن‌ها توصیه شده است. افزایش تفسیرپذیری می‌تواند اعتماد کاربر را جلب کرده و به او در تصمیم‌گیری کمک کند. این موضوع به‌ویژه در سیستم‌های مبتنی بر مدل‌های پیچیده مانند یادگیری عمیق اهمیت دارد.
- حریم خصوصی[3]: سیستم‌های توصیه‌گر برای عملکرد خود به داده‌های شخصی کاربران نیاز دارند. جمع‌آوری و استفاده از این داده‌ها نگرانی‌هایی را در مورد حریم خصوصی ایجاد می‌کند. توسعه تکنیک‌هایی برای توصیه‌گری با حفظ حریم خصوصی (مانند یادگیری فدرال یا استفاده از نویز تفاضلی) ضروری است.
- انصاف و جلوگیری از سوگیری[4]: الگوریتم‌های توصیه‌گر ممکن است سوگیری‌های موجود در داده‌ها را بازتولید یا تقویت کنند، که منجر به توصیه‌های ناعادلانه برای گروه‌های خاصی از کاربران یا دیده نشدن برخی اقلام می‌شود. تحقیق در زمینه شناسایی و کاهش این سوگیری‌ها اهمیت فزاینده‌ای یافته است.
- حملات مخرب[5]: افراد یا ربات‌ها ممکن است با وارد کردن پروفایل‌های جعلی یا امتیازات ساختگی، سعی در دست‌کاری سیستم و ترویج یا تخریب اقلامی خاص داشته باشند. توسعه روش‌های مقاوم در برابر این حملات ضروری است.
- ارزیابی[6]: ارزیابی اثربخشی سیستم‌های توصیه‌گر پیچیده است. معیارهای سنتی آفلاین (مانند RMSE یا Precision@K) ممکن است به‌طور کامل تجربه کاربری واقعی یا اهداف تجاری بلندمدت را منعکس نکنند. نیاز به معیارهای جامع‌تر و تست‌های A/B آنلاین دقیق‌تر احساس می‌شود.

افق‌های تحقیقاتی آینده

- سیستم‌های توصیه‌گر آگاه از زمینه[7]: در نظر گرفتن زمینه (مانند زمان، مکان، وضعیت احساسی کاربر، همراهان کاربر) می‌تواند به‌طور قابل‌توجهی کیفیت توصیه‌ها را بهبود بخشد.
- سیستم‌های توصیه‌گر مکالمه‌ای[8]: این سیستم‌ها از طریق یک رابط مکالمه‌ای (مانند چت‌بات) با کاربر تعامل می‌کنند تا نیازها و ترجیحات او را بهتر درک کرده و توصیه‌های پویاتری ارائه دهند.
- یادگیری تقویتی برای توصیه‌گرها[9]: استفاده از یادگیری تقویتی می‌تواند به سیستم‌ها کمک کند تا توالی‌هایی از توصیه‌ها را بهینه کنند که منجر به حداکثر شدن پاداش بلندمدت (مانند تعامل یا رضایت کاربر) شود.
- توصیه‌گرهای بین دامنه‌ای[10]: انتقال دانش از یک دامنه (مثلاً فیلم) به دامنه دیگر (مثلاً کتاب) می‌تواند به حل مشکل پراکندگی داده و شروع سرد کمک کند
- توصیه‌گرهای گروهی[11]: توسعه سیستم‌هایی که بتوانند توصیه‌های مناسبی برای گروهی از کاربران با سلایق مختلف ارائه دهند (مثلاً برای یک خانواده که می‌خواهند با هم فیلم ببینند) یک حوزه تحقیقاتی فعال است.
- تکنیک‌های پیشرفته یادگیری عمیق: کاوش بیشتر در معماری‌های جدید یادگیری عمیق، مانند شبکه‌های عصبی گرافی (GNNs) برای مدل‌سازی روابط پیچیده بین کاربران و اقلام، و یا مدل‌های مبتنی بر ترنسفورمر برای درک توالی‌های رفتار کاربر، پتانسیل بالایی برای بهبود سیستم‌های توصیه‌گر دارد.

۴- پیشینه تحقیق در سیستم‌های توصیه‌گر

در راستای مرور ادبیات سیستم‌های توصیه‌گر، جمع‌آوری اطلاعات بر اساس جستجوی کلیدواژه‌های مرتبط با سیستم‌های توصیه‌گر در حوزه فیلم، موسیقی و برنامه تلویزیونی انجام پذیرفته است. جمع‌آوری مقالات از طریق موتور جستجوی Google scholar و یا مراجع ذکرشده در مقالات مروری صورت گرفته است. درمجموع ۱۵ مقاله مروری مدنظر قرارگرفته است. تعداد مقالات جمع‌آوری‌شده از طریق جستجوی وب و مراجع مقالات مروری منجر به بررسی ۱۷۴ مقاله گردیده است که این مقالات پس از بررسی اولیه در دسته‌بندی سه‌گانه رویکردهای توصیه‌گر به شرح زیر تقسیم‌بندی شده‌اند:

[1] Accuracy vs. Diversity and Serendipity
[2] Explainability and Transparency
[3] Privacy
[4] Fairness and Bias
[5] Shilling Attacks
[6] Evaluation
[7] Context-Aware Recommender Systems(CARS)
[8] Conversational Recommender Systems
[9] Reinforcement Learning for RS
[10] Cross-Domain Recommendation
[11] Group Recommender Systems

- رویکرد مبتنی بر محتوا: ۳۶ مقاله
- رویکرد پالایش مشارکتی: ۱۰۱ مقاله
- رویکرد ترکیبی: ۳۷ مقاله

با توجه به به محوریت روش‌های مبتنی بر محتوا در پژوهش حاضر، از میان مقالات فوق مطالعه دقیق‌تر مقالاتی که بر روش‌های مبتنی بر محتوا تمرکز داشتند و تعدادی از مقالات ترکیبی مرتبط مدنظر قرارگرفته است و از این تعداد مقالات ۴۷ مقاله بررسی و توضیحات تفضیلی آن در این مقاله ارائه می‌گردد. بر اساس حوزه کاربردی سیستم‌ها مقالات در سه بخش توصیه فیلم، موسیقی، برنامه تلویزیونی و چندرسانه‌ای ارائه شده‌اند.

۴-۱- سیستم‌های توصیه‌گر حوزه فیلم

سیستم‌های مبتنی بر روش‌های سنتی مبتنی بر محتوا

این سیستم‌ها از ویژگی‌های توصیفی فیلم‌ها مانند سبک، بازیگران، کارگردان، و کلیدواژه‌ها برای ساخت پروفایل کاربر و مقایسه آن با فیلم‌های موجود استفاده می‌کنند. معیارهای شباهت مانند فاصله کسینوسی برای پیشنهاد فیلم‌های مشابه به کار می‌روند.

در مقاله ارائه‌شده توسط Pradeep و همکاران [۷] یک رویکرد متداول برای توصیه مبتنی بر محتوا ارائه شده است و نوآوری آن در ارائه یک ابزار با رابط کاربری تحت وب است.

ویژگی‌های استفاده شده در این مقاله بازیگران، کلیدواژه، عوامل سازنده فیلم، و سبک‌های فیلم می‌باشد. بر اساس ۴ ویژگی ذکرشده و با استفاده از معیار فاصله کسینوسی، ۱۰ مورد از شبیه‌ترین فیلم‌ها به پروفایل کاربر یافت شده و به او نشان داده می‌شود.

در مقاله ارائه‌شده توسط Singh و همکاران [۸] با استفاده از ویژگی «محبوبیت فیلم» و « سبک » یک توصیه‌گر مبتنی بر محتوا با استفاده از K نزدیک‌ترین همسایه و فاصله کسینوسی ارائه شده است.

در مقاله ارائه‌شده توسط Almeida و همکاران [۹] بیان‌شده است که سیستم‌های توصیه‌گر، اگر فقط دقت در نظر گرفته شود، ممکن است باعث تولید توصیه‌های بایاس‌شده شود. برای رفع این مشکل، یک الگوریتم تکاملی چندهدفه ارائه شده است. سه هدف دقت، تنوع و تازگی توصیه‌ها در این الگوریتم لحاظ شده است.

در مقاله ارائه‌شده توسط Meel و همکاران [۱۰]، فرض شده است که به ازای هر فیلم، هر کاربر ممکن است با توجه به حسی که نسبت به فیلم داشته است، تعدادی tag به آن نسبت داده باشد. این tag ها می‌توانند یک عبارت یا یک کلمه باشند. فهرست تگ‌هایی که کاربران مختلف به یک فیلم داده‌اند در کنار هم قرار می‌گیرد.

سبک فیلم نیز در کنار این تگ‌ها قرار می‌گیرد و درنهایت به ازای هر فیلم یک فهرست از تگ‌ها وجود خواهد داشت. از word2vec برای به دست آوردن یک بازنمایی برای هر یک از تگ‌ها استفاده می‌شود. سپس بازنمایی هر فیلم با جمع وزن‌دار بردار word2vec هر برچسب متناسب با وزن tf-idf آن برچسب به دست می‌آید. در ادامه از SVD استفاده می‌شود.

ویژگی‌هایی که در مقاله ارائه‌شده توسط Singla و همکاران [۱۱] از فیلم استخراج‌شده است، عبارت‌اند از داستان فیلم، امتیازهای فیلم، کشورهای سازنده و سال ساخت. راه‌حل ابتدایی ارائه‌شده در این مقاله این است که با استفاده از Doc2Vec، از هر داستان فیلم یک بردار ۱۲۸ بعدی ساخته شود. مقایسه این بردارها با پروفایل کاربر، با معیار فاصله کسینوسی انجام می‌شود.

این روش دو مشکل دارد. مشکل اول این است که کیفیت فیلم‌ها را در نظر نمی‌گیرد. یعنی ممکن است فیلمی که صرفاً ازنظر داستان به پروفایل یک کاربر شبیه است ولی امتیاز بسیار کمی دارد را نیز به کاربر پیشنهاد کند. مشکل دوم این است که اگر سبک فیلم در نظر گرفته نشود، ممکن است یک فیلم که به طنز از روی فیلمی دیگر و با همان داستان ساخته شده است نیز به کاربر پیشنهاد شود. برای مثال فیلم ۳۰۰ فیلمی حادثه‌ای است و فیلم «اسپارتان‌ها را ملاقات کنید» فیلمی خنده‌دار است که با داستان فیلم ۳۰۰ ولی به طنز ساخته شده است. روش فعلی صرفاً به خاطر شبیه بودن داستان این دو فیلم، آن دو را مشابه در نظر می‌گیرد. برای رفع مشکل اول، فهرست پیشنهادهایی که صرفاً با شبیه بودن داستان به‌دست‌آمده، با امتیاز فیلم، مکان و سال ساخت پالایه می‌شود و پیشنهادهایی مناسب انتخاب می‌شود. برای رفع مشکل دوم، سبک هر فیلم به‌عنوان کلیدواژه در نظر گرفته می‌شود. هر فیلم می‌تواند تا حداکثر سه سبک داشته باشد. به دلیل این‌که بعضی از کلیدواژه سبک‌ها، مثل عاشقانه، در فیلم‌های زیادی موجود است و برخی مثل آهنگین در فیلم‌های بسیار کمی وجود دارد، مقدار TF-IDF هر یک از این کلیدواژه‌ها را به‌عنوان وزن آن در نظر می‌گیریم. دو ویژگی جدید، به بردار ۱۲۸ بعدی ابتدایی اضافه می‌شود و یک بردار واحد به ازای هر فیلم تهیه می‌شود. این بردار با استفاده از معیار فاصله کسینوسی با پروفایل‌های کاربران مقایسه خواهد شد.

سیستم‌های جدید مبتنی بر زمینه

این سیستم‌ها علاوه بر ویژگی‌های محتوایی و تعاملات کاربر،

اطلاعات زمینه‌ای مانند زمان، مکان، و وضعیت احساسی کاربر را نیز در نظر می‌گیرند. برای مثال، از مدل‌های مارکوف برای پیش‌بینی ترجیحات کاربر در بازه‌های زمانی مختلف استفاده می‌کنند.

مقاله ارائه‌شده توسط Yin و همکاران [۱۳] بر روی مدل کردن علایق کاربر با استفاده از شبکه‌های اجتماعی تأکید دارد. مدلی که در این مقاله ارائه شده است، کاربر را بر اساس دو ویژگی سلیقه شخصی کاربر و اطلاعات زمینه‌ای زمانی یا تأثیری که در زمان‌های مختلف کاربر از بقیه جامعه می‌گیرد، مدل می‌کند. این مدل، مدل آمیخته آگاه از زمینه زمانی نام دارد. علاوه بر این مدل پایه، از آن‌جایی که سلیقه شخصی کاربر نیز در طول زمان متغیر است، برای مدل کردن این رفتار کاربر، مدل TCAM پویا ارائه شده است. در این مقاله، بر اساس دو مدل ارائه شده، برای کاربر پروفایل ساخته شده و یک سیستم توصیه‌گر آگاه از زمینه زمانی ارائه شده است.

در مقاله ارائه‌شده توسط Elashmawi و همکاران [۸۱] جهت توصیه فیلم، ۴ مرحله طی می‌شود: در مرحله نخست، به استخراج فریم‌ها و قطعات صدا و افزایش کیفیت و کاهش نویز در آن‌ها می‌پردازند. در مرحله دوم، با استفاده از تکنیک‌های مختلف نظیر تشخیص اشیا، تشخیص چهره، دسته‌بندی صحنه و تشخیص صدا به استخراج ویژگی‌ها از فریم‌ها و قطعات صدا می‌پردازند. در مرحله سوم، شباهت ویدئوها بر مبنای ویژگی‌های استخراج‌شده با بکارگیری معیارهایی نظیر شباهت کسینوسی، فاصله اقلیدسی و جنبش زمانی پویا را محاسبه می‌کند. در مرحله آخر، سیستم بر اساس میزان مشابهت ویدئوی موردعلاقه کاربر با ویدئوهای موجود در پایگاه داده خود را بر اساس بر ویژگی‌های استخراج‌شده از اشیا موجود، چهره‌ها، صحنه‌ها، صداها و کلمات موجود در آن، را به کاربر پیشنهاد می‌کند.

در مقاله ارائه‌شده توسط Markapudi و همکاران [۸۲] از یک شبکه عصبی چند لایه رو به جلو و روش نمونه‌گیری مونت کارلو جهت پیشنهاد ویدئو به کاربر استفاده می‌شود. در این روش، ابتدا با استفاده از پالایه کاهش نویز گاوسی تطبیقی حرکتی کیفیت و وضوح فریم‌های ویدئو بهبود داده می‌شود. سپس ویدئو به تعدادی قطعه، که هر قطعه از تعدادی فریم تشکیل شده است، تقسیم شده و ویژگی‌های زمانی و مکانی هر قطعه استخراج می‌شود. ازجمله ویژگی‌های زمانی، میانگین حرکات اشیا و ازجمله ویژگی‌های مکانی، رنگ موجود در هر فریم است. در مرحله بعد، ویژگی‌های استخراج‌شده از هر ویدئو در قالب بردار بازنمایی می‌شوند که محتوای هر ویدئو را تفسیر می‌کند. سپس مدل شبکه عصبی چندلایه روبه جلو بر روی بردار ویژگی‌های حاصل از ویدئو آموزش داده می‌شود تا بدین‌وسیله مجموعه‌ای از خروجی‌ها، که مشخص‌کننده ویدئوی پیشنهادی است، تولید شود. سپس از روش نمونه‌گیری مونت کارلو برای تخمین عدم قطعیت پیش‌بینی انجام‌شده استفاده می‌شود تا بدین وسیله مرتبط‌ترین ویدئوها به کاربر پیشنهاد گردد.

ازجمله مزایای این روش عدم نیاز به داده‌های سایر کاربران (حل مشکل شروع سرد برای اقلام جدید) و تفسیرپذیری بالا بر اساس ویژگی‌های مشخصی ارائه‌شده می‌باشد. ازجمله معایب این روش بیش‌تخصصی شدن (Overspecialization) و عدم تنوع در پیشنهادها و وابستگی شدید به کیفیت فراداده‌های فیلم‌ها می‌باشد.

سیستم‌های ترکیبی (محتوا + پالایش مشارکتی)

این سیستم‌ها از ترکیب روش‌های مبتنی بر محتوا و پالایش مشارکتی استفاده می‌کنند. برای مثال، برخی از آن‌ها ابتدا از پالایش مشارکتی برای پیش‌بینی امتیاز کاربران استفاده کرده و سپس با استفاده از ویژگی‌های محتوایی، پروفایل کاربر را بهبود می‌بخشند.

در مقاله ارائه‌شده توسط Liang و همکاران [۵] بر روی دو مشکل اساسی پالایش مشارکتی، تنک بودن ماتریس کاربر-قلم و شروع سرد، تأکید شده است. ابتدا برای ایجاد پروفایل کاربر از کلیدواژه‌های مرتبط با هر فیلم به‌عنوان ویژگی استفاده شده است. سپس در بخش پالایش مشارکتی از الگوریتم‌های تخمین توزیع جهت یادگیری ترجیحات کاربران بر اساس پروفایل‌هایشان استفاده شده است. مدل توصیه‌گر این مقاله یک مدل دوگانه است که بخش مبتنی بر محتوا، با استفاده از پروفایل ساخته‌شده می‌تواند موسیقی‌های جدید به کاربر پیشنهاد بدهد و بخش پالایش مشارکتی با استفاده از این پروفایل، به جای استفاده مستقیم از ماتریس کاربر-قلم، می‌تواند کاربران مشابه با یکدیگر را بیابد.

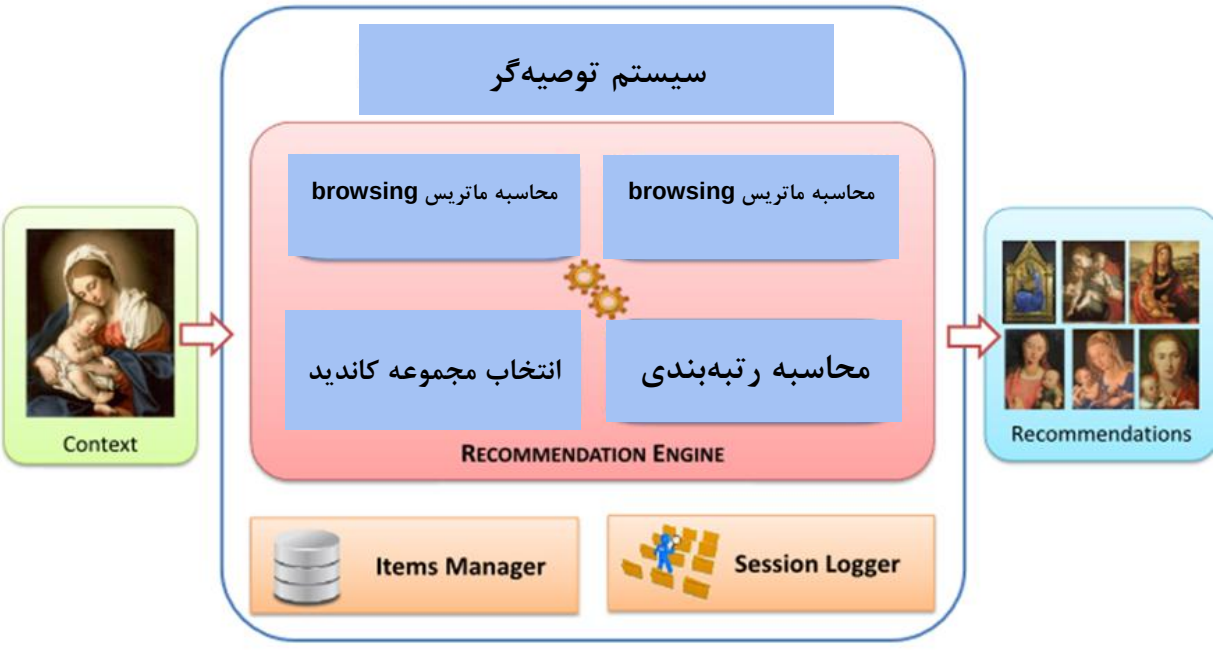
در سیستم‌های توصیه‌گر مبتنی بر محتوای سنتی، پروفایل کاربر بر اساس سابقه فیلم‌هایی که کاربر تماشا کرده است ساخته‌شده و فیلم‌هایی که کاربر ندیده است و مشابه با پروفایل او هستند به او پیشنهاد می‌شود. مقاله ارائه‌شده توسط Chen و همکاران [۶] به این مشکل اشاره می‌کند که تعداد فیلم‌هایی که کاربر مشاهده کرده است، نسبت به تمام فیلم‌های موجود، بسیار کم است. به همین دلیل در این مقاله ابتدا با استفاده از روش پالایش مشارکتی، امتیازی که ممکن است کاربر به فیلم‌هایی که ندیده است بدهد، پیش‌بینی می‌شود. با استفاده از این امتیازها، دو فهرست برای هر کاربر تهیه می‌شود. فهرست فیلم‌هایی که پیش‌بینی می‌شود کاربر بیشترین امتیاز را به آن‌ها بدهد و فهرست فیلم‌هایی که پیش‌بینی می‌شود کاربر کمترین امتیاز را به آن‌ها بدهد. با استفاده از فهرست اول،

پروفایل مثبت کاربر و با استفاده از فهرست دوم پروفایل منفی کاربر ساخته می‌شود. درنهایت با استفاده از روش مبتنی بر محتوا، فیلم‌هایی به کاربر پیشنهاد می‌شود که بیشترین شباهت به پروفایل مثبت کاربر و کمترین شباهت به پروفایل منفی کاربر را داشته باشد.

در مقاله ارائه‌شده توسط Sottocornola و همکاران [۱۲]، سیستم توصیه‌گری ارائه‌شده است که از سه بخش تشکیل شده است. بخش اول این سیستم، یک توصیه‌گر پالایش مشارکتی مبتنی بر الگوریتم K نزدیک‌ترین همسایه و مبتنی بر کاربر است که بر اساس این‌که کاربران پیش از این چه فیلم‌هایی را دوست داشته و چه فیلم‌هایی را دوست نداشته‌اند، احتمال این‌که کاربر فیلم جدیدی را دوست داشته باشد را بیان می‌کند. بخش دوم، یک سیستم توصیه‌گر مبتنی بر محتوا است که داستان هر فیلم را با استفاده از doc2vec به فضای برداری متراکم می‌برد و سپس با استفاده از جنگل تصادفی، میزان احتمال علاقه کاربر به فیلم مورد نظر را به دست می‌آورد. در بخش آخر، یک شبکه عصبی پیش‌خور وجود دارد که به‌عنوان ورودی، میزان احتمال علاقه کاربر تولیدشده توسط بخش پالایش مشارکتی، دقت مدل پالایش مشارکتی، میزان علاقه کاربر تولیدشده توسط بخش مبتنی بر محتوا و دقت مدل مبتنی بر محتوا را دریافت کرده و مقدار احتمال نهایی علاقه کاربر به فیلم جدید ارائه‌شده را به دست می‌آورد.

در مقاله ارائه‌شده توسط Albanese و همکاران [۱۴] یک رویکرد جدید برای حل مسئله پیشنهاد قلم در سیستم‌های جستجوی چندرسانه‌ای ارائه‌شده است. در این رویکرد باهدف محاسبه یک رتبه‌بندی سراسری از همه اقلام موجود در یک مجموعه داده، از روشی مانند نحوه کار موتور جستجوی گوگل استفاده می‌کند و سپس با استفاده از امتیازات سراسری محاسبه شده کاندیداها را رتبه‌بندی می‌کند. برای مشاه معماری می‌توانید به شکل ۴ مراجعه کنید.

در نظریه انتخاب اجتماعی، مجموعه‌ای از رأی‌دهندگان برای رتبه‌بندی مجموعه‌ای از گزینه‌ها فراخوانی می‌شوند و رتبه‌بندی‌های فردی به‌صورت یک رتبه‌بندی سراسری یکپارچه‌سازی می‌شوند. در فرمول‌بندی ما، مجموعه رأی‌دهندگان و مجموعه گزینه‌ها هر دو با مجموعه اقلام موجود در داده‌ها منطبق هستند. ما ابتدا آنچه را که به معنای انتخاب در دامنه جستجو است را تعریف می‌کنیم و سپس مکانیزمی را برای جمع آوری انتخاب‌های افراد در یک رتبه‌بندی سراسری تعریف می‌کنیم. نتیجه این کار یک قالب کلی برای محاسبه پیشنهادات سفارشی با ترکیب ویژگی‌های ذاتی اقلامی چندرسانه‌ای، رفتار گذشته کاربران و رفتار کلی کل جامعه کاربران است. توصیه‌ها با استفاده از الگوریتم رتبه‌بندی اهمیت که شبیه استراتژی معروف PageRank است رتبه‌بندی می‌شوند.



شکل ۴. معماری سیستم پیشنهادی آلبانی و همکاران [۱۴]

رویکرد مقاله یک استراتژی ترکیبی دارد که دو استراتژی محتوایی و مشارکتی را همزمان استفاده می‌کند. این برنامه از سوابق سیستم استفاده می‌کند تا به‌طور ضمنی اطلاعات مربوط به تک‌تک کاربران و جامعه کاربران را به‌طورکلی به‌دست آورد و به‌عنوان رتبه‌بندی یکنواخت از آن استفاده کند. این رویکرد مشابه با روش‌های پالایش تصادفی در پالایش مشارکتی، اطلاعات فعالیت‌های گذشته را بصورت یک گراف جهت‌دار تولید می‌کند و بر اساس این گراف مناسب‌ترین اقلام را برای پیشنهاد انتخاب می‌کند.

شباهت این الگوریتم با روش‌های استخراج مبتنی بر قاعده بدین صورت است که برای محاسبه اهمیت هر قلم از روابط انتقالی بین اقلام استفاده می‌کند. مشابه با روش‌های بازیابی اطلاعات و پالایش، به‌خصوصیات اقلامی که کاربر در حال حاضر مشاهده می‌کند، اهمیت زیادی می‌دهد تا بتواند به‌طور موثر سودمندی موارد دیگر را محاسبه کند. سرانجام، رویکرد پیشنهادی از دانش پیشینی در مورد فراداده و روابط آن‌ها استفاده می‌کند و هر دو نوع اطلاعات سطح پایین (ویژگی‌ها) و سطح بالا (فراداده) را در نظر می‌گیرد، به‌طوری‌که هر دو سطح اطلاعات برای تعیین سودمندی یک قلم در فرآیند توصیه لحاظ شوند.

معماری کلی سیستم در شکل ۴ نشان داده شده است. بر اساس این معماری اجزای اصلی زیر در این مدل قابل تفکیک هستند:

- مدیریت اقلام[1] - یک مخزن داده که اقلامی را که باید با توضیحات مرتبط پیشنهاد شوند ذخیره می‌کند.
- ابزار لاگ زیربخش‌ها[2] - بخشی مختص ضبط و ذخیره اطلاعات مربوط به هر زیربخش اقلام که اطلاعات مربوط به اقلام بازدید شده را نگهداری می‌کند.
- موتور توصیه[3] - مولفه اصلی سیستم است که بر اساس اطلاعات مربوط به شباهت اقلام و تاریخچه مرورگر کاربران جامعه، به‌صورت پویا مجموعه‌ای از اقلام را پیشنهاد می‌کند. اقلام خروجی بر اساس کاهش میزان سودمندی مرتب شده‌اند. این مولفه به‌طور خاص شامل زیرمجموعه‌های زیر است:
  - بخش محاسبه ماتریس مرورگر که داده‌های جمع‌آوری‌شده مرورگر را به دو نوع ماتریس تبدیل می‌کند: یک ماتریس سراسری که رفتار کلی کاربران را در نظر می‌گیرد و یک ماتریس محلی که رفتار یک کاربر را در نظر می‌گیرد.
  - بخش محاسبه ماتریس شباهت که شباهت بین هر جفت از اقلام را محاسبه می‌کند و نتایج را در یک ماتریس ذخیره می‌کند.
  - بخش انتخاب مجموعه کاندیداها که زیر مجموعه اقلامی را که با نیازهای کاربر منطبق‌ترند را محاسبه می‌کند.
  - بخش رتبه‌بندی که کاندیداهای منتخب را برای توصیه رتبه‌بندی می‌کند.

ازجمله مزایای این روش کاهش مشکل شروع سرد برای کاربران و اقلام جدید و بهبود دقت توصیه‌ها با ترکیب سیگنال‌های مختلف می‌باشد. ازجمله معایب این روش پیچیدگی محاسباتی بالاتر نسبت به روش‌های خالص و نیاز به داده‌های تاریخی تعامل کاربران می‌باشد.

در پژوهش انجام شده توسط Rajasekar و همکاران [۱۵] تلاش کرده است تا نقاط ضعف هر دو روش را پوشش داده و دقت پیشنهادات را ارتقا بخشد. از ویژگی‌های برجسته این سیستم، بهره‌گیری از منابع داده‌ای متنوع نظیر متادیتای فیلم‌ها و تاریخچه تماشای کاربران، در کنار ایجاد یک محیط تعاملی برای دریافت بازخوردهای مستمر است که منجر به پالایش مداوم الگوریتم و بهبود تجربه جمعی کاربران در کشف آثار سینمایی می‌گردد.

۴-۲- سیستم‌های توصیه‌گر حوزه موسیقی

سیستم‌های مبتنی بر روش‌های سنتی مبتنی بر محتوا

این سیستم‌ها از ویژگی‌های استخراج‌شده از فایل‌های صوتی مانند MFCC، گام، ریتم، و انرژی برای توصیه موسیقی استفاده می‌کنند. برخی از روش‌ها از یادگیری عمیق برای تحلیل خودکار این ویژگی‌ها بهره می‌برند.

در مقاله ارایه شده توسط Shashaani [۱۶]، با هدف افزایش اعتماد و رضایت کاربران، بر مقوله شفافیت و تبیین‌پذیری در سیستم‌های توصیه‌گر موسیقی تمرکز دارد. با بررسی شکاف‌های موجود در ارائه توضیحات قانع‌کننده به کاربر، این تحقیق تلاش می‌کند راهکارهایی برای ایجاد توصیه‌های تعاملی‌تر و قابل‌اعتمادتر ارائه دهد تا از این طریق، تجربه کاربری در مواجهه با محتوای جدید بهبود یابد.

در مقاله ارائه‌شده توسط Ord و همکاران [۱۷] بیان‌شده است که اکثر سیستم‌های توصیه‌گر مرتبط با موسیقی، از روش پالایش مشارکتی استفاده می‌کنند ولی یکی از اصلی‌ترین مشکلات این روش، شروع سرد است. این امر باعث می‌شود که وقتی داده‌ای در

[1] Items Manager
[2] Session Logger
[3] Recommendation Engine

ابتدای کار وجود ندارد، پیشنهاد دادن موسیقی سخت است و به همین دلیل موسیقی‌های تازه یا موسیقی‌هایی که خیلی معروف نیستند، معمولا در این سیستم‌ها پیشنهاد نمی‌شوند. در این مقاله یک latent factor model برای توصیه معرفی شده است. هنگامی که اطلاعاتی از استفاده کاربران و موسیقی‌ها وجود نداشته باشد، سیگنال‌های صوتی موجود در محتوای موسیقی استخراج می‌شود. برای پردازش سیگنال‌های صوتی در این مقاله روش بازنمایی ساده و سنتی bag-of-words و روش پیشرفته‌تر شبکه‌های عصبی کانولوشنی استفاده شده است و مورد مقایسه قرارگرفته است. دیتاست مورداستفاده، Million Song Dataset بوده است.

در مقاله ارائه‌شده توسط Liu و همکاران [۱۹] روشی برای مقایسه فاصله (عکس شباهت) محتوای قطعه‌های موسیقی مختلف، بر اساس سلیقه کاربر پیشنهاد می‌دهد. این روش بر اساس الگوریتم تصادفی، الگوریتم تکاملی و برنامه‌نویسی تکاملی است. در این مقاله دو دسته ویژگی از هر موسیقی استخراج می‌شود. فرآیند استخراج ویژگی‌های دسته اول به این صورت است که ابتدا برای بخش ملودی اصلی هر موسیقی، یک فایل MIDI به‌صورت دستی تهیه می‌شود. با استفاده از این فایل ویژگی‌هایی مرتبط با صوت موسیقی استخراج می‌شود. میانگین گام (فرکانس پایه)، آنتروپی گام، چگالی گام، مدت زمان متوسط، آنتروپی مدت زمان، آنتروپی فاصله گام نمونه‌هایی از این ویژگی‌ها هستند. همچنین از شکل موج سیگنال موسیقی نیز ویژگی MFCC و فرمنت استخراج می‌شوند.

در مقاله ارائه‌شده توسط Bogdanov و همکاران [۲۰] از صوت هر یک موسیقی‌های یک مجموعه موسیقی که کاربر به آن‌ها علاقه داشته است، ویژگی‌های سطح بالایی استخراج می‌شود. سپس با استفاده از این ویژگی‌ها برای کاربر یک نمایه[1] تعریف می‌شود و کار توصیه با استفاده از آن انجام می‌شود. با توجه به اطلاعات استخراج‌شده از محتوای صوتی شباهت بین پروفایل کاربر و موسیقی‌ها بر اساس مدل‌های معنایی انجام می‌گیرد.

ازجمله مزایای این روش‌ها عدم نیاز به داده‌های تعاملی کاربران (حل مشکل شروع سرد) و قابلیت توصیه موسیقی‌های جدید بر اساس محتوای صوتی می‌باشد. ازجمله معایب این روش‌ها وابستگی به کیفیت استخراج ویژگی‌های صوتی و عدم توجه به سلیقه‌های فردی کاربران در برخی روش‌هاست.

در مقاله ارائه‌شده توسط Wang و همکاران [۲۱] یک مدل احتمالاتی ارائه داده شده است که با ترکیب اطلاعات جمع‌آوری‌شده توسط تلفن همراه کاربر و اطلاعات مرتبط با محتوای موسیقی، به کاربر موسیقی پیشنهاد می‌دهد. در سیستم ارائه‌شده، کاربر می‌تواند به‌صورت دستی فعالیتی که در حال انجام است را انتخاب کند یا اجازه دهد که موبایلش با اطلاعاتی که جمع‌آوری کرده است، مثل اطلاعات مرتبط با شتاب یا نویز محیطی، فعالیت او را به دست آورد. این فعالیت‌ها شامل کار کردن، درس خواندن، دویدن، خواب، پیاده‌روی و خرید کردن است. برای طبقه‌بندی فعالیت فعلی کاربر با استفاده از اطلاعات سنسورها، ۶ طبقه‌بندی درخت تصمیم، رگرسیون لاجیستیک، بیز ساده، ماشین بردار پشتیبان و k نزدیک‌ترین همسایگی با هم مقایسه شده‌اند و درنهایت طبقه‌بندی بیز ساده انتخاب شده است. پس از مشخص شدن فعالیت کاربر، باید برای این فعالیت به او موسیقی پیشنهاد شود. با انجام تحقیقی، نشان داده شده است که کاربران زیادی نسبت به این‌که یک موسیقی برای چه فعالیتی مناسب است، اتفاق نظر دارند. درنتیجه از مجموعه دادگان Grooveshark - که به ازای هر فعالیت حدود ۳۰۰۰ موسیقی دارد - استفاده شده است. برای شخصی‌سازی شدن پیشنهادهایی که به هر کاربر داده می‌شود، از بازخوردهای مثبت کاربر (گوشی دادن کامل به موسیقی یا like کردن آن) و بازخوردهای منفی (رد کردن موسیقی پس از گذشت چند ثانیه از شروع موسیقی) استفاده شده است.

سیستم‌های جدید مبتنی بر محتوا

این سیستم‌ها از مدل‌های پویا مانند یادگیری تقویتی یا مارکوف برای تطبیق توصیه‌ها با تغییر سلیقه کاربران در طول زمان استفاده می‌کنند. به‌علاوه، در این سیستم‌ها از شبکه‌های عصبی برای پردازش مستقیم سیگنال‌های صوتی یا ترکیب آن با داده‌های متنی (مانند متن آهنگ) استفاده می‌کنند.

در مقاله ارائه شده توسط Chiliguano و همکاران [۲۳] با استفاده از شبکه‌های عصبی کانولوشنی، یک بازنمایی سطح بالا از یک قسمت از فایل صوتی موسیقی ساخته می‌شود. این بازنمایی نمایانگر این است که موسیقی مورد نظر با چه احتمالی مربوط به چه سبکی است. برای به دست آوردن رفتار کاربر در گوشی دادن به موسیقی، از الگوریتم‌های تخمین توزیع استفاده می‌شود. با استفاده اطلاعات کاربر و اطلاعات موسیقی، توصیه موسیقی انجام می‌شود.

در مقاله ارائه شده توسط Chou و همکاران [۲۴]، موضوع پیشنهاد موسیقی بعدی به کاربر، بر اساس آخرین موسیقی که کاربر گوش کرده است، مطرح است. در این پژوهش بر روی حل مسئله شروع

[1] Profile

سرد تأکید شده است. برای حل این مسئله، الگوریتمی مبتنی بر تجزیه ماتریس ارائه شده است که از اطلاعات سیگنال صوتی موسیقی استفاده کرده است. با در دست داشتن دنباله موسیقی‌هایی که کاربر گوش کرده است، می‌توان یک ماتریس انتقال ایجاد کرد که در آن احتمال انتقال از هر موسیقی به موسیقی دیگر برای یک کاربر خاص به دست آورده می‌شود. محاسبه این احتمال با شمارش جفت موسیقی‌هایی که کاربر آن‌ها را در کنار هم پخش کرده است انجام می‌شود.

در این پژوهش از ویژگی MFCC به‌عنوان ویژگی سیگنال صوتی موسیقی استفاده شده است.

مقاله ارائه شده توسط Turrin و همکاران [۲۵] بر روی مساله توصیه یک فهرست پخش از موسیقی‌ها تمرکز دارد. نکته مهم در فهرست پخش این است که ترتیب موسیقی‌های موجود در فهرست پخش اهمیتی ندارد. معمولا در مسائل توصیه موسیقی یا فهرست پخش موسیقی، به دست آوردن اطلاعات زمینه سخت است. این مقاله به جای به دست آوردن این چنین اطلاعاتی برای تعیین سلیقه کاربر، از موسیقی‌ای که کاربر در حال حاضر در حال گوش دادن است استفاده می‌کند. در صورتی که کاربر در حال گوش دادن به سبک خاصی از موسیقی باشد، می‌توان سبک این موسیقی را به‌عنوان علاقه کوتاه‌مدت کاربر در نظر گرفت. دو الگوریتم پایه ارائه شده است. در الگوریتم اول، بهترین موسیقی‌های هنرمندی که کاربر در حال گوش دادن به یکی از موسیقی‌های او است به‌صورت فهرست به او معرفی می‌شود. در الگوریتم دوم، بهترین موسیقی هنرمند فعلی و هنرمندانی که مشابه هنرمند فعلی هستند، به‌صورت فهرست پخش ارائه می‌شود. برای یافتن هنرمندان مشابه، از معیار شباهت کسینوسی بر روی ماتریس هنرمند-نشست استفاده می‌شود. هر نشست شامل فهرست تمام موسیقی‌هایی است که یک کاربر در حین استفاده از نرم‌افزار به آن‌ها گوش داده است.

الگوریتم ارائه شده در این مقاله، عمل اکتشاف، جستجو جهت پیدا کردن موسیقی جدید توسط کاربر، را در نظر نمی‌گیرد. در این الگوریتم در صورتی که یک کاربر نشست‌هایی از گوش دادن به موسیقی داشته باشد که تا حدی مشابه با یک‌دیگر باشند، به آن یک فهرست پخش ضمنی گفته می‌شود. شباهت دو نشست با هم توسط معیار شباهت جاکارد کوچک شده محاسبه می‌شود. پس از به دست آوردن فهرست‌های پخش، ماتریسی دو بعدی ساخته می‌شود که یک بعد آن موسیقی‌ها و یک بعد آن فهرست پخش‌های ضمنی است. هنگامی که نشست جدیدی از یک کاربر جدید وارد سیستم می‌شود، برداری با اندازه تعداد کل موسیقی‌ها برای او ساخته می‌شود که در صورتی که کاربر به یک موسیقی در این نشست گوش داده باشد، خانه مربوط به آن موسیقی در این بردار برابر با ۱ خواهد بود. با ضرب این بردار در ماتریس تهیه شده در قسمت قبل، امتیاز تشابه هر یک از فهرست پخش‌های ضمنی به نشست فعلی به دست می‌آید. با ترکیب فهرست‌پخش‌های ضمنی‌ای که بیشترین امتیاز را دریافت کرده‌اند، فهرست پخش مرتبط با نشست فعلی به دست می‌آید.

در مقاله ارائه شده توسط Vasile و همکاران [۲۶] یک روش جدید برای محاسبه شباهت بین اقلام با در نظر گرفتن فراداده (metadata) موجود ارائه شده‌است. برای محاسبه تعبیه‌های سطح پایین اقلام، از تعامل‌های گذشته کاربر با اقلام و ویژگی‌های اقلام استفاده می‌شود. تعبیه اقلام ایده‌ای است که از مدل skip-gram در تعبیه کلمات الهام گرفته شده است و سعی دارد بر اساس دنباله انتخاب اقلام توسط هر کاربر یک تعبیه برای هر قلم ارائه دهد. ایده این مقاله مانند ایده product2vec هست که آن هم از تعبیه کلمات الهام گرفته است. همچنین اطلاعات فراداده قلم به‌عنوان اطلاعات جانبی به بهبود مدل کمک می‌کند.

یکی از مسائل مهم در سیستم‌های توصیه‌گر، در نظر گرفتن همزمان رعایت سلیقه کاربر و معرفی اقلام جدید به کاربر است. به معرفی اقلامی که دقیقا مشابه سلیقه کاربر است استخراج[1] و به معرفی قلم جدید به کاربر اکتشاف[2] می‌گوییم. در مقاله ارائه شده توسط Wang و همکاران [۲۷]، مساله انتخاب میزان مناسب اکتشاف و استخراج به‌صورت مساله راهزن چند-دسته[3] - که یک مساله کلاسیک در یادگیری تقویتی است - مدل شده است. برای یادگیری سلیقه کاربر، از یک مدل بیزین استفاده شده است که محتوای موسیقی و جدید بودن اقلام پیشنهادی را در نظر می‌گیرد.

در مساله راهزن چند دسته، مجموعه محدود و با اندازه ثابتی از منابع داریم و هدف ما این است که این منابع را به گونه‌ای تخصیص دهیم که امتیاز دریافتی بیشینه شود. در این مساله، ویژگی‌های انتخاب منابع، در زمان تخصیص، تا حدی شناخته شده است. مساله سیستم توصیه‌گر به این صورت مدل می‌شود که موسیقی‌ها، همان دست[4] است و و امتیازی که کاربر کاربر می‌دهد، payoff است. ویژگی‌هایی که از صوت موسیقی در این پژوهش مورداستفاده قرارگرفته‌است عبارت‌اند از: نرخ عبور از صفر (Zero Crossing

[1] Exploitation
[2] Exploration
[3] Multi-armed bandit
[4] Arm

(Rate، Centroid، Rolloff، Flux، MFCC، کروما (Chroma) و SFC و SFM.

در مقاله ارائه شده توسط Rao و همکاران [۲۸]، یک مدل مبتنی بر شبکه عصبی کانولوشنی طراحی‌شده است که در ورودی، اسپکتروگرام موسیقی را می‌گیرد. در ادامه چندین لایه کانولوشنی وجود دارد و در لایه یکی مانده به آخر، یک لایه تماما متصل با ۱۰۰ نورون و در لایه آخر، یک لایه softmax با ۲۰ هزار نورون. هدف این شبکه این است که با گرفتن اسپکتروگرام یک موسیقی، ۵۰۰۰ نفری که بیشتر از همه کاربران دیگر به این موسیقی گوش داده‌اند را پیش‌بینی کند. پس از آموزش این شبکه، وقتی یک موسیقی جدید به‌عنوان ورودی به شبکه داده شود، خروجی لایه یکی مانده به آخر آن به‌عنوان بردار ویژگی آن موسیقی استفاده می‌شود.

در مقاله ارائه شده توسط Hosseinzadeh و همکاران [۲۹] این چالش مطرح شده است که در سیستم‌های توصیه‌گر معمولا ترجیح عمومی کاربر را به‌عنوان پروفایل او در نظر می‌گیرند و اهمیت زیادی برای وضعیت فعلی کاربر قائل نیستند. در این پژوهش یک مدل مخفی مارکوف سلسله‌مراتبی برای کشف تغییر سلیقه کاربر ارائه شده است. در صورتی که بازخورد کاربر بر روی دنباله‌ای از اقلام موجود باشد، می‌توان کاربر را یک پروسه مخفی مارکوف سلسله‌مراتبی در نظر گرفت و زمینه یا شرایط فعلی کاربر را، متغیر مخفی این مدل در نظر گرفت.

به ازای هر کاربر، این مدل بیشینه درست‌نمایی دنباله انتقال‌ها بین حالت‌های زمینه‌ای مختلف را به دست می‌آورد و توزیع احتمال زمینه بعدی بر اساس عمل بعدی کاربر را پیش‌بینی می‌کند. حالت کاربر در این مقاله می‌تواند استراحت، رانندگی، ورزش و غیره باشد.

در مقاله ارائه شده توسط Jannach و همکاران [۳۰] این چالش مطرح شده است که موسیقی‌هایی که به کاربر پیشنهاد می‌شود، نباید صرفاً باسلیقه عمومی کاربر تطابق داشته باشد، بلکه باید با موسیقی‌هایی که کاربر به تازگی گوش کرده نیز متناسب باشد. روش ارائه شده در این مقاله برای حل این مشکل، به دو مرحله تقسیم می‌شود. در مرحله اول، هدف این است که یک فهرست از موسیقی‌های پیشنهادی برای کاربر ساخته شود. برای این کار، تکرار حضور هر دو موسیقی در کنار هم و در یک فهرست بخش عمومی یکسان بررسی می‌شود. در این مرحله از الگوریتم K نزدیک‌ترین همسایه، برای پیدا کردن فهرست موسیقی‌ای که موسیقی فعلی بیشترین شباهت به آن را دارد، استفاده خواهد شد. روش این بخش، مشابه با پالایش مشارکتی است. در بخش دوم، هدف این است که با به دست آوردن تم هر فهرست پخش، موسیقی‌هایی به کاربر پیشنهاد شود که به فهرست پخش فعلی او (موسیقی‌هایی که به تازگی پخش کرده است) نزدیک باشد. برای این کار با روش‌های مختلف می‌توان یک امتیاز میزان شباهت موسیقی به فهرست پخش فعلی به دست آورد. یک روش استفاده از تگ‌هایی است که در شبکه اجتماعی Last.fm به موسیقی‌ها انتساب داده شده است. برای این‌که تاثیر تگ‌هایی با ارزش معنایی کم، مثلاً برچسب «دوست داشتم» کم شود، از ضریب TF-IDF به‌عنوان وزن هر برچسب استفاده شده است. روش دیگر، استفاده از ویژگی‌های از صوت مانند گام و بلندی استفاده خواهد شد. در ادامه، از تمام فهرست پخش‌هایی که کاربر در طول زمان ساخته است، میانگین گرفته می‌شود تا یک پروفایل برای کاربر به دست آید. از جمع وزن‌دار امتیازهای به‌دست‌آمده، امتیاز نهایی میزان علاقه کاربر به موسیقی جدید به دست می‌آید. وزن‌ها، هایپر پارامترهای مساله هستند.

در مقاله ارائه شده توسط Vall و همکاران [۳۱]، مساله گسترش یک فهرست پخش موسیقی، با موسیقی‌های جدید، مطرح شده است. برای حل این مساله می‌توان از رویکرد پالایش مشارکتی استفاده کرد ولی استفاده از این رویکرد دو چالش دارد. چالش اول این است که فقط به فهرست پخش‌هایی می‌توان موسیقی اضافه کرد که در زمان آموزش، این فهرست‌ها در نظر گرفته شده باشند. چالش دیگر این است که این رویکرد به‌طورکلی نسبت به موسیقی‌هایی که در فهرست پخش‌های زیادی حضور نداشته‌اند، ضعیف عمل می‌کند. برای حل این دو مشکل، در این پژوهش رویکردی مبتنی بر محتوا ارائه شده است که احتمال تعلق هر موسیقی به هر فهرست پخش را محاسبه می‌کند. ویژگی‌هایی که در این پژوهش استفاده می‌شود، شامل ویژگی‌های مرتبط با فایل صوتی مثل بردار-i (که در ابتدا در مسئله تشخیص گوینده استفاده شده است)، اطلاعات word2vec استخراج‌شده از تگ‌های مربوط به هر موسیقی در شبکه‌های اجتماعی و اطلاعات پالایشی مرتبط با تعداد دفعات پخش موسیقی توسط هر کاربر است.

یک شبکه عصبی تعریف شده است که ورودی آن، تمام موسیقی‌های یک فهرست پخش و یک موسیقی جدید است. هدف این شبکه عصبی این است که مشخص کند آیا موسیقی جدید مرتبط با این فهرست پخش هست یا نه. به ازای هر یک از ورودی‌های شبکه، اطلاعات تجمیع شده از ویژگی‌های آن موسیقی به شبکه داده می‌شود. معماری شبکه در شکل زیر مشاهده می‌شود. شبکه شامل دو بخش است که در بخش ابتدایی، ویژگی‌های انتزاعی‌تری از موسیقی‌های فهرست پخش به‌دست‌آمده و با هم میانگین گرفته می‌شود. این ویژگی‌ها از موسیقی هدف نیز استخراج می‌شود.

درنهایت با استفاده از ویژگی‌های استخراج‌شده، دسته‌بندی دو کلاسه مورد نظر انجام خواهد شد. در شکل ۵، DE لایه Dense است. RE به معنای Relu است. BN برابر با Batch Normalization و DR برابر با Dropout است.

مقاله ارائه شده توسط Vall و همکاران [۳۲] ازنظر ویژگی‌های استفاده شده مشابه با مقاله Vall و همکاران [۳۱] است. تنها تفاوت موجود این است که در این مقاله، یک موسیقی به‌عنوان ورودی به شبکه عصبی داده می‌شود و در خروجی به تعداد فهرست پخش‌ها، نورون موجود است و فهرست پخش‌هایی که موسیقی فعلی احتمالا می‌تواند به آن‌ها تعلق داشته باشد، فعال می‌شوند. معماری شبکه عصبی این مقاله از سه لایه مخفی و ۱۰۰ نورون به ازای هر لایه مخفی تشکیل شده است. تابع فعال‌سازی استفاده شده، تانژانت هایپربولیک است. در لایه‌های شبکه از Batch Normalization و Dropout نیز استفاده شده است.

در مقاله ارائه شده توسط Gharahighehi و همکاران [۳۳] برای حل تعارض شخصی‌سازی تنوع در مقابل دقت در سیستم‌های توصیه‌گر روش جدیدی پیشنهاد شده است. یکی از مهم‌ترین نگرانی‌ها در مورد سیستم‌های توصیه‌کننده پدیده حباب پالایه است. درحالی‌که سیستم‌های توصیه‌گر سعی در شخصی‌سازی اطلاعات دارند، آن‌ها حباب پالایه را در اطراف کاربران تنگ‌تر می‌کنند و آن‌ها را از محتوای گسترده محروم می‌کنند. برای غلبه بر این مشکل، می‌توان فهرست توصیه‌های شخصی را متنوع کرد. یک فهرست متنوع معمولاً محتوای گسترده‌تری را به کاربر ارائه می‌دهد. سیستم‌های توصیه‌گر مبتنی بر نشست انواع توصیه‌هایی هستند که در آن‌ها فقط نشست فعلی کاربر را مورد توجه قرار می‌دهند و بنابراین، با توجه به موارد موجود در نشست فعلی آن‌ها باید قلم بعدی را توصیه کنند. درحالی‌که متنوع‌سازی در دیگر سیستم‌های توصیه‌گر مرسوم به خوبی مورد توجه واقع شده است، اما در سیستم توصیه‌گر مبتنی بر نشست کمتر مورد توجه قرارگرفته است. تنوع و دقت معمولاً همبستگی منفی دارند، یعنی با بهبود یکی دیگر کاهش می‌یابد. در این مقاله افزایش تنوع و دقت بر اساس روش استخراج قاعده و روش نزدیک‌ترین همسایه مبتنی بر نشست ارائه شده است. سرانجام یک رویکرد تعادل عملکرد را پیشنهاد می‌دهند که هم تنوع و هم دقت این سیستم‌ها را بهبود می‌بخشد.

	X_p^1 (D)	...	$X_p^{T_p}$ (D)	x_s (D)
f :	**DE₁** + RE (512)	...	**DE₁** + RE (512)	**DE₁** + RE (512)
	BN₁ + DR (512)	...	**BN₁** + DR (512)	**BN₁** + DR (512)
	DE₂ + RE (512)	...	**DE₂** + RE (512)	**DE₂** + RE (512)
	BN₂ + DR (512)	...	**BN₂** + DR (512)	**BN₂** + DR (512)
	Average (512)			
	Concatenate (1024)			
g :	DE_3 + RE (512)			
	BN_3 + DR (512)			
	DE_4 + RE (512)			
	BN_4 + DR (512)			
	DE_5 + RE (1)			
	Loss			

شکل ۵ . معماری شبکه ارائه شده توسط وال و همکاران

در مقاله ارائه شده توسط Ali و همکاران [۸۵] از مجموعه دادگان موجود در Kaggle شامل اطلاعات حدود ۱۶۰۰۰ آهنگ از Spotify و Genius استفاده کرده است. در این پژوهش، از API های Spotify و Genius جهت استخراج ویژگی‌های موسیقی و آوازها استفاده شده است. ازجمله ویژگی‌های موسیقی هر آهنگ، میزان انرژی، بلندی، محاوره‌ای بودن و آکوستیک بودن آن است. ازجمله ویژگی‌های آواز فرکانس کلمات، فرکانس کلمات بدون تکرار، مثبت بودن یا نبودن احساس و نوع تاپیک مربوطه در آن می‌باشد. در ادامه، یک شبکه سیامس آموزش داده می‌شود. این شبکه از تابع خطای سه‌تایی به منظور مقایسه، آهنگ موردعلاقه کاربر با آهنگی دیگر که موردعلاقه کاربر بوده و آهنگ دیگری که موردعلاقه کاربر نیست، استفاده می‌شود. سپس در فرآیند آموزش فاصله آهنگ هدف با آهنگ موردعلاقه کاربر حداقل و از آهنگ دیگر حداکثر گردد. به این ترتیب، آهنگ‌های مشابه با آهنگ موردعلاقه کاربر بیشترین امتیاز و سایر آهنگ‌ها کمترین امتیازها را به خود اختصاص می‌دهند. به‌علاوه، این شبکه از دو زیرشبکه چندلایه رو به جلو تشکیل می‌شود. دو زیرشبکه ویژگی‌های مربوط به موسیقی و آواز آهنگ را به‌عنوان ورودی دریافت می‌کنند و یک بردار ویژگی به‌عنوان خروجی تولید می‌کند. بردارهای ویژگی ورودی و خروجی به هم متصل شده و به‌عنوان ورودی به یک شبکه چندلایه رو به جلو در نظر گرفته می‌شود و میزان شباهت ویدئوها را تعیین می‌کند. درنهایت بر اساس امتیاز محاسبه شده برای آهنگ‌های مختلف و با در نظر گرفتن حد آستانه، آهنگ‌هایی با امتیاز بیشتر از حد آستانه به‌عنوان شبیه‌ترین آهنگ‌ها به سلیقه کاربر، به کاربر پیشنهاد می‌شود.

در مقاله ارائه شده توسط Burch و همکاران [۸۶] سیستمی طراحی شد که از ۳ جز اصلی تشکیل شده است که دو جز آن برای پیشنهاد آهنگ‌های مشابه به علاقه کاربر و جز دیگر برای تعیین احساسات و

عواطف موجود در آهنگ برای ایجاد فضای متناسب با احساسات تشخیص داده شده بر روی ویدئو استفاده شده است. در جز نخست از الگوریتم خوشه‌بندی جهت گروه‌بندی آوازها بر مبنای ویژگی‌های آهنگ‌ها استفاده می‌شود. به این ترتیب برای هر کاربر مجموعه آهنگ‌های مشابه تعیین می‌شود. درجز دوم از یک الگوریتم تطبیق الگوی سریع جهت تطبیق سیگنال شبیه‌ترین آهنگ‌ها به آهنگ موردعلاقه کاربر در هر گروه مورداستفاده قرار می‌گیرد. در جز سوم، از یک شبکه عصبی پیچشی جهت تعیین عواطف و احساسات موجود در هر آهنگ نظیر شادی و اندوه استفاده می‌شود. ورودی شبکه از طریق بازنمایی به روش طیف مل آهنگ مانند یک تصویر در نظر گرفته می‌شود به این ترتیب که محور عمودی زمان و محور افقی فرکانس و شدت رنگ خواهد بود. احساسات و عواطف استخراج‌شده در فرآیند پس پردازش در محیط موتور Unreal۵ برای ایجاد فضای متناسب با حس موجود در آهنگ مورداستفاده قرار می‌گیرد.

ازجمله مزایای این روش‌ها انعطاف‌پذیری در یادگیری ترجیحات پویای کاربر، قابلیت تعادل بین اکتشاف (Exploration) و بهره‌برداری (Exploitation)، استخراج خودکار ویژگی‌های پیچیده از موسیقی و قابلیت ترکیب اطلاعات چندرسانه‌ای (صوت، متن، احساس) می‌باشد. ازجمله معایب این روش نیاز به داده‌های رفتاری زیاد برای آموزش مدل، پیچیدگی در پیاده‌سازی و تنظیم پارامترها، نیاز به منابع محاسباتی بالا و تفسیرپذیری پایین به دلیل ماهیت مدل‌های عمیق می‌باشد.

سیستم‌های ترکیبی (محتوا + پالایش مشارکتی)

این سیستم‌ها از ترکیب روش‌های مبتنی بر محتوا (ویژگی‌های صوتی) و پالایش مشارکتی استفاده می‌کنند. برخی نیز از اطلاعات زمینه‌ای مانند فعالیت کاربر (ورزش، مطالعه و ...) برای بهبود توصیه‌ها بهره می‌برند.

سیستم‌های توصیه‌گر مبتنی محتوای سنتی معمولا ابتدا یک سری ویژگی از صوت موسیقی استخراج می‌کنند (مانند MFCC) و سپس با استفاده از این ویژگی‌ها، پیشنهادهای مربوط به هر کاربر را به دست می‌آورند. به دلیل این‌که ویژگی‌های اولیه استخراج‌شده، برای مسئله توصیه موسیقی طراحی نشده‌اند، معمولا عملکرد این روش‌ها چندان خوب نیست. در مقاله ارائه شده توسط Wang و همکاران [۱۸] با ترکیب شبکه‌های باور عمیق و مدل‌های گرافیکی احتمالی، دو مرحله ذکرشده در بالا با هم ترکیب‌شده‌اند و به‌صورت همزمان استخراج ویژگی از صوت موسیقی و توصیه موسیقی انجام می‌شود. در ادامه این مقاله، این روش با روش پالایش مشارکتی ترکیب‌شده است و نتایجی بهتر از روش‌هایی صرفاً مبتنی بر پالایش مشارکتی و روش‌های مبتنی بر ویژگی سنتی ارائه داده است.

در مقاله ارائه شده توسط Barragáns-Martínez و همکاران [۲۲] سیستم توصیه‌گری ارائه شده است که ترکیبی از رویکردهای مبتنی بر محتوا و پالایش مشارکتی است.سیستم ارائه شده این امکان را مهیا می‌کند که کاربران با نسبت دادن تگ‌هایی به اقلام مختلف، هر قلم را توصیف کنند. با استفاده از تگ‌ها، برای کاربران و برای اقلام، ابر کاربر و ابر قلم ایجاد می‌شود. ابر برچسب قلم از فهرست تگ‌هایی ساخته می‌شود که به یک قلم نسبت داده شده است. ابر تگ‌های کاربر از تگ‌هایی که کاربر هنوز به هیچ قلمی نداده است ساخته می‌شود. اهمیت هر برچسب نسبت داده شده به هر قلم، متناسب با تعداد دفعاتی است که توسط کاربران به آن نسبت داده شده است. پیشنهاد مبتنی بر محتوا به این صورت است که با مقایسه ابر تگ‌های اقلام با ابر تگ‌های یک کاربر، بهترین قلم به یک کاربر به دست می‌آید و به او پیشنهاد می‌شود. نوع دیگر پیشنهاد، استفاده از روش پالایش مشارکتی است. برای این کار در این سیستم با جمع وزن‌دار تگ‌های کاربرانی که به یک قلم علاقه نشان داده‌اند، ابر برچسب یک کاربر نوعی که این قلم برای او مناسب است به دست می‌آید. با مقایسه این ابر تگ‌ها با ابر برچسب کاربران مختلف، کاربر مناسب یک قلم پیدا می‌شود. در واقع این یک روش پالایش مشارکتی مبتنی بر کاربر است چون درنهایت با مقایسه کاربران با هم، پیشنهاد انجام می‌شود. این سیستم توانسته است coverage و diversity را بهبود ببخشد.

ازجمله مزایای این روش کاهش مشکل شروع سرد با استفاده از ویژگی‌های محتوایی و افزایش دقت با ترکیب سیگنال‌های مختلف می‌باشد. ازجمله معایب این روش پیچیدگی محاسباتی بالاتر و نیاز به مهندسی ویژگی‌های دقیق است.

۴-۳- سیستم‌های توصیه‌گر حوزه تلویزیون

سیستم‌های مبتنی بر روش‌های سنتی مبتنی بر محتوا

این روش‌ها از ویژگی‌های محتوایی مانند سبک، بازیگران، شبکه پخش‌کننده و ... برای پیشنهاد برنامه‌ها استفاده می‌کنند.

در رابطه با تماشای تلویزیون این چالش وجود دارد که افراد معمولا شبکه موردعلاقه خود را تماشا می‌کنند و ممکن است حتی اگر برنامه‌ای متناسب با علاقه آن‌ها در شبکه‌ای دیگر پخش شود، متوجه آن نشوند. برای رفع این مشکل، در مقاله ارائه شده توسط

Engelbert و همکاران [۳۴] سیستم توصیه‌گری مبتنی بر طبقه‌بندی بیزین بر اساس بازخورد صریح کاربر و این‌که کاربر درگذشته چه برنامه‌هایی را مشاهده کرده است، ارائه شده است. با استفاده از این سیستم توصیه‌گر، برنامه‌هایی که ممکن است کاربر به آن‌ها علاقه داشته باشد ضبط می‌شود و کاربر می‌تواند بعدا به آن‌ها دسترسی پیدا کند. ویژگی‌هایی که در این سیستم استفاده می‌شود عبارت‌اند از شبکه تلویزیونی که برنامه از آن پخش‌شده است، عنوان برنامه، زیرنویس برنامه، دسته برنامه (مثلاً فیلم، مسابقه تلویزیونی یا ...)، سبک، بازیگران، توضیحات و سال ساخت.

در این پژوهش دو نوع پروفایل برای هر کاربر تعریف می‌شود. در ابتدا که کاربر در سیستم ثبت‌نام می‌کند، توسط پنل کاربری خود، علایق خود (مثلاً سبک موردعلاقه، بازیگر یا کارگردان موردعلاقه) را وارد می‌کند و پروفایل ثابت خود را تکمیل می‌کند. این پروفایل در طول زمان ثابت می‌ماند و تغییری نمی‌کند. سابقه تماشای برنامه‌های مختلف توسط کاربر به‌عنوان بازخورد ضمنی در نظر گرفته می‌شود و با توجه به رفتار کاربر، هر برنامه تماشا شده در یکی از دو دسته «موردعلاقه» یا «عدم علاقه» قرار می‌گیرد. با استفاده از این بازخوردها، پروفایل پویای کاربر تهیه می‌شود. کاربر این امکان را دارد که به هر برنامه به‌صورت صریح بازخورد دهد و پروفایل خود را دقیق‌تر کند. هدف این سیستم این است که در شروع کار که اطلاعات زیادی از برنامه‌های تماشا شده توسط کاربر وجود ندارد، از پروفایل ثابت کاربر استفاده شود و به تدریج، سیستم توصیه‌گر به سراغ استفاده از پروفایل پویای کاربر برود. دسته‌بندی یک برنامه جدید در دو دسته «موردعلاقه» یا «عدم علاقه» با استفاده از طبقه‌بندی بیزین و بر اساس ویژگی‌های استخراج‌شده از هر برنامه تلویزیونی و با استفاده از هر دو پروفایل یک کاربر انجام می‌شود.

ازجمله مزایای این روش‌ها عدم نیاز به داده‌های تاریخی کاربران دیگر (حل مشکل شروع سرد) و شفافیت در توصیه‌ها (قابل تفسیر بودن معیارهای پیشنهاد) می‌باشد. ازجمله معایب این روش‌ها محدودیت در کشف سلیقه‌های پیچیده کاربران و وابستگی به کیفیت متادیتای برنامه است.

سیستم‌های جدید مبتنی بر محتوا

این روش‌ها از ویژگی‌های زمینه‌ای (زمان، مکان، احساس کاربر) و تکنیک‌های هوش مصنوعی (شبکه‌های عصبی، مدل‌های گرافی) برای شخصی‌سازی عمیق‌تر استفاده می‌کنند.

مقاله ارائه شده توسط Quan و همکاران [۳۵] این موضوع را هدف گرفته است که معمولا مشاهده تلویزیون به‌صورت گروهی انجام می‌شود. برای توصیه برنامه به گروهی از کاربران، ابتدا با استفاده از شبکه‌های بیزین، و بر اساس سابقه برنامه‌هایی که هر کاربر تماشا کرده است، سبک و برنامه‌هایی که هر کاربر ترجیح می‌دهد به دست می‌آید و سپس با استفاده از پروسه سلسله‌مراتبی آنالیزی[1] سلیقه گروهی افراد به دست می‌آید. تأکید این پژوهش بر روی توصیه برنامه به گروهی از کاربران است و تنها بخش فردی این روش، آموزش یک شبکه بیزین با استفاده از اعمال برآورد درست‌نمایی بیشینه بر روی سابقه تماشای کاربر، برای به دست آوردن سلیقه آن کاربر است.

مقاله ارائه شده توسط Oh و همکاران [۳۶] بر روی یک چالش مهم در توصیه برنامه تلویزیون تمرکز کرده است. چالش مطرح شده این است که تماشای تلویزیون به‌صورت گروهی اتفاق می‌افتد و درنتیجه اطلاعاتی که از تاریخچه رفتار کاربران در دسترس است مربوط به یک کاربر خاص نیست بلکه افراد مختلف یک خانواده را شامل می‌شود. برای حل این چالش، در این مقاله یک روش ساخت پروفایل وابسته به زمان پیشنهاد شده است. این روش با تجزیه و تحلیل گزارش ساعت بر اساس زمان روز، پروفایل ایجاد می‌کند. ایده اصلی این است که معمولا بینندگان برنامه که اعضای مختلف یک خانواده هستند با توجه به زمان روز تغییر می‌کند. به‌عنوان‌مثال، تماشای برنامه تلویزیونی در نیمه شب به احتمال زیاد توسط والدین انجام می‌شود زیرا کودکان معمولاً زود می‌خوابند. در مقابل، تاریخچه تماشای تلویزیون در حدود ساعت ۴ تا ۵ عصر غالبا مربوط به کودکان است زیرا والدین در آن زمان معمولاً در محل کار خود هستند. بنابراین، برای تشخیص هر اولویت باید گزارش ساعت را بر اساس زمان روز تجزیه و تحلیل کرد. در همین راستا در این مقاله پس از تقسیم ساعات زمانی روز برای ساخت پروفایل مجزا برای هر کاربر از یک الگوریتم خوشه‌بندی سلسله‌مراتبی استفاده کرده است.

هدف مقاله ارائه شده توسط Turrin و همکاران [۳۷]، پیشنهاد برنامه تلویزیونی به‌صورت شخصی‌سازی شده در هر زمان به هر کاربر است. از آن‌جایی که کاربران مختلف در بازه‌های زمانی خاص ممکن است علاقه‌مندی‌های متفاوتی داشته باشند. در این پژوهش، سه مفهوم «بهترین شبکه»، «بهترین شبکه برای هر کاربر» و «بهترین شبکه برای هر کاربر و هر بخش زمانی» به ترتیب معادل با شبکه‌ای که بیشترین زمان تماشا را داشته باشد، شبکه‌ای که بیشترین زمان تماشا برای یک کاربر خاص را داشته باشد و شبکه‌ای که بیشترین زمان تماشا برای یک کاربر خاص و در یک بخش زمانی خاص را

[1] Analytic Hierarchy Process (AHP)

داشته باشد تعریف شده است. چالشی که در این مقاله مطرح شده است این است که ممکن است بر اساس سابقه تماشای کاربر، این طور برداشت شود که یک کاربر در بازه زمانی ۱۰ تا ۱۱ صبح، به برنامه کودک در شبکه دیزنی علاقه‌مند است ولی در ساعت ۱۱ تا ۱۲ صبح به برنامه کودک در شبکه دیزنی علاقه‌مند نیست. دلیل این موضوع می‌تواند کم بودن سابقه تماشای کاربر در بازه زمانی دوم باشد. برای رفع این مشکل، سابقه‌های تماشای کاربر در هر بخش زمانی با بخش‌های زمانی اطرافش ترکیب می‌شود. به این کار هموارسازی گفته می‌شود. ویژگی‌های استفاده شده در این مقاله، سبک برنامه (مثلاً برنامه کودکان)، زیر سبک (مثلاً کارتون)، شبکه (مثلاً دیزنی) و فهرست بخش‌های زمانی‌ای که کاربر در آن‌ها این برنامه را تماشا کرده است، می‌باشد.

در مقاله ارائه شده توسط da Silva و همکاران [۳۸] از دو دسته ویژگی‌های زمینه‌ای (تاریخ و زمان و مکان ساخت برنامه تلویزیونی و مکان کاربر) و ویژگی‌های صریح که مستقیماً از کاربر گرفته می‌شود (زبان، جنسیت، سن و شغل) و ویژگی‌هایی که مانند سبک که به‌صورت اتوماتیک در برنامه‌های تلویزیونی استخراج می‌شود، استفاده می‌شود. در این پژوهش برای به دست آوردن سبک کلاس‌ها، چندین روش یادگیری ماشین، مانند طبقه‌بندی درخت تصمیم، طبقه‌بندی بیزین ساده، شبکه عصبی پرسپترونی و استدلال مبتنی بر مورداستفاده شده است.

مقاله ارائه شده توسط Zibriczky و همکاران [۳۹] به‌طور خاص بر روی تلویزیون‌های خطی (تلویزیون‌های عادی که برنامه‌ها را به‌صورت از پیش برنامه‌ریزی شده برای همه کاربران نمایش می‌دهند) در مقابل VODها (مانند نتفلیکس) تمرکز دارد. به دلیل این‌که در تلویزیون‌های خطی، بازخورد صریح (امتیازدهی) از کاربر نداریم، نمی‌توان از روش‌های رایج پالایش مشارکتی برای این نوع تلویزیون استفاده کرد و درنتیجه باید با استفاده از بازخوردهای ضمنی و احتمالا نویزی، توصیه انجام شود. از چالش‌های دیگر تلویزیون خطی می‌توان به این اشاره کرد که معمولا بخش زیادی از برنامه‌های تلویزیون خطی جدید هستند و مشکل شروع سرد برای این برنامه‌ها وجود دارد. در این پژوهش از log نحوه تعامل کاربر با تلویزیون به‌عنوان بازخورد ضمنی استفاده می‌شود. رفتارهایی ازجمله عوض کردن بسیار سریع کانال‌ها یا تماشای هر برنامه برای مدت کوتاهی، نویز محسوب می‌شود که در مرحله اول این نویز باید حذف شود. در روش مبتنی بر محتوا، از ویژگی‌های عنوان، سبک، کانال پخش‌کننده، دسته[1] و رده‌بندی سنی استفاده می‌شود. برای مقایسه کاربر با برنامه‌های تلویزیونی، از معیار شباهت کسینوسی استفاده شده است. در روش پالایش مشارکتی، از IALS۱ استفاده شده است. این روش، روشی مبتنی بر تجزیه ماتریس برای مجموعه داده با بازخورد صریح یا ضمنی است. در روش ترکیبی، از ترکیب دو روش بالا و اطلاعات metadata استفاده شده است. برای حل مشکل شروع سرد، پیشنهاد شده است که برنامه‌ها بر اساس معیاری با هم تجمیع شوند. تجمیع بر اساس کانال پخش کننده برنامه صورت گرفته است. یعنی در واقع مساله به پیش‌بینی کانال موردعلاقه کاربر از بین کانال‌های تلویزیون تبدیل می‌شود. دلیل تجمیع بر روی کانال‌ها این است که معمولا کاربرها به برنامه خاصی در کانال خاصی علاقه‌مند هستند و در یک بازه زمانی آن را مشاهده می‌کنند و سپس به سراغ برنامه بعدی موردعلاقه خود در کانال بعد می‌روند. با این تعاریف، دو روش برای پیشنهاد کانال در هر بازه زمانی پیشنهاد شده است که یکی از آن‌ها iTALS است که بر اساس بازخورد ضمنی کار می‌کند. روش پیشنهادی دیگر، در نظر گرفتن طول زمان مشاهده کاربر از هر کانال و استفاده از آن در سیستم توصیه‌گر است.

در مقاله ارائه شده توسط Song و همکاران [۴۰]، چهار نوع ویژگی‌های زمینه معرفی شده است.

ویژگی‌های زمینه کاربر: به دو نوع ثابت و پویا تقسیم می‌شود. ویژگی‌های ثابت شامل نمایه کاربر (نام کاربری، نام، سن و جنسیت کاربر) ترجیح ثابت کاربر (زبان زیرنویس و زبان صدایی که کاربر ترجیح می‌دهد، نوع محتوای موردعلاقه، بازیگر موردعلاقه و کارگردان موردعلاقه) است. این اطلاعات به‌صورت مستقیم از کاربر دریافت می‌شود. ویژگی‌های پویا شامل اطلاعاتی است که به‌صورت دائم در حال تغییر است، مثلاً موقعیت جغرافیایی کاربر، agendaی کاربر و احساس کاربر. بعضی تغییر رفتارهای کاربر، مثلاً تماشای مسابقات المپیک توسط کاربری که برنامه ورزشی نگاه نمی‌کند نیز جزو ویژگی‌های پویا محسوب می‌شود:

- ویژگی‌های زمینه‌ای مرتبط با دستگاه کاربر: اندازه صفحه، میزان حافظه، امکانات ارتباطی مثل G۴ یا وای‌فای و وضعیت دستگاه (روشن، خاموش و سطح باتری).
- ویژگی‌های زمینه‌ای شبکه: نوع دسترسی به شبکه، پهنای باند، کیفیت سرویس اینترنت (میزان تاخیر و میزان از دست رفتن پاکت‌ها).
- ویژگی‌های زمینه‌ای سرویس: اطلاعات عمومی در رابطه با یک ویدئو مثل عنوان، کلیدواژه، کارگردان، بازیگر، توضیحات و

[1] Category

زبان اثر. همچنین اطلاعات تخصصی تر مثل فرمت فایل (MPEG2 یا MPEG4 یا AVC)، رزولوشن و کمترین نرخ بیت موردنیاز برای تماشای این فایل.

برای توصیه به کاربر، از ساختار سیستم توصیه‌گر لایه‌ای (در واقع یک شبکه‌عصبی پرسپترون چند لایه) استفاده شده است. لایه ابتدایی این ساختار، لایه ویژگی‌ها است و هر گره[1] یکی از ویژگی‌های زمینه‌ای است. بعد از این لایه یک لایه با دو نورون وجود دارد و در لایه آخر یک نورون وجود دارد که میزان احتمال این‌که ویدئوی فعلی برای این کاربر پیشنهاد شود را مشخص می‌کند.

مقاله ارائه شده توسط Chaudhry و همکاران [۴۱] به‌طور خاص بر حل مساله پیشنهاد برنامه به کاربر در تلویزیون‌هایی که به‌صورت همگانی پخش می‌شوند[2] تأکید دارد. این مقاله رابطه بین هر کاربر و برنامه‌های تلویزیونی را ناهمگون در نظر می‌گیرد و به استخراج روابط ناهمگون بین کاربر مدنظر و برنامه‌های تلویزیونی، با استفاده از بازخورد ضمنی، می‌پردازد.

روش دیگر ارائه شده برای توصیه یک برنامه جدید، یک روش هیبرید است که در آن از روی گراف موجود، ماتریس‌های کاربر-توئیت (این‌که چه کاربری چه توئیتی را نوشته است)، توئیت-آدرس (این‌که یک هر توئیت به چه صفحه‌ای از DBPedia اشاره دارد)، آدرس-آدرس (ارتباط معنایی بین لینک‌های DBPedia) وآدرس-برنامه (این‌که هر برنامه توسط چه صفحه‌ای از DBPedia توضیح داده شده است) ساخته شده و پس از ضرب ماتریسی تمام این ماترس‌ها، ماتریس کاربر-برنامه به دست می‌آید. برای محاسبه ماتریس آدرس-آدرس از روش ارائه شده در مقاله [۴۳] زیر استفاده شده است که روش مبتنی بر محتوا است.

هدف مقاله ارائه شده توسط Yuan و همکاران [۴۴]، ارائه یک سیستم توصیه‌گر برای IPTV مورداستفاده در پردیس دانشگاه لنکستر است. در مساله بررسی شده، سبک، شبکه پخش کننده برنامه و نام برنامه درمجموعه داده آموزشی جمع‌آوری‌شده موجود است. برای انتخاب ویژگی مناسب، برای هر یک از ویژگی‌ها، آنتروپی نرمال‌شده محاسبه شده است و درنهایت بهترین گزینه، یعنی نام برنامه، انتخاب شده است. برای انتخاب ویژگی‌های زمینه‌ای نیز با تحلیلی مشابه تحلیل بالا، از بخشی از زمان روز که برنامه در آن پخش می‌شود و زنده بودن پخش برنامه یا این‌که برنامه به‌صورت ویدئوی همیشه در دسترس استفاده شده است و روزی از هفته که برنامه در آن روز پخش‌شده است، کنار گذاشته شده است. برای ساختن پروفایل کاربر، می‌توان از تخصیص نهان دیریکله استفاده کرد. برای این‌که بتوان ویژگی‌های زمینه‌ای را نیز دخیل کرد، مدل توسعه یافته شده‌ای به نام تخصیص نهان دیریکله آگاه از زمینه ارائه شده است. همچنین جهت جلوگیری از همگرایی به اکسترمم محلی توسط الگوریتم بیشینه‌سازی امید ریاضی، از نمونه‌برداری گیبز استفاده شده است.

پاسخ به سوال کلیدی «برای تماشای بعدی ویدئویی توصیه می‌شود؟» مدت‌ها است که در مرکز تحقیقات سیستم‌های توصیه‌گر قرار دارد. بااین‌حال، بهره‌برداری کافی از سرنخ‌های پنهان شده در توالی اقدامات تاریخچه جستجوی کاربر، به منظور آشکار کردن اهداف کوتاه مدت کاربران، اخیراً مورد توجه واقع شده است. در مقاله ارائه شده توسط Symeonidis و همکاران [۴۵]، بر اساس یک سناریوی کاربردی در دنیای واقعی، یک ماتریس احتمال انتقال مبتنی بر مارکوف ساخته می‌شود تا ترجیحات کوتاه مدت افراد به‌طور موثر نشان داده شود. بر اساس آزمایشات در این مقاله نشان داده شد در نظر گرفتن توالی جستجوها در تاریخچه جستجوی کاربر برای بهبود اثربخشی پیشنهادها کمک می‌کند، زیرا گرفتن ترجیحات کوتاه مدت کاربران در ارتباط با اقلام با طول عمر کوتاه مانند برنامه‌های تلویزیونی بسیار مهم است (اخبار، برنامه‌های تلویزیونی، و غیره.).

Dudekula و همکاران [۸۷] بر اساس تاریخچه برنامه‌های مشاهده شده توسط کاربر و تشخیص چهره به پیشنهاد برنامه‌های مرتبط می‌پردازند. در مرحله نخست، تصویر کاربر توسط بخش دوربین تلویزیون هوشمند گرفته می‌شود. سپس با استفاده از الگوریتم شبکه عصبی پیچشی مورد ارزیابی قرار می‌گیرد. این شبکه بر اساس پیکره تصاویر چهره افراد مشهور و پیکره تصاویر برچسب خورده آموزش داده می‌شود. در مرحله دوم، ویژگی‌های تصویر کاربر به‌وسیله شبکه عصبی پیچشی استخراج می‌شود و با پروفایل کاربر در پیکره دادگان تطبیق داده می‌شود. در مرحله سوم، بر مبنای تاریخچه کاربر، روش‌های پالایش مبتنی بر محتوا با پالایش مبتنی بر همکاری ترکیب می‌شود. در پالایش مبتنی بر محتوا، از متادیتاهایی نظیر سبک، بازیگران و کارگردان جهت پیشنهاد برنامه‌های مشابه با برنامه‌های موردعلاقه کاربر استفاده می‌کند. در پالایش مبتنی بر مشارکت، از بازخوردهای دریافتی توسط کاربران مشابه با کاربر هدف و میزان رای مثبت به هر برنامه استفاده می‌شود. درنهایت، با استفاده از ترکیبی از فیلترینگ‌های انجام‌شده به کاربر هدف برنامه‌های مناسب پیشنهاد می‌گردد.

[1] Node

[2] broadcast

ازجمله مزایای این روش‌ها در نظر گرفتن شرایط لحظه‌ای کاربر (مثلاً زمان روز) و انعطاف‌پذیری بیشتر با داده‌های ناهمگون (مانند توئیت‌ها) است. ازجمله معایب این روش‌ها پیچیدگی محاسباتی و نیاز به داده‌های غنی و مشکل در تفسیرپذیری نتایج می‌باشد.

سیستم‌های ترکیبی (محتوا + پالایش مشارکتی)

این روش‌ها هم از محتوای برنامه‌ها و هم از رفتار کاربران مشابه برای توصیه استفاده می‌کنند.

در مقاله ارائه شده توسط Barraza-Urbina و همکاران [۴۲]، به طراحی یک سیستم توصیه‌گر برای یک شبکه تلویزیونی که امکان مشاهده مجدد برنامه را به کاربران می‌دهد پرداخته‌است. سه چالش اصلی در این مساله وجود دارد که آن را از سیستم‌های توصیه‌گر سنتی متفاوت می‌کند. مساله اول این است سیستم مورد نظر، امکان امتیازدهی کاربران را فراهم نکرده است. مساله دوم این است که که هیچ سابقه تماشای برنامه‌ای توسط کاربران ذخیره نشده است و مساله سوم این است که برنامه‌های جدید تلویزیون به‌مرور زمان به این سیستم اضافه می‌شوند و برنامه‌های قدیمی از سیستم حذف می‌شوند. بنابراین، هدف این پژوهش ارائه سیستم توصیه‌گری برای توصیه برنامه تلویزیونی به کاربر ناشناس و صرفاً بر اساس نشست فعلی کاربر و اطلاعات مرتبط با توئیتر کاربر مورد نظر است. با در نظر گرفتن فعالیت کاربران در توئیتر، یک مدل داده گرافی ناهمگون ارائه شده است. در این مدل گره «کاربر»ها مربوط به کاربرهایی است که توئیت مرتبطی در رابطه با یکی از برنامه‌های تلویزیونی شبکه مورد نظر نوشته‌اند. هر کاربر به توئیتی که نوشته است متصل است و هر توئیت می‌تواند شامل یک پیوست یا یک لینک باشد. همچنین احساس مرتبط با هر توئیت توسط یک الگوریتم آماده استخراج‌شده و در گره مربوط به هر توئیت ثبت می‌شود. گره‌های مرتبط با اپیزود و برنامه نیز موجود است که هر گره اپیزود به یک گره برنامه متصل است و هر گره توئیت می‌تواند به یک اپیزود یا یک برنامه متصل باشد. ارتباط بین توئیت‌ها و برنامه‌ها با استفاده از هشتگ‌های به‌کاررفته در هر توئیت به دست می‌آید. ارتباط بین اپیزودها و برنامه‌ها، با استفاده از تاریخ و ساعت نوشته شدن توئیت به دست می‌آید. جهت کامل‌تر کردن اطلاعات گراف، هر برنامه به یک گره از آدرس صفحه مربوط به آن برنامه در DBpedia و آدرس صفحه IMDB متصل خواهد شد.

اولین روشی که برای توصیه برنامه جدید به کاربر پیشنهاد شده است، استفاده از پالایش مشارکتی است. با استفاده از گراف تهیه شده، می‌توان ماتریس کاربر-توئیت و هچنین ماتریس توئیت-برنامه را طبق لینک‌های بین گره‌ها به دست آورد. در مقاله بیان‌شده است که در حال تحقیق برای تعریف امتیازی برای این یال‌ها هستند و احتمالا در آینده از امتیاز احساس توئیت‌های کاربر استفاده خواهد شد. در حال حاضر لینک‌ها را می‌توان باینری در نظر گرفت. با ضرب دو ماتریس کاربر-توئیت و توئیت-برنامه، ماتریس کاربر-برنامه به دست می‌آید و با روش سنتی پالایش مشارکتی، توصیه برنامه جدید انجام می‌شود.

با توجه به اینکه مقاله ارائه شده توسط Barragáns-Martínez و همکاران [۲۲] در هر دو دامنه موزیک و برنامه تلویزیونی ارائه شده است توضیحات آن در بخش Music ارائه شده است.

ازجمله مزایای این روش‌ها کاهش مشکل شروع سرد با ترکیب دو روش و بهبود دقت با استفاده از همکاری بین کاربران می‌باشد. ازجمله معایب این روش‌ها نیاز به داده‌های تاریخی کاربران (مشکل در محیط‌های با داده‌های کم) و پیچیدگی در یکپارچه‌سازی دو روش می‌باشد.

۴-۴- سیستم‌های توصیه‌گر چندرسانه‌ای

سیستم‌های توصیه‌گر چندرسانه‌ای برای کمک به کاربران در انتخاب محتوای چندرسانه‌ای مناسب طراحی‌شده‌اند و با ترکیب الگوریتم‌های توصیه‌گر و ویژگی‌های چندرسانه‌ای، پیشنهادات شخصی‌سازی شده ارائه می‌دهند. این سیستم‌ها با چالش‌هایی مانند پردازش داده‌های پیچیده چندرسانه‌ای، شکاف معنایی[1]، و نیاز به مقیاس‌پذیری در مواجهه با حجم زیاد داده‌ها روبه‌رو هستند. در این سیستم‌ها نیز از روش‌های مختلفی مانند پالایش مشارکتی، پالایش مبتنی بر محتوا، و سیستم‌های ترکیبی برای ارائه توصیه‌ها استفاده می‌شود. این سیستم‌ها در حوزه‌های مختلفی مانند موزه‌های مجازی، شبکه‌های اجتماعی، آموزش الکترونیکی، و سرگرمی (مثل موسیقی و فیلم) به کار می‌روند. با رشد داده‌های چندرسانه‌ای، نیاز به سیستم‌های توصیه‌گر هوشمند و کارآمد بیشتر شده است و تحقیقات آینده بر ترکیب فناوری‌های جدید مانند اینترنت اشیا و داده‌های حجیم[2] متمرکز خواهد بود[۹۲].

در مقاله‌ای که توسط دلجو و همکاران [۹۳] رائه شده است، با بررسی بیش از ۱۵ سال تحقیقات، سیستم‌ها را بر اساس نوع رسانه، تکنیک‌های استخراج ویژگی‌ها، و الگوریتم‌های توصیه دسته‌بندی کرده‌اند. در این پژوهش با تأکید بر اهمیت ویژگی‌های بصری و

[1] Semantic Gap

[2] Big Data

شنیداری در مدلسازی ترجیحات کاربران موردبررسی قرار می‌گیرد و از محتوای چندرسانه‌ای (صوتی، تصویری، متنی) برای پیشنهاد اقلام رسانه‌ای (مانند موسیقی، فیلم) و غیررسانه‌ای (مانند پوشاک، غذا) استفاده می‌کنند.

در مقاله ارائه شده توسط دلجو و همکاران [۹۴] به بررسی تحقیقات پیشرفته در سیستم‌های توصیه‌گر چندرسانه‌ای می‌پردازد. در این پژوهش، تمرکز اصلی بر روش‌هایی است که محتوای چندرسانه‌ای را به‌عنوان اطلاعات جانبی در مدل‌های توصیه‌گر ادغام می‌کنند. این ویژگی‌ها برای توصیه اقلام رسانه‌ای یا غیررسانه‌ای (مانند تصاویر لباس) به کار می‌روند. به‌علاوه، به تشریح رویکردهای پردازش محتوا و الگوریتم‌های نوین MMRS (ترکیبی، عصبی، مبتنی بر گراف) می‌پردازد. ازجمله کاربردهای MMRS استفاده به‌عنوان سیستم‌های توصیه‌گر در حوزه‌هایی چون غذا، مد، موسیقی و ویدیو می‌باشد.

جیانگ و همکاران [۹۴]، یک سیستم توصیه‌گر چندوجهی جدید معرفی می‌کند که از مدل‌های انتشار گراف برای بهبود یادگیری نمایندگی کاربران با آگاهی از چندوجهی استفاده می‌کند. این روش با ترکیب یک مدل انتشار گراف چندوجهی و یادگیری تضاد بین وجهی، همترازی بهتری بین ویژگی‌های چندوجهی و مدل‌سازی تعاملات کاربر- قلم ایجاد می‌کند. DiffMM از قابلیت‌های تولیدی مدل‌های انتشار برای ایجاد یک گراف کاربر- قلم آگاه از وجهی استفاده می‌کند و از طریق مکانیسم تزریق سیگنال وجهی، اطلاعات چندوجهی را در فرآیند انتشار ادغام می‌کند. این روش همچنین از یادگیری خودنظارتی برای غلبه بر پراکندگی داده‌ها استفاده میکند و تفسیرپذیری توصیه‌ها را بهبود میبخشد. نتایج نشان میدهد که DiffMM در سناریوهای توصیه چندوجهی، به‌ویژه در داده‌های پراکنده، مؤثرتر است.

مالیتستا و همکاران [۹۶]، به معرفی یک چارچوب رسمی برای سیستم‌های توصیه‌گر چندرسانه‌ای با استفاده از یادگیری عمیق چندوجهی می‌پردازد و سه سوال کلیدی «کدام داده‌های چندوجهی»، «چگونگی پردازش ویژگی‌ها» و «زمان ادغام وجه‌ها» را بررسی می‌کند. نویسندگان با مرور جامع ادبیات موجود، الگوهای تکرارشونده در سیستم‌های توصیه‌گر چندرسانه‌ای را شناسایی و آن‌ها را در یک طرح نظری یکپارچه سازماندهی می‌کنند. این پژوهش یک چارچوب چندمرحله‌ای شامل داده‌های ورودی چندوجهی، پردازش ویژگی‌ها، نمایش چندوجهی، و ادغام ویژگی‌ها ارائه می‌دهد و آن را بر روی چندین مدل پیشرفته اعتبارسنجی می‌کند. نتایج مقایسه‌ای نشان می‌دهد که سیستم‌های چندوجهی مانند LATTICE و FREEDOM در معیارهای دقت عملکرد بهتری دارند، اما در معیارهای فراتر از دقت مانند تنوع و نوآوری، روش‌های کلاسیک بهتر عمل می‌کنند. چالش‌های فنی مانند عدم وجود داده‌های چندوجهی کامل، استفاده از استخراج‌کننده‌های ازپیش‌آموخته، و انتخاب راهبردهای نمایش و ادغام به‌عنوان مسائل باز مطرح شده‌اند.

۵- سیستم‌های توصیه‌گر مطرح در صنعت بین‌الملل

۵-۱- سیستم توصیه‌گر نتفلیکس

نتفلیکس از یک سیستم رتبه‌بندی مبتنی بر جدول دوبعدی استفاده می‌کند، جایی که رتبه‌بندی اتفاق می‌افتد:

- درون هر ردیف (قوی‌ترین توصیه‌ها در سمت چپ)
- در میان ردیف‌ها (قوی‌ترین توصیه‌ها در بالا)

نتفلیکس برای سیستم توصیه‌گرخود از بحث شخصی‌سازی اطلاعات استفاده می‌کند. برای این منظور مثلاً در فهرست برنامه‌های پیشنهادی نتفلیکس تصاویری از برنامه مورد نظر که به هر کاربر نمایش داده می‌شود بر اساس تمایلات او خصوصی‌سازی می‌شوند. [۴۶]

روش‌های یادگیری ماشین زیر برای شخصی‌سازی اکانت‌ها در نتفلیکس استفاده می‌شود:

- یادگیری عمیق
- یادگیری تقویتی

در شکل ۶ هر سطر یک موضوع خاص را نشان می‌دهد (به‌عنوان مثال Top 10، Trending، Horror و غیره)، و معمولاً با استفاده از یک الگوریتم تولید می‌شود. صفحه اصلی هر کاربر، بسته به دستگاهی که از آن استفاده می‌کند، تقریباً از ۴۰ ردیف شامل حداکثر ۷۵ قلم تشکیل شده است. برای رتبه‌بندی سطری و ستونی از الگوریتم‌های مختلفی استفاده شده است. در ادامه به بررسی این الگوریتم‌های رتبه‌بندی می‌پردازیم.

شکل ۶. نمونه‌ای از اطلاعات محتوایی کاربران [7]

الگوریتم‌های رتبه‌بندی

مهم‌ترین رویکردهای مختلف رتبه‌بندی را می‌توان به‌صورت زیر بیان نمود:

- رتبه‌بندی ویدئویی شخصی(PVR): این روش یک الگوریتم چندمنظوره است که معمولاً کاتالوگ برنامه‌ها را با معیارهای خاصی پالایه می‌کند (به‌عنوان‌مثال برنامه‌های تلویزیونی خشونت آمیز، برنامه‌های تلویزیونی آمریکایی، عاشقانه و غیره)، همچنین از ویژگی‌های جانبی دیگری ازجمله ویژگی‌های شخصی کاربر و محبوبیت برنامه‌ها هم استفاده می‌کند.
- رتبه‌بندی ویدئویی n-برترین: مشابه PVR عمل می‌کند با این تفاوت که PVR فقط به راس رتبه‌بندی نگاه می‌کند ولی این روش کل کاتالوگ را بررسی می‌کند. با استفاده از معیارهایی که به راس رتبه‌بندی کاتالوگ نگاه می‌کنند و با روش مناسب بهینه‌سازی می‌شود.
- ترندهای پیشگو: این الگوریتم رویدادهای زمانی که بر اساس استنتاج نتفلیکس پیشگوهای قوی هستند را شناسایی می‌کند. این رویدادهای کوتاه مدت می‌توانند در حد چند دقیقه در چند روز باشند. این رویدادها / ترندها به‌طور معمول عبارت‌اند از:

۱. رویدادهایی که روند فصلی دارند و تکرار می‌شوند (به‌عنوان‌مثال روز ولنتاین منجر به افزایش میزان جذب فیلم‌های عاشقانه می‌شود)
۲. رویدادهای کوتاه مدت (به‌عنوان‌مثال ویروس کرونا یا سایر بلایا، که منجر به علاقه کوتاه مدت به مستندهای مربوط به آن‌ها می‌شود)

- رتبه‌بندی مشاهده پیوسته: این الگوریتم مواردی را که کاربر برای مشاهده انتخاب کرده ولی تا انتهای برنامه تماشا نکرده و مشاهده آن را به اتمام نرسانده است بررسی می‌کند، به‌طور معمول:

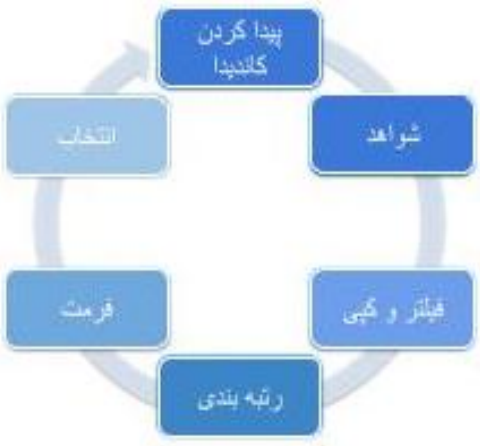
۱. محتوای اپیزودیک (به‌عنوان‌مثال مجموعه‌های نمایشی)
۲. محتوای غیر اپیزودیک که در بازه‌های گسسته قابل پیگیری هستند (به‌عنوان‌مثال فیلم‌هایی که نیمه تمام هستند، یا مجموعه‌های تلویزیونی که هر قسمت آن یک اپیزود مستقل است مانند آینه سیاه)

این الگوریتم احتمال ادامه مشاهده توسط کاربر (مشاهده پیوسته) را محاسبه می‌کند. این الگوریتم شامل سیگنال‌های آگاه از زمینه (به‌عنوان‌مثال زمان سپری شده از زمان مشاهده، نقطه رهاشدن مشاهده، دستگاه کاربر که برای تماشا استفاده شده و غیره) است.

در این الگوریتم از شبکه‌های RNN برای پیش بینی دنباله‌های حساس به زمان استفاده شده است. نتفلیکس می‌تواند از برنامه‌های مشاهده شده درگذشته یک کاربر خاص در کنار اطلاعات محتوایی استفاده کند و از این برای پیش‌بینی اینکه برنامه انتخابی بعدی این کاربر چه خواهد بود استفاده کند. به‌طور خاص، استفاده از زمان پیوسته همراه با محتوای زمان گسسته به‌عنوان ورودی بهترین عملکرد را دارد.

- رتبه‌بندی شباهت ویدئو-ویدئو (چون تماشا کردی) (BYW): این الگوریتم اساساً شبیه الگوریتم پالایش محتوایی است. بر اساس اقلامی مشاهده شده کاربر، الگوریتم سایر موارد مشابه را شناسایی می‌کند (با استفاده از ماتریس شباهت قلم- قلم) و مشابه‌ترین موارد را برمی‌گرداند. در میان الگوریتم‌های دیگر، این یکی شخصی نیست زیرا هیچ ویژگی جانبی دیگری استفاده نمی‌شود. بااین‌حال، شخصی‌سازی به این معناست که نمایش اقلام مشابه با برنامه خاص در فهرست مشاهده شده‌های کاربر، انتخابی مناسبی بر اساس سلیقه کاربر است.

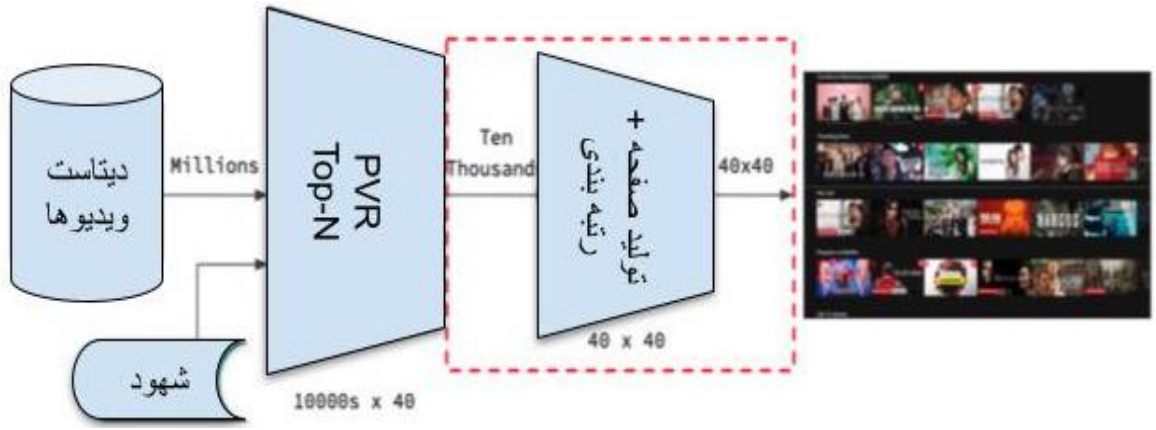
هر پنج الگوریتمی که در بالا بیان شد برای تولید سطرها در جدول پیشنهاد برنامه به کاربر از یک روال تکراری استفاده می‌کند. این روال در شکل ۷ مشاهده می‌شود. به‌عنوان‌مثال، اگر PVR سبک عاشقانه را پیشنهاد دهد، نامزدهایی را پیدا می‌کند که متناسب با این سبک باشند، و یک ردیف از برنامه‌های متناسب با این سبک (به‌عنوان‌مثال فیلم‌های عاشقانه که کاربر قبلاً تماشا کرده است) ارائه می‌دهد. این الگوریتم انتخاب شواهد در هر الگوریتم رتبه‌بندی دیگر که در بالا ذکرشده است نیز استفاده شده است. شکل ۸ تصویر گردش کار مدل نتفلیکس را نمایش می‌دهد.



شکل ۷. روال تکراری پیشنهاد برنامه به کاربر

ساخت صفحه کاربر

بعد از تولید سطرهای پیشنهادی برای اکانت یک کاربر خاص، که گاهی به ۱۰۰۰ سطر می‌رسد، مرحله بعد تولید پیشنهاد نهایی از میان سطرهای ساخته شده است. همانطور که در شکل ۹ مشاهده می‌شود ابتدا بر اساس سابقه کاربر یک فهرست n-بهترین از سبک‌های موردعلاقه کاربر بر اساس اکانت او استخراج می‌شود. این سبک‌ها که هر کدام یک سطر از جدول پیشنهادات را می‌سازند ممکن است بیش از ۱۰۰۰ سطر باشند که لازم است مختصرسازی شوند در غیراینصورت برای کاربر ایجاد ابهام و سردرگمی می‌کنند. برای این منظور برنامه‌های پیشنهاد شده باید رتبه‌بندی شوند تا درنهایت یک جدول محدود از پیشنهادها برای کاربر باقی بماند. الگوریتم‌های رتبه‌بندی مورداستفاده در ادامه معرفی شده‌اند.



شکل ۸. روال کامل توصیه‌گر نتفلیکس

شکل ۹. ساختار سیستم توصیه‌گر در نتفلیکس

- روش مبتنی بر الگو (در شروع)

نتفلیکس از یک روش مبتنی بر الگو برای حل این مشکل تولید صفحه استفاده کرده است. برای حل این مسئله نه تنها به دقت، بلکه به‌طور همزمان تنوع، قابلیت دسترسی و ثبات نیز باید مورد توجه واقع شود. ملاحظات دیگری شامل قابلیت‌های سخت افزاری (از چه دستگاهی استفاده می‌شود) و اینکه کدام ردیف‌ها/ستون‌ها در نگاه اول و هنگام پیمایش قابل مشاهده هستند، مورد توجه قرار می‌گیرند.

مدل نتفلیکس تلاش می‌کند آنچه را که کاربر می‌خواهد تماشا کند را به‌طور دقیق پیش‌بینی کند، اما این نکته را لحاظ می‌کند که وی ممکن است بخواهد فیلم‌هایی را که نیمه کاره رها کرده را انتخاب کند. در عین حال می‌خواهد با ارائه اقلامی تازه تنوع برنامه‌ها در کاتالوگ خود را برجسته کند. سرانجام فاکتور ثبات در پیشنهادات ضروری است تا زمانیکه کاربر برای مدتی با نتفلیکس تعامل داشته باشند و به روش‌های مورداستفاده در پیمایش صفحه و باقی ویژگی‌ها عادت کنند.

بر اساس تمام الزامات بیان‌شده فوق، می‌توان دریافت که چرا یک روش مبتنی بر الگو می‌تواند برای شروع کاملاً خوب کار کند. چرا که می‌توان معیارهای ثابت معینی را که باید مرتباً برآورده شوند به‌صورت قانون لحاظ کرد. با این وجود استفاده مدام از روش‌های مبتنی بر الگو ممکن است اکانت کاربران را به سمت بهینه‌های محلی هدایت کند بجای اینکه تجربه خوشایندی برای اعضا ایجاد نماید.

- رویکرد ردیف محور: رویکرد مبتنی بر ردیف برای امتیازدهی به هر ردیف و رتبه‌بندی آن‌ها بر اساس آن نمرات از رویکردهای توصیه شده یا یادگیری برای رتبه‌بندی موجود استفاده می‌کند. این روش می‌تواند نسبتاً سریع باشد اما فاقد تنوع است. ممکن است نتیجه نهایی برای یک کاربر صفحه‌ای پر از ردیف‌های مختلف باشد که به‌طورکلی با علاقه وی مطابقت دارد، اما ممکن است ازنظر شباهت بسیار نزدیک به هم بوده و فاقد تنوع لازم باشند.
- رویکرد مرحله‌ای: بهبود رویکرد افزایش ردیف استفاده از رویکرد مرحله‌ای است که در آن هر ردیف مانند روش فوق امتیازدهی می‌شود. با این تفاوت که ردیف‌ها به ترتیب از ابتدا انتخاب و جایگذاری می‌شوند و هر زمان که ردیف جدیدی انتخاب می‌شود، شباهت آن با ردیف‌های قبلی محاسبه می‌شوند تا رابطه آن با ردیف‌های قبلی و همچنین موارد پیدا شده در صفحه مشخص شود. این یک رویه ساده مرحله‌ای حریصانه است. این روش را می‌توان با استفاده از رویکرد K-سطر پیشرو بهبود بخشید. در این رویکرد برای انتخاب ردیف بعدی، محاسبه امتیازدهی برای k ردیف بعدی انجام می‌شود و سپس بهترین ردیف انتخاب می‌شود، این روش ساده وضعیت را بهبود می‌بخشد. بااین‌حال هیچ یک از این رویکردها بهینه سراسری را به‌دست نمی‌آورد.
- رویکرد یادگیری ماشین: رویکردی که نتفلیکس در یادگیری ماشینی استفاده می‌کند آموزش یک مدل بر اساس تاریخچه رفتار هر اکانت است. در این روش اطلاعاتی مثل اینکه کاربر چه برنامه‌هایی را برای دیدن انتخاب کرده و نحوه تعامل او با اقلام پیشنهادی به چه صورت بوده از تاریخچه اکانت استخراج و استفاده می‌شود.

ویژگی‌ها و راه‌های بسیار دیگری وجود دارد که می‌تواند یک ردیف خاص را در صفحه اصلی برای کاربر انتخاب کند. این می‌تواند به سادگی استفاده از همه متا داده‌های قلم (به‌عنوان یک بردار تعبیه) و تجمیع آن‌ها باشد. صرف نظر از اینکه از چه ویژگی‌هایی برای چینش اقلام در صفحه استفاده می‌شود، هدف اصلی ایجاد یک صفحه فرضی از انتخاب‌های نهایی برای کاربر است تا بدین وسیله با دریافت انتخاب‌های واقعی کاربر (بعد از اینکه قلمی را انتخاب و مشاهده کرد) به رویکرد الگوریتم امتیازدهی شود. امتیازدهی با استفاده از معیارهای مشابه Precision@k و Recall@k انجام می‌شود

۵-۲- سیستم توصیه‌گر یوتیوب

یوتیوب بزرگترین بستر جهانی برای ایجاد، انتشار و کشف محتوای ویدئویی است. یوتیوب دارای بیش از یک میلیارد کاربر فعال است که این آمار در حال رشد است. در ادامه به رویکرد یادگیری عمیق که اخیراً در سیستم توصیه‌گر ویدئویی یوتیوب استفاده شده است، می‌پردازیم. این مطالب در مقاله [۴۷] منتشرشده است.

چالش‌هایی که پیشنهاد فیلم در یوتیوب با آن مواجه است به قرار زیر است:

- مقیاس: بسیاری از الگوریتم‌های توصیه‌گرموجود در مقیاس‌های کوچک به خوبی کار می‌کنند، اما همان الگوریتم‌ها در مواجهه با مقیاس گسترده یوتیوب با مشکل برخورد می‌کنند.
- تازگی: یوتیوب دارای یک مجموعه بسیار پویا است و در هر ثانیه تعداد زیادی فیلم جدید بارگذاری می‌شود. لذا سیستم توصیه‌گر در فرایند تصمیم‌گیری این موارد جدید را لحاظ کند. ایجاد تعادل بین ویدئوهای تازه بارگذاری شده و قدیمی‌تر

اهمیت بسیار دارد.

- نویز: پیش‌بینی رفتار کاربر در یوتیوب با استفاده از تاریخچه رفتارهای او به دلیل کمبود و تنوع عوامل خارجی غیرقابل مشاهده بسیار دشوار است. ما به ندرت رضایت واقعی کاربر را به‌دست می‌آوریم و در عوض سیگنال بازخوردهای ضمنی او را مدل می‌کنیم. علاوه بر این، فراداده‌های مرتبط با محتوا بدون وجودشناسی کاملاً مشخص ساختار ضعیفی ارائه می‌دهند.

الگوریتم‌هایی که برای سیستم توصیه‌گردر یوتیوب انتخاب می‌شوند لازم است که نسبت به این ویژگی خاص در داده‌های آموزشی مقاومت بسیار داشته باشند.

ساختار کلی سیستم توصیه‌گر یوتیوب در شکل ۱۰ نشان داده شده است. این سیستم از دو شبکه عصبی تشکیل شده است. یکی برای تولید کاندیداها و دیگری برای رتبه‌بندی. شبکه تولید کاندیداها رویدادها را از تاریخچه فعالیت یوتیوب کاربر به‌عنوان ورودی گرفته و بر اساس این اطلاعات زیرمجموعه‌های کوچک (در حد صد ویدئو) را از پیکره بزرگ ویدئوها استخراج می‌کند. شبکه کاندیداهای تولیدشده شخصی‌سازی گسترده را فقط از طریق پالایش مشارکتی فراهم می‌کند. شباهت بین کاربران بر اساس ویژگی‌های درشت محاسبه می‌شود؛ ویژگی‌هایی مانند شناسه ویدئوهای مشاهده شده و عبارت‌های جستجو شده کاربر. ارائه فهرست چند «بهترین» پیشنهاد به ویژگی‌های ریزتری نیاز دارد تا اهمیت‌های نسبی بین کاندیداها با فراخوانی بالا شناسایی شود. شبکه رتبه‌بندی با اختصاص یک امتیاز به هر فیلم بر اساس یک تابع هدف مورد نظر، که با استفاده از مجموعه‌ای غنی از ویژگی‌ها برای توصیف فیلم و کاربر، این کار را انجام می‌دهد. ویدئوهای دارای بالاترین امتیاز، به ترتیب امتیاز آن‌ها به کاربر ارائه می‌شوند.

معماری مدل

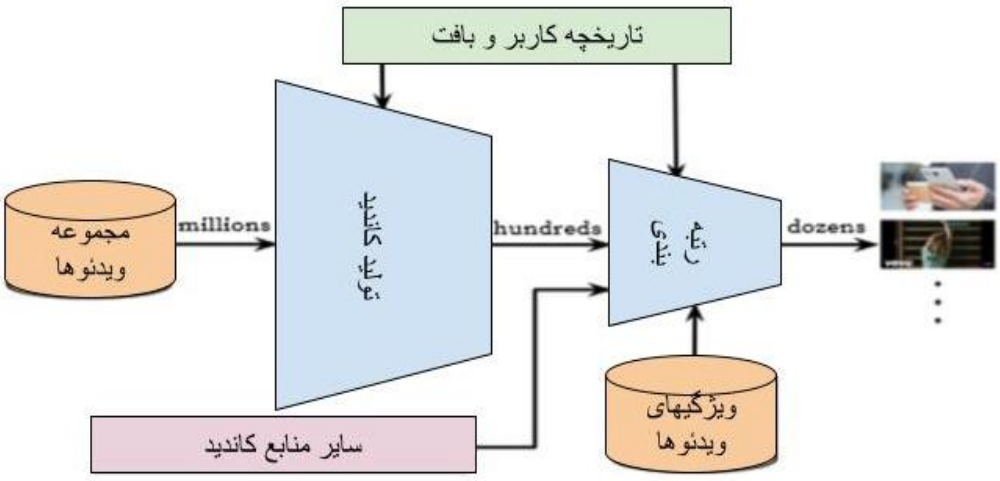
طبق شکل ۱۱ با استفاده از مدل زبانی کیسه کلمات[۴۷] یک بردار تعبیه با ابعاد بالا برای هر ویدئو یاد گرفته می‌شود و این بردارهای تعبیه به‌عنوان ورودی به یک شبکه عصبی تغذیه‌ای روبه جلو داده می‌شود. تاریخچه ویدئوهای تماشا شده کاربر با یک دنباله طول متغیر از شناسه ویدئوها نشان داده می‌شود. سپس این بردار به یک بردار تعبیه در فضای چگال نگاشته می‌شود. این شبکه به ورودی‌های متراکم با اندازه ثابت نیاز دارد لذا با استفاده از روش میانگین‌گیری بردارهای ورودی به یک بردار تبدیل می‌شوند. در لایه اول بردارهای تعبیه مشترک در ارتباط با پارامترهای دیگر آموزش داده می‌شوند. ویژگی‌ها به یک لایه اول متصل می‌شوند و به دنبال آن چندین لایه کاملاً متصل از نوع ReLU است.

در این مدل برای بررسی جدید یا قدیمی بودن ویدئوها یکی از ویژگی‌ها سن در نظر گرفته شده است و مقدار آن برابر با مدت زمانی است که از بارگذاری ویدئو بر روی یوتیوب می‌گذرد. شکل ۱۱ معماری شبکه را نشان می‌دهد.

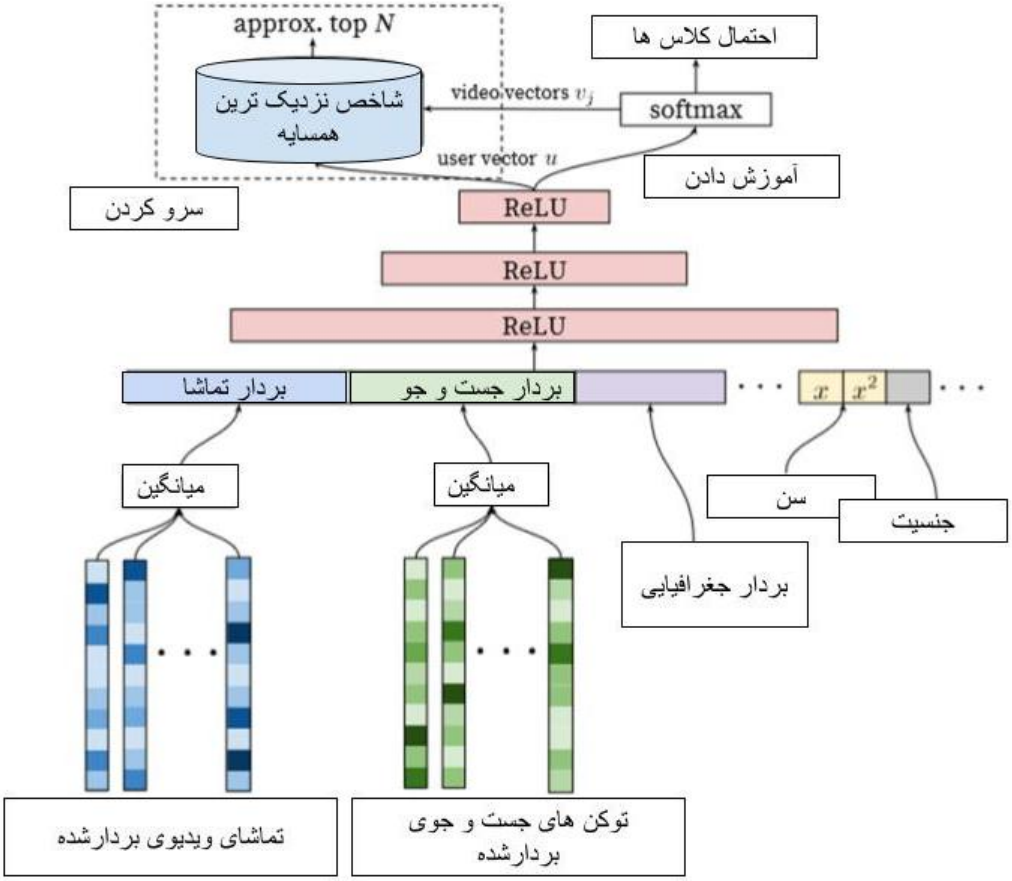
رتبه‌بندی

رتبه‌بندی در این روش بر اساس فاکتور اثر تدوین شده است. فاکتور اثر به معنای میزان اثرگذاری توصیه پیشنهاد شده به کاربر است که به میزان این پیشنهاد از طرف کاربر پذیرفته شده است. در مرحله رتبه‌بندی، به ویژگی‌های بیشتری در ارتباط با توصیف ویدئو و ارتباط کاربر با ویدئو دسترسی داریم، زیرا به دلیل اینکه یک مرحله کاندیداهای مناسب برای کاربر تولیدشده‌اند، در این مرحله تنها چند صد فیلم به جای میلیون‌ها قلم مورد ارزیابی قرار می‌گیرند.

.



شکل ۱۰. ساختار سیستم توصیه‌گر در یوتیوب



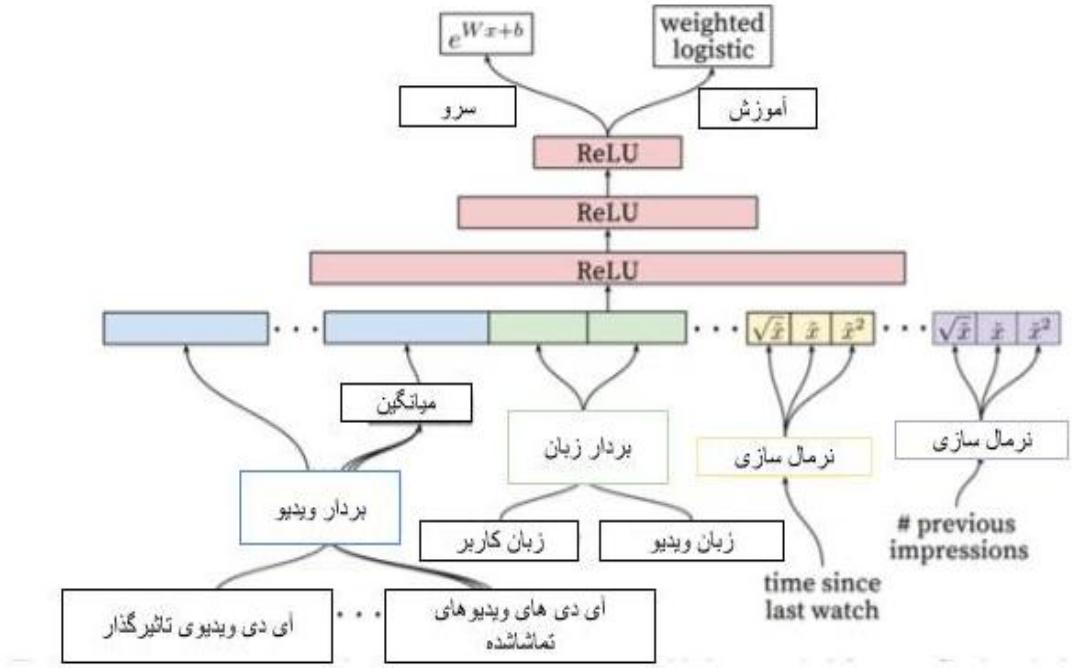
شکل ۱۱. معماری شبکه عمیق تولید کاندیداها درمدل توصیه‌گر یوتیوب [45]

برای رتبه‌بندی از یک شبکه عصبی عمیق با معماری مشابه شبکه استفاده شده برای تولید کاندیداها استفاده می‌کنیم تا با استفاده از رگرسیون لجستیک به هر نمایش فیلم نمره مستقلی اختصاص دهیم (شکل ۱۲). سپس فهرست فیلم‌ها بر اساس این امتیاز مرتب شده و به صفحه کاربری کاربر منتقل می‌شود. رتبه‌بندی با استفاده از روش نرخ کلیک اغلب باعث تبلیغ فیلم‌هایی می‌شود که ظاهر فریبنده‌ای دارد و اغلب کاربران به‌طور کامل آن را تماشا نمی‌کنند ("clickbait") در حالی که زمان صرف شده برای تماشای یک ویدئو تعاملات بهتری از کاربر را ضبط می‌کند .

مهندسی ویژگی

از صدها ویژگی در مدل‌های رتبه‌بندی استفاده شده است. علی رغم ویژگی یادگیری عمیق برای حذف مهندسی دستی ویژگی‌ها، ماهیت داده‌های خام در این مسئله به‌گونه‌ای است که به راحتی امکان ورود مستقیم به شبکه‌های عصبی پیش‌رو را فراهم نمی‌کند.

مشاهده می‌کنیم که مهم‌ترین سیگنال‌ها آنهایی هستند که تعامل قبلی کاربر با یک قلم و سایر موارد مشابه را توصیف می‌کنند و با تجربه دیگران در رتبه‌بندی تبلیغات مطابقت دارند. به‌عنوان نمونه، سابقه گذشته کاربر با کانالی که ویدئو را در حال امتیازگذاری بارگذاری کرده در نظر بگیرید - کاربر چند فیلم از این کانال تماشا کرده است؟ آخرین باری که کاربر ویدئویی با این موضوع تماشا کرده چه زمانی بوده است؟ این ویژگی‌های مداوم که اقدامات گذشته کاربر را در موارد مرتبط توصیف می‌کند، بسیار مهم هستند زیرا به خوبی قابل تعمیم به موارد متفاوت هستند. همچنین انتشار اطلاعات و روش انتخاب کاندیداهای تولیدشده برای رتبه‌بندی بسیار مهم است که این انتقال اطلاعات می‌تواند در بردار ویژگی‌ها لحاظ شود. به‌عنوان‌مثال اینکه یک ویدئوی کاندیدا توسط چه منابع اطلاعاتی معرفی شده است و یا اینکه با چه امتیازاتی این ویدئو کاندیدا شده است. این اطلاعات از مرحله تولید کاندید به مرحله رتبه‌بندی قابل انتقال است.



شکل ۱۲. معماری شبکه رتبه‌بندی در یوتیوب [45]

بردارهای تعبیه ویژگی‌های طبقه‌بندی شده

مشابه کاندیداهای پیشنهادی، از بردار تعبیه برای ترسیم ویژگی‌های طبقه‌ای تنک هم استفاده می‌شود تا در یک نمایش چگال به‌عنوان ورودی برای شبکه‌های عصبی مناسب باشند. در [۴۸] یک پیاده‌سازی از مدل پیشنهادی یوتیوب ارائه شده است.

۵-۳- سیستم توصیه‌گر آمازون

در مقاله‌ای تحت عنوان «راز سیستم توصیه‌گر آمازون برای فروش آنلاین بیشتر» به بررسی روش‌هایی که در سیستم توصیه‌گر آمازون استفاده می‌شود پرداخته شده است [۴۹].

«با توجه به موفقیت آمازون، می‌توان ادعا کرد که سیستم توصیه‌گر به‌خوبی کار می‌کند. این شرکت از رشد فروش ۲۹ درصدی در ۱۲ ماهه دوم مالی خود به ۱۲,۸۳ میلیارد دلار خبر داد، در حالی که در مدت مشابه سال گذشته ۹,۹ میلیارد دلار بود. بدون شک بسیاری از این رشد مربوط به روش‌های پیشنهادی است که آمازون در هر بخش از فرآیند خرید اعمالی کرده است ... »

آن‌ها میلیاردها داده استفاده می‌کنند تا بسیاری موارد مختلف را آزمایش کنند تا به‌سرعت بفهمند چه عواملی موثرند. طی سال‌های گذشته دو نحوه ارسال پیشنهاد در آمازون را شاهد هستیم: از طریق ایمیل و از طریق سایت. به‌علاوه مشخص است که توصیه‌های آمازون از طریق ایمیل بهتر از توصیه‌های درون سایت عمل می‌کند.

آمازون همچنین پیشنهادهایی را از طریق ایمیل به کاربران ارائه می‌دهد. در واقع، نرخ تبدیل و کارایی اینگونه ایمیل‌ها «بسیار زیاد» است، به‌طور قابل‌توجهی مؤثرتر از توصیه‌های روی سایت است.

انواع روش‌های پیشنهادی

۱) توصیه‌های موجود در سایت

در ادامه برخی از روش‌های توصیه کالا بر روی سایت آمازون بیان‌شده است. [۴۹]

- «برای شما توصیه می‌شود، توماس»- با کلیک بر روی پیوند «توصیه‌های شما» کاربران به صفحه‌ای پر از محصولات پیشنهادی منتقل می‌شوند که فقط برای آن‌ها توصیه می‌شوند. آمازون طیف وسیعی از محصولات را از دسته‌های مختلف که مرور کرده‌اند، یا محصولاتی که احتمالاً روی آن‌ها کلیک کرده‌اند با جزئیات بیشتری درباره آن‌ها ارائه می‌دهد.
- «معمولا با هم خریده شده‌اند»- این پیشنهادات یک هدف اصلی دارد: افزایش متوسط ارزش سفارش. توصیه‌های «اغلب با هم خریداری شده» با ارائه پیشنهاد محصول بر اساس موارد موجود در سبد خریدار، آن‌ها را ترغیب به خرید زیر کالاهای دیگری می‌کنند.
- «موارد بازدید شده اخیر و پیشنهاد ویژه شما» - آمازون به محصولی که آن را جستجو کرده‌اند نگاه می‌کند و محصولات بسیار مشابه با اشکال، اندازه‌ها و مارک‌های مختلف را برای

کمک به کاربر در یافتن قلمی بسیار مشابه با محصولی که قبلاً به آن علاقه نشان داده‌اید، توصیه می‌کند. آن‌ها مارک‌ها، رنگ‌ها، اشکال و اندازه‌های مختلف را پرتاب می‌کنند با این امید که محصولی را در مقابل شما قرار دهند که در برابر آن مقاومت نکنید.

- تاریخچه جستجوی شما - اگر کاربران قبلاً به محصولی نگاه کرده‌اند، به این معنی است که آن‌ها کمی به آن علاقه‌مند هستید و آمازون آن را می‌داند، بنابراین اگر می‌خواهید سریع برگردید و چیزی را خریداری کنید که قبلاً به آن علاقه داشتید، سابقه مرور شما را به شما نشان می‌دهد.

۲) توصیه‌های خارج از سایت

در اینجا نمونه‌هایی از توصیه‌هایی است که آمازون برای ارسال آن‌ها از طریق ایمیل انجام داده است:

- «بهترین فروش مدل Canon در این هفته»- اولین ایمیل طیف وسیعی از مدل‌های پرفروش از دسته کالایی است که بازدید کرده‌اند. از آنجا که فقط مدل‌های Canon در این ایمیل به نمایش گذاشته شده‌اند، می‌توان مطمئن بود که کاربر مربوطه یکبار در حال جستجوی آن مارک دوربین در سایت بوده است و یا حتی دوربین Canon را به سبد خرید خود اضافه کند.
- «قاب دوربین بخرید» - این ایمیل پیشنهادی حاوی مواردی است که اغلب با هم خریداری می‌شوند. با این هدف که کاربر بخواهد یک دوربین خریداری کند و لوازم جانبی آن را نیز خریداری کند. بدین صورت ارزش سفارش و میزان درآمد حاصل از هر مشتری افزایش می‌یابد.

ساختار مدل

در اصل سیستم توصیه‌گر غول خرده فروشی بر اساس تعدادی از عناصر ساده بنا شده است: [۵۰]

- آنچه کاربر درگذشته خریده است،
- چه کالاهایی را در سبد خرید مجازی خود دارد،
- چه کالاهایی را رتبه‌بندی کرده و دوست داشته است،
- سایر مشتریان چه مواردی را مشاهده و خریداری کرده‌اند.

آمازون این روش خود را «پالایش اشتراکی قلم- قلم» می‌نامد و از این الگوریتم برای شخصی‌سازی شدید تجربه جستجو برای اطمینان از بازگشت مجدد مشتری برای خرید استفاده می‌کند.

مقاله [۵۱] در سال ۲۰۰۳ در ارتباط با سیستم توصیه‌گر آمازون منتشرشده است. بر اساس این مقاله برای حل مسئله توصیه سه روش رایج وجود دارد:

- پالایش مشترک مشارکتی،
- مدل‌های خوشه‌ای،
- روش‌های مبتنی بر جستجو.

الگوریتم مدل پیشنهادی آمازون یک پالایش مشارکتی قلم- قلم است. بر خلاف پالایش مشترک مشارکتی، محاسبه آنلاین الگوریتم به‌طور مستقل از تعداد مشتری و تعداد اقلام موجود در فهرست محصولات مورد قیاس قرار می‌گیرد. الگوریتم آمازون پیشنهاداتی را در زمان واقعی و در مقیاس مجموعه‌های گسترده داده تولید می‌کند و توصیه‌هایی با کیفیت بالا پیشنهاد می‌کند.

بیشتر الگوریتم‌های پیشنهادی با یافتن مجموعه‌ای از مشتریانی شروع می‌شوند که اقلام خریداری شده آن‌ها و امتیازدهی و رتبه‌بندی اقلام در فهرست خریدشان با کاربر مورد نظر ما مطابقت داشته باشد. دو نسخه متداول از این الگوریتم‌ها پالایش مشارکتی و مدل‌های خوشه‌ای هستند. الگوریتم‌های دیگر - ازجمله روش‌های مبتنی بر جستجو و پالایش مشارکتی قلم- قلم‌که در مدل آمازون استفاده می‌شود - بر یافتن اقلام مشابه بجای مشتریان مشابه تمرکز می‌کنند. الگوریتم تلاش می‌کند برای هر یک از اقلام خریداری شده و رتبه‌بندی شده کاربر، موارد مشابه را پیدا کند. سپس موارد مشابه را جمع کرده و به او پیشنهاد می‌کند.

۵-۴- سیستم توصیه‌گر فیسبوک

شبکه‌های اجتماعی آنلاین بخشی از زندگی روزمره ما شده است و یکی از سایت‌های محبوب آنلاین شبکه‌های اجتماعی در اینترنت فیسبوک است، جایی که کاربران با دوستان خود ارتباط برقرار می‌کنند، به گروه‌ها می‌پیوندند، گروه‌ها را ایجاد می‌کنند، گروه‌ها را بازی می‌کنند، و در سراسر جهان دوست می‌یابند. همچنین، تعداد زیادی از گروه‌ها برای علل و عقاید مختلف ایجاد می‌شوند. با وجود این گستردگی در گروه‌ها کاربران را برای انتخاب یک گروه مناسب برای عضویت دچار مشکل می‌کند. برای حل این مشکل، باتاراژاو و همکاران [۵۲] سیستم توصیه‌گر گروهی[1] با استفاده از ترکیبی از روش خوشه‌بندی سلسله مراتبی و درخت تصمیم را معرفی کرده‌اند. با استفاده از این ابزار گروه‌های شبکه اجتماعی فیسبوک را می‌توان بر اساس مشخصات اعضا شناسایی کرد[۵۲].

[1] GRS

شکل ۱۳ معماری ساده‌ای از سیستم توصیه‌گر گروهی (GRS) فیسبوک را نشان می‌دهد. این مدل شامل سه جز اصلی زیر است:

- استخراج ویژگی‌های نمایه کاربر
- موتور طبقه‌بندی
- توصیه نهایی

مجموعه دادگان- برای این پژوهش با استفاده از یک api به حساب‌های دانشجویی که در آن تنظیمات حریم خصوصی برای دسترسی به شبکه دوستان آن‌ها مجاز بود (تنظیم پیش فرض)، دسترسی پیدا کرده‌اند. برای این منظور از شبکه اجتماعی دانشگاه شمال تگزاس[1] در فیسبوک استفاده شده است. در طی این جستجو، امکان دسترسی به حساب‌های ۱۵۸۰ کاربر وجود داشته است. از این حساب‌ها اطلاعات زیر جمع‌آوری‌شده:

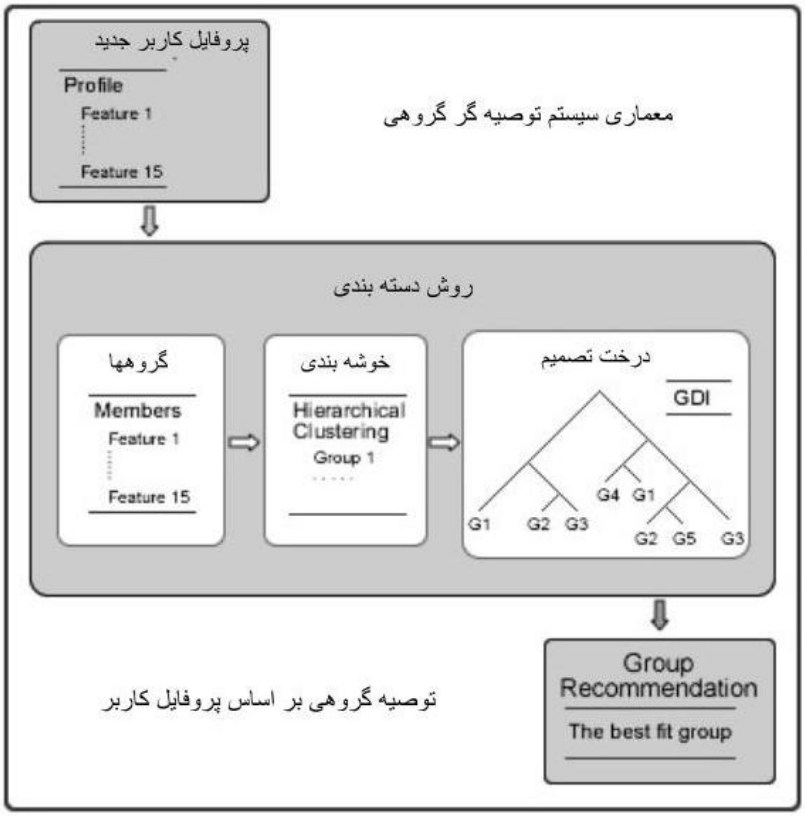
- اطلاعات نمایه کاربران،
- ارتباطات دوستانه و گروه‌هایی که عضو به آن‌ها بودند.

اولین گام در سیستم توصیه‌گر گروهی تجزیه و تحلیل و شناسایی ویژگی‌هایی است که تمایلات کاربر را ازنظر علاقه، ارتباط اجتماعی، اطلاعات اساسی مانند سن، جنس، تعداد دیوارها، تعداد یادداشت‌ها و بسیاری از موارد مشابه تفکیک می‌کند.

یکی از تکنیک‌های متداول برای یافتن شباهت بین گره‌ها در فضای چند بعدی، تجزیه و تحلیل خوشه‌بندی سلسله‌مراتبی است. برای استنتاج تشابه بین اعضا در این مدل، از فاصله اقلیدسی استفاده شده است.

ضریب خوشه‌بندی

هر گروه دارای یک ویژگی منحصر به فرد است، که آن را از دیگران متمایز می‌کند، بااین‌حال برخی از اعضای یک گروه ممکن است مشخصات مختلفی داشته باشند. با افزایش این اختلافات تا حدی این اعضا به‌عنوان «نویز» اجتناب‌ناپذیر برای خوشه‌بندی ظاهر می‌شوند. برای شناسایی و کاهش این نویزها یک گروه را توسط اعضای اصلی آن مشخص می‌کنند. منظور از اعضای اصلی آنهایی هستند که داخلی‌ترین قسمت گروه را ایجاد می‌کنند. برای این منظور مفهوم ضریب خوشه مطرح می‌شود.



شکل ۱۳. معماری سیستم توصیه‌گر گروهی فیسبوک

[1] UNT

این ضرایب بر اساس فرمول زیر بیان می‌شود:

$$C = \frac{N_{R_i}}{R_i},$$

در این فرمول N_k تعداد نرمال شده اعضا در فاصله k از مرکز است و R_i فاصله اقلیدسی نرمال شده از مرکز عضو i است و بر اساس فرمول زیر محاسبه می‌شود:

$$R_i = \frac{r_i}{\max_j(r_j)},$$

r_i فاصله از مرکز عضو i است.

درخت تصمیم

ماهیت سیستم توصیه‌گر گروهی (GRS) مبتنی بر طبقه‌بندی است. GRS بر اساس ویژگی‌های نمایه کاربر، مناسب‌ترین گروه‌ها را برای کاربر پیدا می‌کند. یک راه‌حل متداول برای مسئله طبقه‌بندی، الگوریتم درخت تصمیم مبتنی بر تقسیم‌بندی بازگشتی باینری است. چندین روش تقسیم در این الگوریتم وجود دارد: جینی، توئینگ و انحراف . در طی آزمایشات برای یافتن نتیجه بهتر، هر یک از قوانین تقسیم در GRS ادغام شده است. بااین‌حال هیچ پیشرفت قابل‌توجهی در دقت مشاهده نشد، به این معنی که درخت نهایی به قاعده تقسیم برای ساخت درخت بستگی ندارد. هدف اصلی این الگوریتم‌های تقسیم، یافتن بهترین تقسیم داده‌ها با حداکثر همگنی در هر طرف است. در هر تکرار بازگشتی داده‌ها تصفیه می‌شوند تا زمانی که الگوریتم به گره‌های نهایی(کلاس‌ها) برسد.

پیاده‌سازی سیستم توصیه‌گر فیسبوک

یک نمونه سیستم توصیه‌گر پیشرفته مبتنی بر یادگیری عمیق با استفاده از سیستم منبع باز PyTorch و Caffe۲ پیاده‌سازی شده است. DLRM[۵۱] با ترکیب اصول مربوط به هر دو روش پالایش مشارکتی و رویکردهای مبتنی بر تجزیه و تحلیل، مدل پیشرفته‌ای ارائه و نتایج پیشرفته‌ای را ارائه دهد. DLRM[۵۱] مدلی است که در گروه تحقیقاتی فیسبوک طراحی‌شده است.

۵-۵- سیستم توصیه‌گر اینستاگرام

بیش از نیمی از جامعه اینستاگرام هر ماه برای کشف عکس‌ها، فیلم‌ها و داستان‌های جدید مرتبط با علایق خود از ابزار جستجو در اینستاگرام استفاده می‌کنند. پیشنهاد مرتبط‌ترین مطالب از بین میلیاردها گزینه در زمان واقعی چالش‌های یادگیری ماشین متعددی به همراه دارد که برای حل آن‌ها به راه‌حل‌های جدید مهندسی نیاز است [۵۲].

IGQL - زبان اختصاصی توصیه‌گر اینستاگرام

پژوهشگران در تیم تحقیقاتی فیسبوک مجموعه‌ای از زبان‌های جستجوی سفارشی شامل تکنیک‌های سبک‌سازی مدل و ابزارهایی که امکان آزمایش با سرعت بالا را ایجاد می‌کنند برای جستجوی بهتر ایجاد کرده‌اند تا با این چالش‌ها مقابله کنند. این سیستم‌ها ضمن افزایش کارایی توسعه‌دهنده، از مقیاس‌های بزرگ جستجو در اینستا پشتیبانی می‌کنند. درمجموع، این راه‌حل‌ها را می‌توان به‌عنوان یک سیستم هوش‌مصنوعی مبتنی بر یک قیف رتبه‌بندی ۳ بخشی بسیار کارآمد توصیف کرد که ۶۵ میلیارد ویژگی را استخراج می‌کند و هر ثانیه ۹۰ میلیون پیش بینی مدل را انجام می‌دهد. زبان طراحی‌شده IGQL نام دارد [۵۴].

ضرورت ساخت یک موتور توصیه‌گر در شبکه‌ای مثل اینستاگرام که روزانه حجم وسیعی فایل شامل تعداد زیادی عکس و فیلم در آن بارگذاری می‌شود، تولید ابزارهای اساسی است که سه نیاز مهم زیر در این فضا را برآورده کنند:

- توانایی انجام آزمایش‌های سریع در مقیاس وسیع اینستاگرام را داشته باشد،
- نیاز است که امکان دستیابی به سیگنال‌های اطلاعاتی قوی‌تری در مورد کاربران را ایجاد کند،
- ازنظر محاسباتی برای اطمینان از کیفیت بالا و پیشنهاد توصیه‌های جدید روشی کارآمد باشد.

ساخت الگوریتم‌ها و تکنیک‌های توصیه بهینه یک زمینه تحقیقاتی است و روند انتخاب سیستم مناسب می‌تواند بسته به هدف بسیار متفاوت باشد. به‌عنوان‌مثال، در حالی که یک الگوریتم ممکن است منافع طولانی مدت را به‌طور موثر در نظر بگیرد، دیگری ممکن است در پیشنهاد توصیه‌ها بر اساس محتواهای اخیر عملکرد بهتری داشته باشد. تیم مهندسی اینستاگرام به روشی نیاز دارد که هم ایده‌های جدید را به‌طور کارآمد امتحان کند و هم ایده‌های امیدوار کننده را به راحتی و بدون نگرانی بیش از حد درباره منابع محاسباتی مانند پردازنده و حافظه موردنیاز به کار گیرد. IGQL یک فرازبان خاص دامنه سفارشی است که سطح انتزاعی مناسب را فراهم می‌کند و همه الگوریتم‌ها را در یک مکان جمع می‌کند.

بردار تعبیه حساب کاربری برای رتبه‌بندی شخصی

کاربران اینستاگرام معمولا میلیاردها قطعه فیلم با کیفیت بالا را که دارای مجوز واجد شرایط برای کاوش هستند، در اینستاگرام به

اشتراک می‌گذارند. یکی از چالش‌های جستجو در این فضا ارائه یک طبقه‌بندی قابل تغییر به‌صورت یک کاتالوگ واضح و در حال تحول برای طیف گسترده‌ای از سلایق مختلف است که از موضوعات مربوط به خوشنویسی عربی تا مدل قطارها تا انواع جلبک‌ها قابل تغییر هستند. درنتیجه مدل‌های مبتنی بر محتوا در جستجو در فضایی با چنین تنوعی که ناشی از ماهیت مبتنی بر علاقه کاربر است، با مشکل مواجه می‌شود.

ig2vec - بردار تعبیه کاربران

از آنجا که اینستاگرام تعداد زیادی حساب کاربری متمرکز بر اساس موضوعات خاص دارد، یک خط لوله بازیابی ایجاد شده است که به‌جای تمرکز رسانه محور، کاربر محور بوده و با اطلاعات مربوط به سطح حساب کار می‌کند. با ساختن بردارهای تعبیه حساب‌ها، شناسایی اینکه کدام حساب‌ها ازنظر موضوعی مشابه یکدیگر هستند، با کارایی بیشتری انجام می‌گبرد. بردار تعبیه حساب‌های کاربری با استفاده از ig2vec، یک مدل بردار تعبیه مشابه word2vec ساخته می‌شوند. به‌طور معمول، چارچوب بردار تعبیه word2vec نمایشی از یک کلمه را بر اساس محتوایی که (در جملات مجموعه آموزش) در آن‌ها ظاهر شده می‌آموزد. ig2vec شناسه‌های حساب کاربرانی که کاربری خاص با آن‌ها ارتباط برقرار می‌کند به‌عنوان دنباله‌ای از کلمات در یک جمله تلقی می‌کند.

۵-۶- سیستم توصیه‌گر تیک‌تاک

تیک‌تاک یک پلتفرم رسانه اجتماعی است که به کاربران اجازه می‌دهد که ویدئوهای کوتاه را با سایر کاربران به اشتراک گذارند. این پلتفرم قادر است بر اساس سلایق و علایق کاربران به ارائه پیشنهادات متعددی به کاربران می‌پردازند. سیستم توصیه‌گر بلادرنگ به کار گرفته شده در تیک‌تاک Monolith است. این سیستم در واقع یک مدل یادگیری عمیق است که بر اساس اعمال قبلی کاربر به پیش‌بینی تمایلات کاربر می‌پردازد. این مدل از یک جدول بازنمایی بدون برخورد جهت ذخیره‌سازی و بروزرسانی بازنمایی کاربران و ویدئوها استفاده می‌کند. در این جدول از تابع هش و پالایه بلوم استفاده و تضمین می‌کند که هر کاربر و هر ویدئو تنها یکبار بازنمایی و ذخیره گردد. به‌علاوه، برای بروزرسانی پارامترهای مدل بر اساس بازخوردهای دریافتی از کاربران از طریق زمان تماشای ویدئو، علاقه‌مندی و کامنت‌های کاربران از آموزش برخط استفاده می‌کند[۸۷].

معماری سیستم توصیه‌گر تیک‌تاک

معماری مدل پیشنهادی از ۳ جز اصلی تشکیل شده است: (۱) سیستم خدمت‌رسانی آنلاین شامل متعادل‌کننده بار، لایه کش و لایه مدل است. متعادل‌کننده بار درخواست‌ها را بر روی لایه کش که وظیفه ذخیره و نگهداری بازنمایی کاربران و ویدئوها و پارامترهای مدل را دارد، توزیع می‌کند. لایه کش درخواست‌ها را به سمت لایه مدل که وظیفه امتیازدهی به ویدئوهای کاندید و رتبه‌بندی آن‌ها بر اساس امتیازات محاسبه شده دارد، هدایت می‌کند. (۲) سیستم آموزش برخط، مسئولیت بروزرسانی مدل بر اساس بازخوردهای دریافتی از کاربران را دارد و شامل جمع‌آوری دادگان، پردازش دادگان و تنظیم پارامترها می‌باشد. در بخش جمع‌آوری دادگان، بازخورد کاربران جمع آوری و به بخش پردازش دادگان ارسال می‌شود. در بخش پردازش دادگان، به پیش‌پردازش دادگان و محاسبه گرادیان برای تنظیم پارامترهای مدل می‌پردازد. گرادیان‌های محاسبه شده به بخش تنظیم دادگان جهت بروزرسانی مدل سپرده می‌شود. (۳) سیستم آموزش آفلاین، مسئولیت آموزش مدل از پایه یا تنظیم بهینه مدل به‌صورت دوره‌ای را بر عهده دارد. در این سیستم انبار دادگان بازخورد کاربران، تاریخچه استفاده و ویژگی ویدئوها را در خود ذخیره می‌کند. خط لوله به کار گرفته شده در این سیستم، به استخراج دادگان و ویژگی‌ها می‌پردازد. دادگان استخراج‌شده جهت آموزش مدل خوشه‌بندی توزیع‌شده، مورداستفاده قرار می‌گیرند و خروجی حاصل به سیستم خدمت‌رسانی آنلاین ارسال می‌شود.

بررسی کیفیت سیستم توصیه‌گر تیک‌تاک

در پژوهش ارائه شده توسط Boeker و Urman [۸۸] پیشنهادات ارائه شده توسط سیستم توصیه‌گر ارائه شده در تیک‌تاک بر اساس علایق کاربر شخصی‌سازی می‌شود و شامل مراحل زیر است:

- اعتبارسنجی حساب کاربری

در این مرحله، ۱۵۹ کاربر واقعی از ۲۱ کشور مختلف در جهان، جهت اعتبارسنجی حساب کاربری تیک‌تاک دعوت شدند. از شرکت‌کنندگان در این بررسی کاربرد و میزان رضایت از سرویس‌های تیم‌تاک مورد سوال قرار گرفت. به‌علاوه، از شرکت کنندگان خواسته شد که در حساب‌های کاربری ایجاد شده توسط پژوهشگران این پژوهش به محتواهای پیشنهاد شده توسط تیک‌تاک امتیاز دهند

- تحلیل داده

در این مرحله به تحلیل دادگان جمع‌آوری‌شده از ۱۵۹ حساب

کاربری تیک‌تاک، که در بخش قبل توضیح داده شد، پرداخته می‌شود. به این ترتیب که ابتدا به منظور انجام تحلیل توصیفی از آمارهای توصیفی جهت خلاصه‌سازی دادگان جهت تعیین میزان رضایت کاربران از سرویس تیک‌تاک استفاده می‌شود. به‌علاوه، از رویکردهای مصورسازی جهت بیان الگوهای رفتاری کاربران و موارد مورد توجه کاربران در بازه‌های زمانی مختلف بر اساس رخداد اتفاقات مهم در جهان استفاده می‌شود. از سوی دیگر، به تحلیل استنباطی دادگان از طریق بررسی پاسخ سوالات موردبررسی در این پژوهش می‌پردازد. به‌علاوه، تحلیل احساسات، هیجانات و عقاید موجود نسبت به محتواهای پیشنهادی در حساب کاربری شرکت‌کنندگان نیز انجام می‌شود. از سوی دیگر، تحلیل فاکتورهای شخصی‌سازی شده در ارائه پیشنهاد و محتوای پیشنهاد شده نیز انجام می‌شود.

۵-۷- سایر سیستم‌های تجاری موجود

در ارتباط با سیستم‌های توصیه‌گر ابزارهای زیادی طراحی‌شده است که به ارائه خدمات می‌پردازند. از این ابزارهای تجاری به دلیل ماهیت اطلاعات زیادی منتشر نشده است. در مورد مدل و ساختار پیاده‌سازی اطلاعاتی در دست نیست. در ادامه به بیان معرفی ارائه شده از این ابزارها پرداخته شده است.

واندرکیوب [۵۵] - پلتفرم واندرکیوب راه‌حل ایده‌آل برای مدیریت یکپارچه پخش تلویزیونی است. کل مراحل پخش همگانی، از زمانبندی برنامه‌ریزی برنامه تا پخش، را گام به گام پوشش می‌دهد. استراتژی‌های برنامه‌ریزی برنامه، ازنظر تولید روزانه، در قلب برنامه‌ریزی تلویزیون هستند. واندرکیوب به شما امکان می‌دهد فقط با چند حرکت سریع و آسان از طریق نرم‌افزارهای یکپارچه مختلف از برنامه‌ریزی سالانه به پخش روزانه بروید. تدوین فهرست پخش و برنامه‌ریزی برنامه به لطف اتوماسیون پردازش‌ها و دسترسی به بایگانی‌های کاملا دیجیتالی به روندی روان و فوری تبدیل شده‌اند. هر برنامه نرم‌افزاری با مجموعه‌ای از توابع همراه است که به‌طور خاص قابلیت اطمینان، انعطاف‌پذیری و دوام سیستم در هر شرایطی را تضمین می‌کند. قابلیت‌های زیر برای ابزار زمانبندی در این پلتفرم قرار دارد:

- زمانبندی برنامه‌ها
- ایجاد، مشاهده و ویرایش برنامه‌ریزی
- مدیریت اطلاعات برنامه‌ها
- چاپ زمانبندی برنامه‌ها
- ارسال نامه الکترونیکی برنامه زمانبندی به فهرست پستی قابل تنظیم
- انتشار خودکار زمانبندی برنامه‌ها در وب
- مدیریت وقفه‌های تبلیغاتی

بر اساس مقاله [۵۶] فهرستی از برترین سیستم‌های تجاری ارائه شده است. در ادامه برخی از این موارد ذکرشده است.

- Optimizely [۵۷] - یکی از پیشروترین پلتفرم‌ها در حوزه توصیه‌گرها در جهان است، برای تیم‌های بازاریابی و تیم محصول امکان آزمایش، یادگیری و استقرار تجربیات دیجیتال را فراهم می‌کند.
- einstain [۵۸] - بنیانگذار Salesforce.com مارک بنیوف سهم خود را در ایجاد انقلابی در صنعت نرم‌افزار با ارائه «اولین هوش مصنوعی جامع جهانی برای CRM» انجام داده است. برای افرادی که از قبل مشتری Salesforce هستند، توصیه‌های انیشتین می‌تواند به راحتی در سیستم توصیه‌گر ادغام شود.
- Clarifai [۵۷] - مت زیلر (Matt Zeiler) بنیانگذار و مدیر عامل شرکت Clarifai یک متخصص هوش مصنوعی با دکترای تخصصی در یادگیری ماشین از NYU است. شرکت Clarifai از زمان کسب پنج مقام برتر در حوزه طبقه‌بندی تصاویر در مسابقات ImageNet ۲۰۱۳، قابلیت‌های برتر سیستم خود در مسئله تشخیص تصویر را ثابت کرده است. از دیگر محصولات موفق این شرکت سیستم توصیه‌گر است که به شرکت‌ها در حوزه تجارت الکترونیکی قابلیت ارائه پیشنهاد کالاهای مشابه به‌عنوان محصولات پیشنهادی را می‌دهد.
- Jinni [۶۰] - این ابزار به‌صورت رایگان در دسترس است و این امکان را فراهم می‌آورد که یک فیلم را انتخاب کنید و فیلم‌های مشابه با آن فیلم را ببینید. همچنین با استفاده از این وبسایت می‌توانید در بین فیلم‌ها بر اساس حالت، داستان، سبک، سال ساخت یا مخاطب (مثلاً مناسب برای کودکان) جستجو کنید و فیلم‌های جدیدی پیدا کنید. در صورتی که در این سایت، حساب بسازید، با لایک یا دیسلایک کردن فیلم‌ها می‌توانید پیشنهادات آینده را برای خود شخصی‌سازی کنید. این ابزار شامل فیلم‌های سینمایی، فیلم کوتاه و سریال‌های تلویزیونی است. همچنین امکان اتصال جینی به نتفلیکس برای اضافه کردن راحت‌تر فیلم‌ها به فهرست تماشا نیز وجود دارد.
- Rotten Tomatoes [۶۱] - این سایت، یک ابزار پیشرفته امتیازدهی به فیلم‌ها است. در کنار امتیازدهی و رده‌بندی فیلم‌ها، در این سایت می‌توانید بهترین فیلم‌های موجود برای کرایه کردن یا خرید را بررسی کنید، اخبار و نقدهای نوشته

شده در مورد فیلم‌ها را مطالعه کنید، تریلر فیلم‌ها را مشاهده کنید و حتی در صفحه فیلم‌ها، نظر بگذارید تا کاربران دیگر ازنظر شما استفاده کنند. هدف این سایت این نیست که به شما فیلم‌های مشابه یک فیلم را پیشنهاد دهد، بلکه هدفش این است که به شما اطلاعات کافی در مورد فیلم‌های مختلف بدهد تا بتوانید آگاهانه برای انتخاب یک فیلم را برای تماشا انتخاب کنید.

- Movie Lens [۶۲] - این سایت یک سرویس پشنهاددهنده فیلم است که توسط گروه تحقیقاتی لنز در دانشگاه مینسودا تهیه شده است. با استفاده از این سرویس می‌توانید همزمان که با استفاده از دیتابیس بزرگ این ابزار، پیشنهادهایی در مورد فیلم‌ها می‌گیرید، با تیم توسعه‌دهنده این ابزار کمک کنید تا الگوریتم‌هایی را تست کنند و توسعه دهند. یکی از ابزارک‌های این سایت، تنظیم‌کننده فیلم نام دارد. با استفاده از این ابزارک می‌توانید به دنبال فیلم‌هایی بگردید که مثلاً شبیه به Pulp Fiction است ولی میزان حادثه‌ای بودن فیلم بیشتر است.
- Criticker [۶۳] - این سایت از الگوریتمی با نام نشان‌دهنده امتیاز احتمال برای پیشنهاد دادن فیلم به شما استفاده می‌کند. نحوه کار این الگوریتم به این صورت است که از میانگین سلیقه ۱۰۰۰ کاربر که سلیقه آن‌ها بیشتری شباهت به شما را دارد برای پیشنهاد فیلم به شما استفاده می‌کنند. این ۱۰۰۰ کاربر از بین میلیون‌ها کاربر این سایت انتخاب می‌شود. در این سایت، در حین بررسی فیلم‌های مختلف، امتیازی در کنار هر فیلم نمایش داده می‌شود که این امتیاز، درصد احتمال این است که شما آن را فیلم را دوست داشته باشید. می‌توانید بر روی این درصد احتمال کلیک کنید تا از نحوه محاسبه آن و حتی رتبه‌بندی کاربرانی که در محاسبه این درصد احتمال دخیل بوده است مطلع شوید.
- Best similar [۶۴] - این سایت هر فیلم را با ۸ موضوع مشخص می‌کند و این امکان را دارد که بعد از انتخاب یک فیلم، بتوانید فیلم‌هایی را در هر یک از ۸ موضوع مرتبط با فیلمی که انتخاب کرده‌اید پیشنهاد بگیرید یا این به‌طورکلی فیلم‌های مشابه دیگری را پیدا کنید که بیشترین اشتراک موضوع با فیلم انتخابی شما را داشته باشند.
- Peach [۶۵] - این ابزار در ابتدا خود را سیستم توصیه‌گری برای پخش عمومی معرفی کرده است اما با بررسی دقیق‌تر مستندات آن، مشخص است که هدف این ابزار توصیه فیلم به کاربران به‌صورت شخصی است و پلتفرم هدف آن، سایت‌هایی مانند نماوا یا فیلیمو است.

این ابزار بخش وسیعی از نیازهای یک سازمان جهت ساخت یک سیستم توصیه‌گر را پوشش می‌دهد.

- دو نرم‌افزار موبایل برای دو سیستم‌عامل اندروید و iOS تهیه شده است که با استفاده از آن‌ها می‌توان اطلاعاتی از فیلم‌هایی که کاربران تماشا می‌کنند جمع‌آوری کرد و آن‌ها را به سرور فرستاد.
- شامل وبسایت‌های آماده‌ای برای ثبت‌نام و ورود کاربر و همچنین داشبوردی برای مدیریت کاربران توسط مدیر سایت است.
- ابزاری به نام Peach Lab تهیه شده است که مشابه با Jupyter Lab است. با استفاده از آن می‌توان با استفاده از داده‌های جمع‌آوری‌شده از کاربران، الگوریتم‌های تهیه شده توسط Peach و همچنین ابزارهای پایتون از پیش نصب شده، سیستم توصیه‌گری آموزش داد. چهار الگوریتم در این بخش موجود است که می‌توان به‌صورت مستقیم از آن‌ها استفاده کرد یا با توجه به نیازهای موجود، تغییراتی در آن‌ها ایجاد کرد. این الگوریتم‌ها عبارت‌اند از الگوریتم مبتنی بر محتوا، الگوریتم پالایش مشارکتی با استفاده از تجزیه ماتریس، الگوریتمی برای توصیه اقلام باهدف بالا بردن تنوع و الگوریتمی برای پیشنهاد موارد trend شده.
- در کنار الگوریتم‌های موجود، مجموعه قوانینی متنوعی نیز تعریف شده است. برای مثال می‌توان تعدادی از برنامه‌ها را در فهرست سیاه قرار داد تا هیچ وقت انتخاب نشوند یا می‌توان برای برنامه‌ها طول عمر تعریف کرد تا پس از اتمام این زمان، دیگر برنامه‌ها انتخاب نشوند.

کدهای بخش‌های اصلی این ابزار موجود نیست و فقط کدهایی از نرم‌افزارهای اندروید و iOS آن موجود است. اطلاعات قیمت‌گذاری و این‌که آیا امکان خرید آن وجود دارد نیز موجود نیست. به نظر می‌رسد که peach در فازهای ابتدایی توسعه است و هنوز به‌صورت رسمی ارائه نشده است.

۶- پیکره‌های آموزشی سیستم توصیه‌گر

با توجه به اینکه اکثریت (بر اساس دانش ما تا این لحظه) مدل‌های مطالعه شده با ناظر بودند، مدل‌های مطرح‌شده برای حل مسئله سیستم توصیه‌گر با استفاده از روش‌های با ناظر آموزش داده می‌شوند. عنصر اصلی در مدل‌های با ناظر در اختیار داشتن داده‌های آموزشی کافی و مناسب است. برای این منظور تلاش‌هایی برای تولید و ساخت پیکره‌های آموزشی انجام‌شده است.

جدول ۱. ژانر در پیکره متن‌باز MMTF14k

#	Genre	#pos	#neg	skewness
1	Action	1,766	11,857	6.71
2	Adventure	1,202	12,421	10.33
3	Animation	482	13,141	27.26
4	Children	592	13,031	22.01
5	Comedy	4,139	9,484	2.29
6	Crime	1,473	12,150	8.25
7	Documentary	1209	12414	10.27
8	Drama	6,592	7,031	1.07
9	Fantasy	737	12,886	17.48
10	Film-Noir	151	13,472	89.22
11	Horror	1,453	12,170	8.38
12	Musical	509	13,114	25.76
13	Mystery	754	12,869	17.07
14	Romance	2,003	11,620	5.80
15	Sci-Fi	938	12,685	13.52
16	Thriller	2,233	11,390	5.10
17	War	543	13,080	24.09
18	Western	323	13,300	41.18
Avg.		**1,505.5**	**12,118**	**18.65**

در مقاله [۲] یک پیکره آموزشی ساخته و منتشرشده است. در این مقاله، مجموعه داده MMTF-14K پیشنهاد و منتشرشده است. این دادگان برای استفاده در ساخت سیستم توصیه‌گر فیلم با استفاده از آخرین پیشرفت در نمایش محتوای دیداری و شنیداری طراحی‌شده است. این مجموعه داده به‌صورت عمومی در دسترس است و شامل ۱۳،۶۲۳ پیش‌پرده فیلم از نوع هالیوود است که توسط ۱۳۸،۴۹۲ کاربر با تقریب ۱۲،۵ میلیون رتبه تولیدشده است. مجموعه داده MMTF-14K اولین مجموعه داده در مقیاس بزرگ در جامعه سیستم‌های توصیه‌گر است که انواع توصیفات مبتنی بر محتوا را دارد. جدول ۱ مجموعه سبک‌ها و آمار آن‌ها در پیکره را نمایش می‌دهد [۵۰].

مجموعه داده MovieLens یکی دیگر از مجموعه‌های آموزشی قابل‌دسترس برای سیستم‌های توصیه‌گر MovieLens است. این مجموعه داده شامل نظرات مختلف کاربران در ارتباط با ۱۷۰۰ فیلم است. تعداد کاربران به کار گرفته‌شده برای ساخت این دادگان ۱۰۰۰ کاربر بوده است و در کل ۱۰۰۰۰۰ رتبه‌بندی از دیدگاه‌های مختلف برای مجموعه فیلم‌ها تولیدشده است [۶۲].

پیکره دادگان Spotify Audio Features، شامل آهنگ‌های سال‌های ۲۰۱۰ تا ۲۰۱۷ از سرویس معروف Spotify است. ویژگی‌های صوتی نظیر میزان بلندی، سرعت پخش و انرژی برای ۱۳۰ هزار آهنگ به‌وسیله Spotify Web API استخراج و شامل ۲۰۱۷ ردیف و ۱۷ ستون است. میزان بلندی، سرعت و انرژی آهنگ‌ها به ترتیب به‌صورت میانگین ۷،۰۳- دسی‌بل، ۱۲۱،۶ ضربه در هر دقیقه و ۰،۶۸(در بازه ۰ تا ۱) است [۹۰].

در پیکره دادگان معرفی‌شده در [۸۹]، مجموعه آهنگ‌ها و متون آن‌ها بین سال‌های ۱۹۵۰تا ۲۰۱۹ با استفاده از ترکیب Echo Nest API و بسته پایتون Spotify استخراج شد. حجم این پیکره حدود ۲۶،۴ مگابایت و شامل ۲۰۱۷ ردیف و ۱۷ ستون است.

۷- بحث و مقایسه

در جدول ۲ مقاله‌های سیستم‌های توصیه‌گر مبتنی بر محتوا در حوزه فیلم، موسیقی و تلویزیون، که در بخش ۴ توضیحات کامل آن‌ها ارائه‌شده، بررسی‌شده‌اند. در ادامه توضیحات بیشتری در تکمیل بررسی هر یک از این مقالات بیان خواهد شد.

جدول۲. مقاله‌های سیستم‌های توصیه‌گر مبتنی بر محتوا

ردیف	مقاله	الگوریتم	سال چاپ	حوزه	مجموعه داده	اطلاعات جانبی
۱	Liang و همکاران [۷]	ترکیبی	۲۰۱۴	فیلم	MovieLens	کلیدواژه
۲	Chen و همکاران [۸]	ترکیبی	۲۰۲۱	فیلم		
۳	Pradeep و همکاران [۹]	مبتنی بر محتوا	۲۰۲۰	فیلم		کلیدواژه
۴	Singh و همکاران [۱۰]	مبتنی بر محتوا	۲۰۲۰	فیلم	بدون دیتاست	
۵	Almeida و همکاران [۱۱]	مبتنی بر محتوا	۲۰۲۰	فیلم	MovieLens, IMDB	ایجاد تنوع در توصیه‌ها
۶	Meel و همکاران [۱۲]	ترکیبی	۲۰۲۱	فیلم	MovieLens	تگ
۷	Singla و همکاران [۱۳]	مبتنی بر محتوا	۲۰۲۰	فیلم	IMDB	متن داستان فیلم
۸	Sottocornola و همکاران [۱۴]	ترکیبی	۲۰۱۷	فیلم	MovieLens	متن داستان فیلم
۹	Yin و همکاران [۱۵]	مبتنی بر محتوا	۲۰۱۵	فیلم	Digg, MovieLens Douban Movie, Delicious	شبکه‌های اجتماعی
۱۰	Albanese و همکاران [۱۶]	ترکیبی	۲۰۱۳	فیلم	MovieLens	
۱۱	Elashmawi و همکاران [۸۳]	مبتنی بر محتوا	۲۰۲۲	فیلم	YouTube-8M	ویدئوهای یوتیوب
۱۲	Rajasekar و همکاران [۱۵]	ترکیبی	۲۰۲۵	فیلم	بدون دیتاست	
۱۳	Markapudi و همکاران [۸۴]	مبتنی بر محتوا	۲۰۲۳	فیلم	UGC	ویدئوهای پویانمایی، سخنرانی، موزیک ویدئو و ویدئوهای ورزشی

ردیف	مقاله	الگوریتم	سال چاپ	حوزه	مجموعه داده	اطلاعات جانبی
۱۴	Ord و همکاران [۱۷]	مبتنی بر محتوا	۲۰۱۳	موسیقی	The Million Song Dataset	محتوای صوتی موسیقی
۱۵	Wang و همکاران [۱۸]	مبتنی بر محتوا ترکیبی	۲۰۱۴	موسیقی	The Echo Nest Taste Profile Subset	محتوای صوتی موسیقی
۱۶	Liu و همکاران [۱۹]	مبتنی بر محتوا	۲۰۱۳	موسیقی	Custom -- 560 MP3+MIDI	محتوای صوتی موسیقی
۱۷	Bogdanov و همکاران [۲۰]	مبتنی بر محتوا	۲۰۱۳	موسیقی		محتوای صوتی موسیقی
۱۸	Wang و همکاران [۲۱]	مبتنی بر محتوا	۲۰۱۲	موسیقی	CAL500, Grooveshark	آگاه از زمینه
۱۹	Barragáns-Martínez و همکاران [۲۲]	مبتنی بر محتوا پالایش مشارکتی	۲۰۱۰	موسیقی		تگ
۲۰	Chiliguano و همکاران [۲۳]	مبتنی بر محتوا	۲۰۱۶	موسیقی	The Million Song Dataset	محتوای صوتی موسیقی
۲۱	Chou و همکاران [۲۴]	مبتنی بر محتوا	۲۰۱۶	موسیقی	Custom - A real-world dataset	محتوای صوتی موسیقی
۲۲	Turrin و همکاران [۲۵]	مبتنی بر محتوا	۲۰۱۵	موسیقی	30Music	تاریخچه کوتاه‌مدت کاربر
۲۳	Vasile و همکاران [۲۶]	ترکیبی	۲۰۱۶	موسیقی	30Music	تعبیه اقلام
۲۴	Wang و همکاران [۲۷]	مبتنی بر محتوا	۲۰۱۴	موسیقی	Custom - 400 Songs	محتوای صوتی موسیقی
۲۵	Rao و همکاران [۲۸]	مبتنی بر محتوا	۲۰۲۱	موسیقی		محتوای صوتی موسیقی
۲۶	Hosseinzadeh و همکاران [۲۹]	مبتنی بر محتوا	۲۰۱۵	موسیقی	Last.fm	آگاه از دنباله
۲۷	Jannach و همکاران [۳۰]	ترکیبی	۲۰۱۵	موسیقی	Last.fm, AotM, 8tracks	آگاه از دنباله
۲۸	Vall و همکاران [۳۱]	ترکیبی	۲۰۱۸	موسیقی	The Million Song Dataset, AotM, 8track	تگ
۲۹	Vall و همکاران [۳۲]	ترکیبی	۲۰۱۷	موسیقی	The Million Song Dataset, AtoM	محتوای صوتی موسیقی
۳۰	Gharahighehi و همکاران [۳۳]		۲۰۲۱	موسیقی	30music, Last.fm, Nowplaying	مبتنی نشست
۳۱	Ali و همکاران [۸۵]	مبتنی بر محتوا	۲۰۲۳	موسیقی	Kaggle (Spotify, Genius)	آهنگ‌های سال‌های ۲۰۱۰ تا ۲۰۱۷
۳۲	Burch و همکاران [۸۶]	مبتنی بر محتوا	۲۰۲۳	موسیقی	top five playlists for each of the four moods using Spotify's API	دیتا برای کاربرد خاص خود که پیش‌بینی مود کاربران هست جمع‌آوری شده
۳۳	Shashaani	مبتنی بر محتوا	۲۰۲۴	موسیقی		
۳۴	Engelbert و همکاران [۳۴]	مبتنی بر محتوا	۲۰۱۱	تلویزیون	Custom - Simulation	
۳۵	Quan و همکاران [۳۵]	مبتنی بر محتوا	۲۰۱۴	تلویزیون	Custom - Survey from 20 people for 1 month	رفتارهای گروهی کاربران
۳۶	Oh و همکاران [۳۶]	مبتنی بر محتوا	۲۰۱۲	تلویزیون		رفتارهای گروهی کاربران
۳۷	Turrin و همکاران [۳۷]	مبتنی بر محتوا	۲۰۱۴	تلویزیون	Custom - 13664 users, 217 channels, 4 months	علاقه‌مندی کاربر به یک شبکه تلویزیونی
۳۸	daSilva و همکاران [۳۸]	مبتنی بر محتوا	۲۰۱۲	تلویزیون	Custom	
۳۹	Zibriczky و همکاران [۳۹]	مبتنی بر محتوا پالایش مشارکتی ترکیبی	۲۰۱۲	تلویزیون	SaskTel	
۴۰	Song و همکاران [۴۰]	مبتنی بر محتوا	۲۰۱۲	تلویزیون		
۴۱	Chaudhry و همکاران [۴۱]	مبتنی بر محتوا	۲۰۱۵	تلویزیون		
۴۲	Barraza-Urbina و همکاران [۴۲]	ترکیبی	۲۰۱۵	تلویزیون	Twitter Data	اطلاعات شبکه‌های اجتماعی
۴۳	Yuan و همکاران [۴۴]	مبتنی بر محتوا	۲۰۱۵	تلویزیون	Custom - Campus IPTV, 1 year	
۴۴	Symeonidis و همکاران [۴۵]		۲۰۲۰	تلویزیون		آگاه از دنباله
۴۵	Dudekula ، و همکاران [۸۷]	ترکیبی	۲۰۲۳	تلویزیون	CelebFaces Attribute Dataset, Labeled Faces in the Wild-People	تلویزیون‌های هوشمند

دسته‌ای از مقاله‌ها، از رویکردهای پایه مبتنی بر محتوا استفاده کرده‌اند. وجود مقالاتی بر مبنای رویکردهای پایه مبتنی بر محتوا بر ارزشمند بودن این رویکردها صحه می‌گذارد و نشان می‌دهد اگرچه رویکردهای پایه عمر طولانی‌ای دارند همچنان کاربردی هستند و هنوز به تاریخ انقضا نرسیده‌اند. درنتیجه این رویکردها و ویژگی‌های استفاده شده در آن‌ها می‌تواند در پروژه‌های پژوهشی و صنعتی کاربردی باشد.

به‌عنوان‌مثال در مقاله [۹] از فهرست بازیگران، کلیدواژه، عوامل

سازنده فیلم و سبک‌های فیلم استفاده شده است. مقاله [۱۰] از ویژگی‌های محبوبیت فیلم و سبک فیلم استفاده کرده است. در مقاله [۱۳] از ویژگی‌های داستان فیلم، امتیازهای فیلم، کشور سازنده فیلم و سال ساخت استفاده شده است. در مقاله [۳۲]، از ویژگی‌هایی ازجمله شبکه تلویزیونی که برنامه از آن پخش‌شده است، عنوان برنامه، زیرنویس برنامه، دسته برنامه (مثلاً فیلم، مسابقه تلویزیونی یا ...)، سبک، بازیگران، توضیحات و سال ساخت استفاده شده است. در مقاله [۳۸] از ویژگی‌های زمینه‌ای (تاریخ و زمان و مکان ساخت برنامه تلویزیونی و مکان کاربر)، ویژگی‌های صریح که مستقیماً از کاربر گرفته می‌شود (زبان، جنسیت، سن و شغل) و ویژگی‌هایی که به‌طور خودکار استخراج می‌شوند (مانند سبک) استفاده شده است. در مقاله [۳۹] از ویژگی‌های عنوان، سبک، کانال پخش‌کننده، دسته (category) و رده‌بندی سنی استفاده شده است. مقاله [۴۰]، از ویژگی‌های ثابتی شامل نمایه کاربر (نام کاربری، نام، سن و جنسیت کاربر) و ترجیح ثابت کاربر (زبان زیرنویس و زبان صدایی که کاربر ترجیح می‌دهد، نوع محتوای موردعلاقه، بازیگر موردعلاقه و کارگردان موردعلاقه) استفاده کرده است. مقاله [۴۴] بر روی انتخاب ویژگی‌های مناسب با استفاده از تخصیص نهان دیریکله تمرکز کرده است. در مقاله [۸۳] به استخراج ویژگی‌های اشیا موجود مانند چهره‌ها و صحنه‌ها، صداها و کلمات موجود می‌پردازد. در مقاله [۸۴] ویژگی‌های زمانی و مکانی هر قطعه به همراه ویژگی مربوط به رنک به‌کاررفته درصحنه‌های مختلف را استخراج می‌کند.

برخی از مقاله‌ها بر روی ایجاد تنوع در توصیه‌ها تمرکز کرده‌اند. نمونه‌ای از این مقاله‌ها، مقاله [۱۱] است. مقاله [۲۷] نیز بر روی ایده اکتشاف که باعث ایجاد تنوع در توصیه می‌شود تمرکز کرده است.

تعدادی از مقاله‌ها، از کلیدواژه‌ها به‌عنوان بردار ویژگی استفاده کرده‌اند. در برخی دادگان برای هر فیلم خلاصه اثر نیز موجود است. می‌توان از خلاصه‌های موجود، کلیدواژه استخراج کرد و سپس از رویکردهای ارائه‌شده در این مقاله‌ها بهره برد. مقاله‌های [۷] و [۹] ازجمله مقالاتی هستند که از اطلاعات کلیدواژه‌ها استفاده کرده‌اند.

تعدادی از مقاله‌ها از تگ‌ها در طراحی سیستم توصیه‌گر استفاده کرده‌اند. به‌عنوان‌مثال در مقاله [۱۲] ایده پردازش معنایی تگ‌ها ارائه شده است. در مقاله [۳۱]، از word2vec برای استخراج اطلاعات معنایی از تگ‌ها استفاده شده است. در مقاله [۳۰]، از TF-IDF برای یافتن تگ‌های مهم‌تر و مؤثرتر استفاده شده است. مقاله [۲۲] نیز از ایده برچسب استفاده کرده است؛ اما تعریف برچسب در این مقاله توسط کاربر بوده است.

در برخی مقاله‌ها، از ایده‌های حوزه پردازش زبان طبیعی استفاده شده است. به‌عنوان‌مثال، در مقاله [۱۳] و مقاله [۱۴]، الگوریتم‌های داک‌تووک و TD-IDF بر روی متن داستان فیلم اعمال‌شده است. در صورت عدم وجود متن داستان می‌توان این روش‌ها را روی خلاصه‌ای از فیلم اعمال کرد. علاوه بر داک‌تووک سایر بازنمایی‌های قوی‌تر معنایی نیز می‌تواند در این بخش به کار رود.

چالش‌های دیگری در زمینه پیشنهاد برنامه به کاربر وجود دارد. ممکن است فیلمی به کاربر پیشنهاد شود که بسیار شبیه به سلیقه کاربر است، اما امتیاز پایینی دارد. یا این‌که فیلمی با داستان مشابه ولی با سبکی متفاوت به کاربر پیشنهاد شود؛ مثلاً فیلمی که به طنز از روی داستان فیلم دیگری ساخته شده است. مقاله [۱۳] به حل این چالش‌ها پرداخته است.

دسته‌ای از مقاله‌ها نیز ایده جدیدی در بخش یادگیری ماشین ارائه داده‌اند. به‌عنوان‌مثال، مقاله [۱۴]، ایده استفاده از جنگل تصادفی را ارائه داده است. مقاله [۳۴] از طبقه‌بندی بیز ساده استفاده می‌کند که از روش‌های متداول حوزه سیستم توصیه‌گر است. مقاله [۳۵] از شبکه‌های بیزین استفاده کرده است. در مقاله [۴۰] شبکه عصبی به‌کاررفته است که از مدل‌های روز می‌باشد. مقاله [۸۷] از شبکه عصبی پیچشی و متادیتاهای موجود در برنامه استفاده کرده است.

برخی از مقاله‌ها به علاقه‌مندی کاربر به یک شبکه تلویزیونی پرداخته‌اند. علاقه‌مندی کاربر به یک شبکه مفهومی کلی‌تر از علاقه‌مندی به یک برنامه است. نمونه‌ای از این مقاله‌ها، مقاله [۳۷] است. این مقاله به این موضوع پرداخته است که در بازه‌های زمانی علاقه‌مندی کاربر تفاوت می‌کند. مثلاً در ایام خاص سال بر اساس مناسبت‌ها علاقه‌مندی خاصی مطرح است.

پایه اصلی توصیه در برخی از سیستم‌ها، استفاده از اطلاعات شبکه‌های اجتماعی است. این روش‌ها به‌شرطی کاربردی هستند که این اطلاعات در سیستم موجود باشد. نمونه‌ای از این مقاله‌ها، مقاله [۱۵] و [۴۲] است.

تعداد زیادی از مقاله‌های حوزه موسیقی، بر روی محتوای صوتی موسیقی تمرکز کرده‌اند و به استخراج ویژگی‌های صوتی و به دست آوردن بازنمایی از این ویژگی‌ها پرداخته‌اند. مقاله‌های [۱۷, ۱۸, ۱۹, ۲۰, ۲۳, ۲۴, ۲۷, ۲۸, ۳۲, ۸۴, ۸۶] از این دسته هستند.

برخی از پژوهش‌ها در زمینه تعبیه اقلام صورت گرفته است که حجم داده خیلی زیادی از تاریخچه رفتارها موردنیاز است. که این حجم داده در دسترس ما نیست. مقاله [۲۶] از این دسته است.

حالت آگاه از دنباله که در تغییر سلیقه کاربر وجود دارد یکی دیگر

از موضوعات مورد تمرکز در این حوزه است. در این روش‌ها، به حافظه نزدیک اهمیت بیشتری داده می‌شود. مقاله [۲۹] و [۳۰] و [۴۵] از این دسته هستند.

برخی از مقاله‌ها، رویکرد آگاه از زمینه‌دارند. نمونه‌ای از این مقاله‌ها، مقاله [۲۱] است. در این مقاله فعالیت کاربر درک می‌شود و با توجه به آن، به کاربر موسیقی پیشنهاد می‌شود. درنتیجه مثلاً موسیقی‌ای که در هنگام کار به کاربر پیشنهاد می‌شود با موسیقی زمان تفریح متفاوت است.

تمرکز اصلی برخی از مقالات بر روی روش‌های مبتنی نشست است. مقاله [۳۳] از این دسته است.

برخی از مقاله‌ها، بر روی رفتارهای گروهی کاربران تمرکز دارند و به گروهی از کاربران پیشنهاد می‌دهند. نمونه‌ای از این مقاله‌ها، مقاله [۳۵] و [۳۶] است.

۸- خلاصه و جمع‌بندی

این پژوهش باهدف بررسی جامع سیستم‌های توصیه‌گر چندرسانه‌ای و با تمرکز ویژه بر رویکردهای مبتنی بر محتوا، به تحلیل عمیق ۴۷ مقاله علمی و مطالعه معماری سیستم‌های شرکت‌های پیشرو در این صنعت پرداخت. تحلیل ادبیات موضوع نشان می‌دهد که حوزه سیستم‌های توصیه‌گر، مسیری تکاملی و پرشتاب را از مدل‌های سنتی به سمت معماری‌های پیچیده و چندوجهی طی کرده است. این مسیر، صرفاً یک پیشرفت الگوریتمی نبوده، بلکه بیانگر تحولی بنیادین در درک مفهوم «محتوا» و «زمینه» در شخصی‌سازی تجربه کاربر است.

یکی از کلیدی‌ترین نتایج این بررسی، تأیید گذار از مدل‌های خالص به سمت سیستم‌های ترکیبی به‌عنوان راهکار غالب و مؤثر در این حوزه است. در حالی که روش‌های پالایش مشارکتی در شناسایی الگوهای رفتاری کاربران قدرتمند هستند، با چالش‌هایی نظیر «شروع سرد» و «پراکندگی داده»[1] مواجه‌اند. از سوی دیگر، رویکردهای مبتنی بر محتوا، با تحلیل ویژگی‌های ذاتی اقلام، این شکاف‌ها را به خوبی پوشش می‌دهند و قابلیت توضیح‌پذیری بالاتری را فراهم می‌آورند. موفقیت شرکت‌هایی مانند نتفلیکس، گواهی بر کارآمدی ترکیب هوشمندانه این دو رویکرد است که در آن، از پالایش مشارکتی برای درک سلایق جامعه کاربران و از تحلیل محتوا برای توصیه‌های دقیق، متنوع و قابل توجیه بهره گرفته می‌شود.

دومین یافته مهم، بازتعریف و تعمیق مفهوم «محتوا» است. در پژوهش‌های اولیه، تحلیل محتوا به فراداده‌های ساده‌ای مانند سبک، بازیگران یا کلمات کلیدی محدود بود. اما امروزه، با پیشرفت در حوزه‌های پردازش زبان طبیعی، بینایی ماشین و پردازش سیگنال، تحلیل محتوا به لایه‌های عمیق‌تری نفوذ کرده است. مقالات موردبررسی نشان می‌دهند که استفاده از تکنیک‌هایی مانند word2vec و Doc2Vec برای استخراج بازنمایی معنایی از خلاصه داستان یا تگ‌های کاربران، و همچنین به‌کارگیری شبکه‌های عصبی کانولوشنی برای تحلیل مستقیم سیگنال صوتی موسیقی یا فریم‌های ویدئویی، به یک استاندارد برای درک دقیق‌تر اقلام تبدیل شده است. این تکامل، سیستم‌ها را قادر ساخته تا شباهت‌های ظریف و غیربدیهی میان اقلام را کشف کرده و از دام «بیش‌تخصصی شدن» رها شوند.

سومین روند کلیدی، حرکت به سوی سیستم‌های توصیه‌گر آگاه از زمینه و پویا است. پژوهش‌های نوین دریافته‌اند که سلیقه کاربر یک مفهوم ثابت نیست، بلکه تحت تأثیر عوامل متعددی مانند زمان روز، مکان، فعالیت جاری (مانند ورزش یا مطالعه) و حتی همراهان کاربر (در سناریوهای تماشای گروهی تلویزیون) تغییر می‌کند. مدل‌هایی که این اطلاعات زمینه‌ای را ادغام می‌کنند، مانند مدل‌های مبتنی بر زنجیره مارکوف برای درک توالی رفتار کاربر یا پروفایل‌سازی وابسته به زمان برای تفکیک اعضای خانواده، توانسته‌اند دقت و مقبولیت توصیه‌ها را به شکل چشمگیری افزایش دهند. این رویکرد، مسئله توصیه را از «چه چیزی را پیشنهاد دهیم؟» به «چه چیزی را در این لحظه و این شرایط خاص پیشنهاد دهیم؟» ارتقا داده است

درنهایت، بررسی سیستم‌های صنعتی مانند یوتیوب و اینستاگرام نشان داد که مقیاس‌پذیری و کارایی، چالش اصلی در دنیای واقعی است. معماری دو مرحله‌ای «تولید کاندیدا و رتبه‌بندی»[2] یک الگوی رایج و موفق برای مواجهه با این چالش است. در این معماری، ابتدا یک الگوریتم سریع و نسبتاً ساده، از میان میلیون‌ها قلم، چند صد کاندیدای مرتبط را استخراج می‌کند و سپس یک مدل پیچیده‌تر و پرهزینه‌تر (معمولاً مبتنی بر یادگیری عمیق)، این کاندیداها را با دقت بالا رتبه‌بندی می‌نماید. این الگو، تعادلی هوشمندانه میان دقت و هزینه‌های محاسباتی برقرار می‌کند.

در جمع‌بندی، واضح است که آینده سیستم‌های توصیه‌گر چندرسانه‌ای در گروی توسعه مدل‌های ترکیبی، چندوجهی، آگاه از زمینه و مقیاس‌پذیر قرار دارد. چالش‌های پیش رو دیگر صرفاً به بهبود دقت محدود نمی‌شوند، بلکه ایجاد تعادل میان دقت و تنوع، افزایش شفافیت و توضیح‌پذیری مدل‌های یادگیری عمیق، تضمین

[1] Data Sparsity

[2] Candidate Generation & Ranking

انصاف و پرهیز از سوگیری و توسعه رابط‌های کاربری تعاملی و مکالمه‌ای، افق‌های تحقیقاتی آینده را ترسیم می‌کنند. موفقیت در این حوزه، نیازمند هم‌افزایی میان تحلیل عمیق محتوا، درک پویای رفتار کاربر و مهندسی سیستم‌های کارآمد است تا تجربه‌ای غنی، مفید و رضایت‌بخش برای کاربر نهایی فراهم آید.

سپاسگزاری

از شرکت آپاسای داده سیستم که ما را در اجرای تحقیق یاری کرده‌اند و حمایت مالی پژوهشی را بر عهده داشته‌اند سپاسگزاریم.

مراجع

[1] Deldjoo, Y., et al. Multimedia recommender systems. in Proceedings of the 12th ACM Conference on Recommender Systems. 2018.

[2] Deldjoo, Y., et al. MMTF-14K: a multifaceted movie trailer feature dataset for recommendation and retrieval. in Proceedings of the 9th ACM Multimedia Systems Conference. 2018.

[3] Chong, D., Deep dive into netflix's recommender system. 2020, Medium.

[4] Amato, F., et al., SOS: A multimedia recommender System for Online Social networks. Future generation computer systems, 2019. 93: p. 914-923.

[5] Liang, T., et al., A hybrid recommendation model based on estimation of distribution algorithms. Journal of Computational Information Systems, 2014. 10(2): p. 781-788.

[6] Chen, Y.-L., Y.-H. Yeh, and M.-R. Ma, A movie recommendation method based on users' positive and negative profiles. Information Processing & Management, 2021. 58(3): p. 102531.

[7] Pradeep, N., et al., Content based movie recommendation system. International Journal of Research in Industrial Engineering, 2020. 9(4): p. 337-348.

[8] Singh, R.H., et al., Movie recommendation system using cosine similarity and KNN. International Journal of Engineering and Advanced Technology, 2020. 9(5): p. 556-559.

[9] Almeida, M.S. and A. Britto. MOEA-RS: A Content-Based Recommendation System Supported by a Multi-objective Evolutionary Algorithm. in International Conference on Artificial Intelligence and Soft Computing. 2020. Springer.

[10] Meel, P., et al. Movie Recommendation Using Content-Based and Collaborative Filtering. in International Conference on Innovative Computing and Communications. 2021. Springer.

[11] Singla, R., et al. FLEX: A Content Based Movie Recommender. in 2020 International Conference for Emerging Technology (INCET). 2020. IEEE.

[12] Sottocornola, G., et al. Towards a deep learning model for hybrid recommendation. in Proceedings of the International Conference on Web Intelligence. 2017.

[13] Yin, H., et al., Dynamic user modeling in social media systems. ACM Transactions on Information Systems (TOIS), 2015. 33(3): p. 1-44.

[14] Albanese, M., et al., A multimedia recommender system. ACM Transactions on Internet Technology (TOIT), 2013. 13(1): p. 1-32.

[15] Rajasekar, R., Radhakrishnan, N., Sridar, K., Viji, C., Mohanraj, M., Kalpana, C., & Rajkumar, N. (2025). Intelligent movie recommendation system. Salud, Ciencia y Tecnología-Serie de Conferencias, (4), 1438.

[16] Shashaani, S. (2024, October). Explainability in music recommender system. In Proceedings of the 18th ACM Conference on Recommender Systems (pp. 1395-1401).

[17] Van Den Oord, A., S. Dieleman, and B. Schrauwen. Deep content-based music recommendation. in Neural Information Processing Systems Conference (NIPS 2013). 2013. Neural Information Processing Systems Foundation (NIPS).

[18] Wang, X. and Y. Wang. Improving content-based and hybrid music recommendation using deep learning. in Proceedings of the 22nd ACM international conference on Multimedia. 2014.

[19] Liu, N.-H., Comparison of content-based music recommendation using different distance estimation methods. Applied intelligence, 2013. 38(2): p. 160-174.

[20] Bogdanov, D., et al., Semantic audio content-based music recommendation and visualization based on user preference examples. Information Processing & Management, 2013. 49(1): p. 13-33.

[21] Wang, X., D. Rosenblum, and Y. Wang. Context-aware mobile music recommendation for daily activities. in Proceedings of the 20th ACM international conference on Multimedia. 2012.

[22] Barragáns-Martínez, A.B., et al., Exploiting social tagging in a web 2.0 recommender system. IEEE Internet Computing, 2010. 14(6): p. 23-30.

[23] Chiliguano, P. and G. Fazekas. Hybrid music recommender using content-based and social information. in 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2016. IEEE.

[24] Chou, S.-Y., et al. Addressing cold start for next-song recommendation. in Proceedings of the 10th ACM Conference on Recommender Systems. 2016.

[25] R. Turrin, A.C., R. Pagano, M. Quadrana, and P. Cremonesi, Large scale music recommendation, in LSRS. 2015.

[26] Vasile, F., E. Smirnova, and A. Conneau. Meta-prod2vec: Product embeddings using side-information for recommendation. in Proceedings of the 10th ACM Conference on Recommender Systems. 2016.

[27] Wang, X., et al., Exploration in interactive personalized music recommendation: a reinforcement learning approach. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), 2014. 11(1): p. 1-22.

[28] Rao, A., et al. Supervised Feature Learning for Music Recommendation. in International Symposium on Signal Processing and Intelligent Recognition Systems. 2020. Springer.

[29] Hosseinzadeh Aghdam, M., et al. Adapting recommendations to contextual changes using hierarchical hidden markov models. in Proceedings of the 9th ACM Conference on Recommender Systems. 2015.

[30] Jannach, D., L. Lerche, and I. Kamehkhosh. Beyond" hitting the hits" Generating coherent music playlist continuations with the right tracks. in Proceedings of the 9th ACM Conference on Recommender Systems. 2015.

[31] Vall, A., et al. A hybrid approach to music playlist continuation based on playlist-song membership. in Proceedings of the 33rd Annual ACM Symposium on Applied Computing. 2018.

[32] Vall, A., et al. Music playlist continuation by learning from hand-curated examples and song features: Alleviating the cold-start problem for rare and out-of-set songs. in Proceedings of the 2nd Workshop on Deep Learning for Recommender Systems. 2017.

[33] Gharahighehi, A. and C. Vens, Personalizing diversity versus accuracy in session-based recommender systems. SN Computer Science, 2021. 2(1): p. 1-12.

[34] Engelbert, B.B., M.B.; Kruthoff-Bruwer, R.; Morisse, K. A user supporting personal video recorder by implementing a generic Bayesian classifier based recommendation system. in Proceedings of

[54] Metaxiotis, K., et al., Decision support through knowledge management: the role of the artificial intelligence. Information Management & Computer Security, 2003.

[55] Wondercube® TV Broadcasting GSD/HD digital broadcast. Available from: http://www.jvc.slak.si/pdf/tv%20broadcasting.pdf.

[56] Recommendation Systems: Applications, Examples & Benefits. Available from: https://research.aimultiple.com/recommendation-system/#amazoncom.

[57] Optimizely. Available from: https://aimultiple.com/.

[58] CRM. Available from: https://www.cnbc.com/quotes/CRM.

[59] Clarifai. Available from: https://www.clarifai.com/.

[60] Jinni. Available from: http://www.jinni.com/.

[61] RottenTomatoes. Available from: www.rottentomatoes.com.

[62] MovieLens 100K Dataset. Available from: https://grouplens.org/datasets/movielens/100k/.

[63] Criticker. Available from: http://www.criticker.com/.

[64] Bestsimilar. Available from: bestsimilar.com.

[65] PEACH. 2021 [cited 2021 1 June 2021]; Available from: https://peach.ebu.io/.

[66] Sidana, S., Recommendation systems for online advertising. 2018, Université Grenoble Alpes.

[67] Netflix Prize Data Set. Available from: https://academictorrents.com/details/9b13183dc4d60676b773c9e2cd6de5e5542cee9a.

[68] Eliashberg, J., et al., Demand-driven scheduling of movies in a multiplex. International Journal of Research in Marketing, 2009. 26(2): p. 75-88.

[69] Zarandi, M.H.F., et al., A state of the art review of intelligent scheduling. Artificial Intelligence Review, 2020. 53(1): p. 501-593.

[70] Brown, D.E., J.A. Marin, and W.T. Scherer, A survey of intelligent scheduling systems, in Intelligent Scheduling Systems. 1995, Springer. p. 1-40.

[71] Horen, J.H., Scheduling of network television programs. Management Science, 1980. 26(4): p. 354-370.

[72] Welbank, M., A review of knowledge acquisition techniques for expert systems. 1983: Martlesham Consultancy Services Martlesham Heath, Ipswich.

[73] Piroozfard, H., K.Y. Wong, and A. Hassan, A hybrid genetic algorithm with a knowledge-based operator for solving the job shop scheduling problems. Journal of Optimization, 2016. 2016.

[74] Mitchell, T.M., Machine learning. 1997.

[75] Kocsis, T., et al., Case-Based Reasoning system for mathematical modelling options and resolution methods for production scheduling problems: Case representation, acquisition and retrieval. Computers & Industrial Engineering, 2014. 77: p. 46-64.

[76] Panaggio, M.J., et al., Prediction and Optimal Scheduling of Advertisements in Linear Television. arXiv preprint arXiv:1608.07305, 2016.

[77] Ghassemi Tari, F. and R. Alaei, Scheduling TV commercials using genetic algorithms. International Journal of Production Research, 2013. 51(16): p. 4921-4929.

[78] AlShami, H., Optimising television programming and scheduling. 2017: Lancaster University (United Kingdom).

[79] A case study of designing TV schedules. Available from: http://www.kr.inf.uc3m.es/wp-content/uploads/2019/12/AI4SE-Meer.pdf.

[80] Le, Q. and T. Mikolov. Distributed representations of sentences and documents. in International conference on machine learning. 2014. PMLR.

the 2011 IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops). 2011.

[35] Quan, J.-C.C., S.-B, A Hybrid Recommender System Based on AHP That Awares Contexts with Bayesian Networks for Smart TV, in International Conference on Hybrid Artificial Intelligence Systems (HAIS 2014). 2014.

[36] Oh, J.S., Y.; Kim, J.; Humayoun, M.; Park, Y.H.; Yu, H. Time-dependent user profiling for TV recommendation. in Proceedings of the 2nd International Conference on Cloud and Green Computing and 2nd International Conference on Social Computing and Its Applications, CGC/SCA 2012. 2012.

[37] Turrin, R.C., A.; Cremonesi, P.; Pagano, R. Time-based TV programs prediction. in Proceedings of the 1st Workshop on Recommender Systems for Television and Online Video (RecSysTV) 2014. 2014.

[38] da Silva, F.S., L.G.P. Alves, and G. Bressan, Personal TVware: an infrastructure to support the context-aware recommendation for personalized digital TV. International Journal of Computer Theory and Engineering, 2012. 4(2): p. 131.

[39] Zibriczky, D., et al. Personalized recommendation of linear content on interactive TV platforms: beating the cold start and noisy implicit user feedback. in UMAP workshops. 2012.

[40] Song, S., H. Moustafa, and H. Afifi, Advanced IPTV services personalization through context-aware content recommendation. IEEE Transactions on Multimedia, 2012. 14(6): p. 1528-1537.

[41] Chaudhry, M., et al. Heterogeneous information network based TV program recommendation. in Proceedings of the 16th International Symposium on Advanced Intelligent Systems. 2015.

[42] Barraza-Urbina, A., et al., Using social media data for online television recommendation services at RTÉ Ireland. 2015.

[43] Hromic, H., et al., Event panning in a stream of big data. Research Day 2013 Schedule, 2012: p. 44.

[44] Yuan, J., et al., Context-aware LDA: Balancing relevance and diversity in TV content recommenders. 2015.

[45] Symeonidis, P., et al. Recommending the video to watch next: an offline and online evaluation at YOUTV. de. in Fourteenth ACM conference on recommender systems. 2020.

[46] Basilico, J. Recent Trends in Personalization: A Netflix Perspective. 2019. ICML.

[47] Covington, P., J. Adams, and E. Sargin. Deep neural networks for youtube recommendations. in Proceedings of the 10th ACM conference on recommender systems. 2016.

[48] Implementing the YouTube Recommendations Paper in TensorFlow — Part 1. Available from: https://theiconic.tech/implementing-the-youtube-recommendations-paper-in-tensorflow-part-1-d1e1299d5622.

[49] Krawiec, T., The Amazon Recommendations Secret to Selling More Online. Rejoiner http://rejoiner. com/resources/amazon-recommendations-secret-selling-online/website visited, 2018: p. 1-18.

[50] Mangalindan, J. Amazon's recommendation secret. 2012; Available from: https://fortune.com/2012/07/30/amazons-recommendation-secret/.

[51] Linden, G., B. Smith, and J. York, Amazon. com recommendations: Item-to-item collaborative filtering. IEEE Internet computing, 2003. 7(1): p. 76-80.

[52] Baatarjav, E.-A., S. Phithakkitnukoon, and R. Dantu. Group recommendation system for facebook. in OTM Confederated International Conferences" On the Move to Meaningful Internet Systems". 2008. Springer.

[53] Facebook recommender system. Available from: https://github.com/facebookresearch/dlrm.

[81] He, X., et al. Neural collaborative filtering. in Proceedings of the 26th international conference on world wide web. 2017.

[82] Cheng, H.-T., et al. Wide & deep learning for recommender systems. in Proceedings of the 1st workshop on deep learning for recommender systems. 2016.

[83] Roshdy, Youssef, and Mennat Allah Hassan. "An Efficient Content-Based Video Recommendation." Journal of Computing and Communication 1.1 (2022): 48-64.

[84] Markapudi, Baburao, et al. "Content-based video recommendation system (CBVRS): a novel approach to predict videos using multilayer feed forward neural network and Monte Carlo sampling method." Multimedia Tools and Applications 82.5 (2023): 6965-6991.

[85] Ali, Shaik Faizan Roshan, et al. "Recommender System using Audio and Lyrics." 2023 4th International Conference on Electronics and Sustainable Communication Systems (ICESC). IEEE, 2023.

[86] Burch, Charats, Robert Sprowl, and Mehmet Ergezer. "A multi-user virtual world with music recommendations and mood-based virtual effects." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 37. No. 13. 2023.

[87] Dudekula, Khasim Vali, et al. "Convolutional Neural Network-Based Personalized Program Recommendation System for Smart Television Users." Sustainability 15.3 (2023): 2206.

[88] Boeker, Maximilian and Aleksandra Urman. "An Empirical Investigation of Personalization Factors on TikTok." Proceedings of the ACM Web Conference 2022 (2022): n. pag.

[89] Liu, Zhuoran, et al. "Monolith: real time recommendation system with collisionless embedding table." arXiv preprint arXiv:2209.07663 (2022).

[90] Ansari, Farooq. The Spotify Audio Features Hit Predictor Dataset (1960-2019). 1, 4TU.Centre for Research Data, 6 Apr. 2020, doi:10.4121/UUID:D77E74B0-66BC-47AC-8B25-5796D3084478.

[91] Moura, Luan; Fontelles, Emanuel; Sampaio, Vinicius; França, Mardônio (2020), "Music Dataset: Lyrics and Metadata from 1950 to 2019", Mendeley Data, V2, doi: 10.17632/3t9vbwxgr5.2

[92] Deldjoo, Y., Schedl, M., Hidasi, B., Wei, Y., He, X. (2022). Multimedia Recommender Systems: Algorithms and Challenges. In: Ricci, F., Rokach, L., Shapira, B. (eds) Recommender Systems Handbook. Springer, New York, NY. https://doi.org/10.1007/978-1-0716-2197-4_25

[93] Deldjoo, Yashar and Schedl, Markus and Cremonesi, Paolo and Pasi, Gabriella, "Recommender Systems Leveraging Multimedia Content", ACM Computing Surveys, 2020, 53, 5,

[94] A Survey of Multimedia Recommender Systems: Challenges and Opportunities Mouzhi Ge (1Faculty of Informatics, Masaryk University, Brno 60200, Czech Republic) and Fabio Persia (2Faculty of Computer Science, Free University of Bozen-Bolzano, Bozen-Bolzano, 39100, Italy) International Journal of Semantic Computing 2017 11:03, 411-428

[95] Jiang, Y., Xia, L., Wei, W., Luo, D., Lin, K., & Huang, C. (2024, October). Diffmm: Multi-modal diffusion model for recommendation. In Proceedings of the 32nd ACM International Conference on Multimedia (pp. 7591-7599).

[96] Malitesta, D., Cornacchia, G., Pomo, C., Merra, F. A., Di Noia, T., & Di Sciascio, E. (2025). Formalizing multimedia recommendation through multimodal deep learning. ACM Transactions on Recommender Systems, 3(3), 1-33.