**ششمین کنفرانس بین‌المللی**

**«بازی‌های رایانه‌ای؛ فرصت‌ها و چالش‌ها»**

**پیاده‌سازی سیستم توصیه‌گر برای بازی های موبایلی براساس آیتم‌های**

**انتخابی کاربر (Item Base)**

**شمیمه صنیع‌ثالث\***

1. **پژوهشگر و پیاده سازی**

Shamims12@yahoo.com

**چکیده:**

بر اساس تحقیقات، این روزها اکثر افراد در اوقات فراغت به کار با گوشی‌های هوشمند مشغول هستند و بسیاری از کودکان، نوجوانان، بزرگسالان و حتی سالمندان برای سرگرمی در این اوقات به بازی‌های موبایلی روی می‌آوردند. از دغدغه‌های اصلی هر گروه سنی انتخاب ژانر بازی مناسب سلایق آن‌ها و بازی‌های مشابه آن است. بسیاری از والدین می‌خواهند فرزندشان از بازی‌های خشن دور بماند یا در هنگام بازی به یادگیری نیز بپردازد. به طور کلی هر گروه سنی علاقه‌مندی و هدف خود را از بازی کردن دارد و دیدن بسیاری بازی در یک اپلیکیشن مانند گوگل‌پلی، بازار یا ... برای اغلب مردم گیج کننده‌است.

هدف اصلی این سامانه پیاده‌سازی یک سیستم توصیه‌گر مبتنی‌بر داده است. این سامانه در ابتدا، علایق فرد را در سیستم جهت تشکیل پروفایل ذخیره می‌کند. در ادامه، با استفاده از امتیازات کاربر به بازی و اطلاعات بازی از طرف سازنده آن که در سامانه‌های مختلف از جمله گوگل‌پلی موجود است، تا حد امکان با الگوریتم‌های پیاده‌سازی شده، بهترین بازی را به کاربران پیشنهاد می‌دهد.

بخش‌های این مقاله شامل ضرورت انجام پروژه (فصل 2)، پیشینه و معرفی سیستم توصیه‌گر (فصل 3)، استخراج داده‌ها از گوگل‌پلی، تمیزسازی و مرتب‌سازی داده‌ها برحسب نیاز، مدل سازی یک سامانه توصیه‌گر (فصل 4) نتیجه‌گیری (فصل 5) است که در ادامه به تفصیل بیان شده‌است.

**کلمات کليدي: سیستم توصیه‌گر، بازی، داده**

1. **مقدمه:**

امروزه پیشرفت تکنولوژی در زمینه‌های هوش‌مصنوعی، یادگیری‌ماشین و تحلیل دیتا باعث شده‌است سایت‌های مختلفی که به ارائه کالا، داده یا سرگرمی به کاربران مشغول هستند، به این رویه روی آوردند، که خدمات خود را با توجه به نیاز و علاقه‌مندی‌ها به کاربران معرفی کنند. به این معنی که فیلتری خودکار و همیشگی با توجه به سابقه فعالیت کاربران و علاقه‌مندی آن‌ها بر روی خدماتشان اعمال کنند. به این سیستم، سیستم‌توصیه‌گر گفته می‌شود.

سیستم‌های توصیه‌گر[[1]](#footnote-1)، زیر مجموعه‌ای از سیستم‌های فیلتر اطلاعات [[2]](#footnote-2) محسوب می‌شوند. هدف سیستم‌های توصیه‌گر، پیشنهاد مناسب‌ترین آیتم‌ها از جمله داده، کالا و سرگرمی به کاربران است. در سیستم‌های توصیه‌گر، داده‌های مرتبط با رفتار کاربران در خرید کالا، دریافت داده‌ها، تماشای سرگرمی و سایر موارد تحلیل و مدل‌سازی می‌شود. سپس با استفاده از مدل تولید شده، مناسب‌ترین آیتم به کاربران پیشنهاد می‌گردد. از بیشترین استفاده سیستم‌های توصیه‌گر، می‌توان در پیاده‌سازی کاربردهای تجاری به‌ویژه، در حوزه محصولات و خدمات مصرفی کاربران نام برد.

سیستم‌های توصیه‌گر به طور کلی در سه سطح قابل پیاده‌سازی هستند:

* مدل‌‌های محتوا محور[[3]](#footnote-3)
* مدل‌های مبتنی بر فیلتر گروهی [[4]](#footnote-4)
* مدل‌های ترکیبی [[5]](#footnote-5)

مدل محتوا‌ محور بر اساس پروفایل هر شخص به صورت مجزا از بقیه کاربران و بر اساس شباهت بین داده‌ها کار می‌کند. مدل مبتنی بر فیلتر گروهی بر اساس شباهت بین پروفایل کاربران عمل می‌کند، به این صورت که اگر کاربر 1 به داده 1 و 2 و 3 علاقه‌مند باشد و کاربر 2 به داده 1 و 3 علاقه‌‌مند باشد در نتیجه بر اساس شباهت پروفایل‌ها داده 2 به کاربر 2 پیشنهاد داده ‌می‌شود. و مدل ترکیبی به تلفیق هر دو مدل محتوا محور و مبتنی‌بر فیلتر گروهی گفته می‌شود. مدل محتوا محور در ابتدای فعالیت یک سامانه بهترین پیشنهاد است، زیرا نیاز به تعداد کاربر اولیه زیادی ندارد و برای هر کاربر فقط توسط پروفایل خودش پیشنهادات ارائه می‌گردد. در صورتی‌که برای استفاده بهینه از دو مدل دیگر نیازمند تعداد مناسبی کاربر در سامانه هستیم. به‌همین دلیل در این سامانه مدل محتوا محور پیاده شده‌است.

1. **ضرورت انجام پروژه:**

بر اساس تحقیقات انجام شده، اکثریت مردم در سنین مختلف در اوقات فراغت به کار با گوشی‌های هوشمند خود مشغول هستند و اغلب آن‌ها ازجمله کودکان، نوجوانان، بزرگسالان و حتی سالمندان برای سرگرمی به انجام بازی‌های موبایلی روی می‌آوردند. یکی از دغدغه‌های اصلی هر گروه سنی انتخاب سبک بازی و به طور کلی بازی‌ مناسب سلایق و روحیه و سن آن‌ها‌ و پیدا کردن بازی‌های مشابه‌ آن است.

دغدغه‌‌ی عده‌ای از والدین این است که می‌خواهند فرزندشان از بازی‌های خشن دور باشد یا در هنگام بازی به یادگیری نیز بپردازد. همچنین گروهی از سالمندان یا میانسالان برای سرگرمی به بازی‌های پازلی یا تقویت حافظه یا ... علاقه دارند. به طور کلی هر گروه سنی علاقه‌مندی و هدف خود را از بازی کردن دارد و به همین دلیل نیازمند سامانه‌ای جامع و با فیلتر مناسب‌تری از سامانه‌های موجود هستیم.

هدف اصلی این سامانه پیاده‌سازی یک سیستم توصیه‌گر مبتنی‌بر داده است. این سیستم در ابتدا، علایق فرد را در سیستم جهت تشکیل پروفایل ذخیره می‌کند. در ادامه، با استفاده از امتیازات کاربر به بازی و اطلاعات بازی از طرف سازنده آن که در سامانه‌های مختلف از جمله گوگل‌پلی موجود است، تا حد امکان با الگوریتم‌های پیاده‌سازی شده، بهترین بازی را به کاربران پیشنهاد می‌دهد.

1. **پیشینه:**

یک سیستم پیشنهاددهنده[[6]](#footnote-6) مانند یک سیستم رایانه‌ای عمل می‌کند که در مواردی که کاربر مایل به خرید داده، سرگرمی، کالا یا غیره باشد با الگوریتم‌های مختلفی به آن‌ها مشاوره و پیشنهاد می‌دهد. همچنین در این سیستم، این قابلیت وجود دارد که پروفایل هر کاربر شخصی‌سازی شده و پیشنهادات متفاوتی با توجه به سلایق خود دریافت کند. به‌این‌منظور، برای ارائه پیشنهادات شخصی‌سازی شده، یک سیستم توصیه‌گر، خواسته‌ها و علایق کاربر را مدل کرده و سپس با استفاده از مدل تشکیل‌شده برای کاربر، بهترین پیشنهادات به او داده می‌شود.

روند کاری یک سیستم توصیه‌گر، به سه مرحله زیر تقسیم می‌گردد:

* کسب تجربیات از داده‌های ورودی کاربر
* تشخیص پیشنهاد[[7]](#footnote-7) مناسب با استفاده از تکنیک‌های موجود
* ارائه نتایج پیشنهادی به کاربر

هدف سیستم پیشنهاددهنده ایجاد توصیه‌های معنی‌دار برای مجموعه‌ای از کاربران به‌جهت استفاده از خدمات یا کالاهاییست که ممکن‌است مورد علاقه آن‌ها باشد. برای مثال، پیشنهادات کتاب در آمازون ، یا فیلم در Netflix ، نمونه‌هایی واقعی از عملکرد سیستم‌های توصیه‌کننده پرقدرت در صنعت هستند.

طراحی سامانه‌ای به این عنوان، به مشخصات داده‌های موجود برای سامانه بستگی دارد. برای مثال اکثر این سیستم‌ها از رتبه‌بندی در مقیاس 1 تا 5 بر اساس علاقه استفاده می‌کنند.]1[

این سیستم با استفاده از معیارها و رتبه‌هایی که کاربران به کالاها یا خدمات می‌دهند پیشنهادات خود را ارائه می‌دهد. در صورتی‌که کاربر یا کالایی (با توجه به الگوریتم‌های مختلف) رتبه‌ای نداشته باشند، پیشنهادات مناسبی برای آن‌ها نیز از طرف سیستم داده نمی‌شود و برای حل این مشکل باید الگوریتم‌ها را با یکدیگر تلفیق کرده و یا از مقادیر پیش‌فرض استفاده کرد.

انواع سیستم‌های توصیه‌گر را می‌توان به سه دسته اصلی تقسیم کرد:

* محتوامحور
* مبتنی‌بر فیلتر گروهی
* سیستم ترکیبی

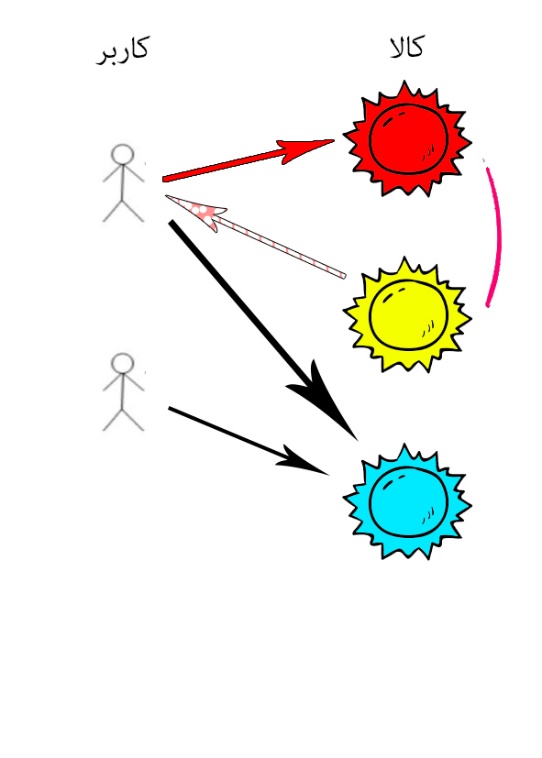
در زیربخش‌های بعدی توضیحات هر یک از این مدل‌ها به تفضیل بیان شده و لازم به‌ ذکر است به دلیل نداشتن کاربر در هنگام پیاده‌سازی این سامانه، از روش محتوا محور استفاده شده‌است‌.

* 1. **سیستم محتوا محور[[8]](#footnote-8):**

این سیستم مبتنی‌بر فعالیت‌ها و پروفایل هر کاربر و به صورت شخصی است. در این سیستم فعالیت‌های کاربران بر روی یکدیگر تاثیری ندارد و پیشنهادات ارائه‌شده صرفاً با توجه به علایق هر فرد و امتیازات داده شده توسط فرد به هر داده است، یعنی بر اساس شباهت داده‌ها کار می‌کند.

مشکل عمده این سیستم، زمانی‌ست که کاربر به تازگی در سامانه ثبت‌نام کرده و اطلاعاتی از فعالیت‌ها و علایق او در دسترس نیست. در این زمان به دو شیوه می‌توان عمل کرد؛ روش اول به این صورت است که سیستم فعالیت‌های کاربر را در سامانه دنبال کرده و علایق او را یاد می‌گیرد و به مرور پیشنهادات قوی‌تری را برای کاربر ارائه می‌دهد. در روش دوم در ابتدای ورود، کاربر لیستی از علاقه‌مندی‌های ابتدایی خود را وارد می‌کند که این روش روند یادگیری سامانه را ارتقا می‌بخشد.

روش کار این سیستم به این صورت است که با توجه به رتبه‌بندی داده‌شده به هر داده و شباهت آن داده با دیگر داده‌*‌*‌ها جدولی تشکیل می‌شود. سپس بالاترین امیتازات جدول که مربوط به قوی‌ترین پیشنهادات است به کاربر ارائه می‌گردد.]2[



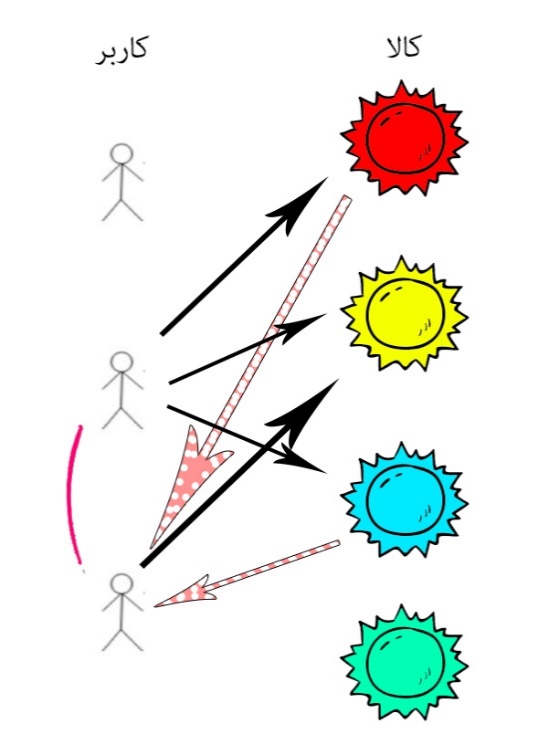
**عکس 1- مدل محتوامحور**

* 1. **سیستم مبتنی‌بر فیلتر گروهی[[9]](#footnote-9):**

یکی از زیرمجوعه‌های این سیستم، مدل مبتی‌بر فعالیت کاربران است. پیشنهادها در سیستم مبتنی‌بر فیلتر گروهی برخلاف روش قبلی که بر اساس تشابه ویژگی کالاهای پیشنهادی با ویژگی‌های کالاهای موردعلاقه کاربر انجام می‌پذیرفت، بر اساس تشابه رفتاری کاربر فعال با دیگر کاربران صورت می‌گیرد. مشکل این روش این است که در صورتی که داده‌ای تازه به سیستم اضافه شده و کاربری به آن علاقه‌ای نشان نداده‌ باشد به کسی پیشنهاد نمی‌شود که این مشکل در رویکرد محتوا محور وجود ندارد. شکل 2 روش مبتنی بر فعالیت‌های کاربران را نمایش می‌دهد.]3[

* 1. **سیستم ترکیبی:**

این سیستم به ترکیب هر دو سیستم بالا گفته می‌شود و از پربازده‌ترین سیستم‌ها به شمار می‌رود.

****

**عکس 2- مدل بر پایه فعالیت‌های کاربران**

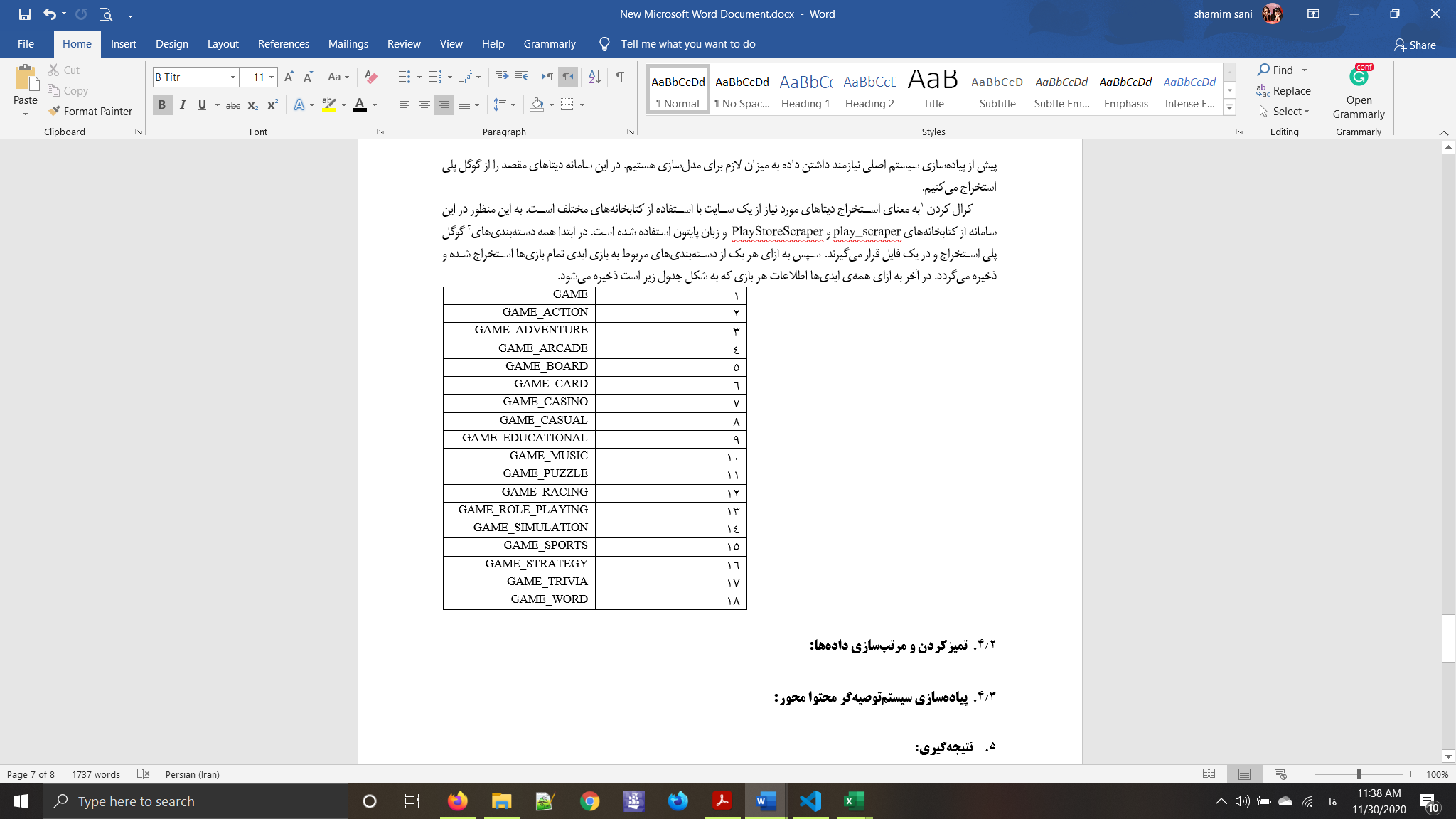
1. **پیاده‌سازی:**

به دلیل نداشتن تعداد یوزر اولیه مناسب در ابتدای کار، پیاده‌سازی این سیستم از روش مبتی‌بر آیتم‌ها انجام گرفته‌است. در ادامه مراحل پیاده‌سازی به‌ترتیب بیان شده‌است.

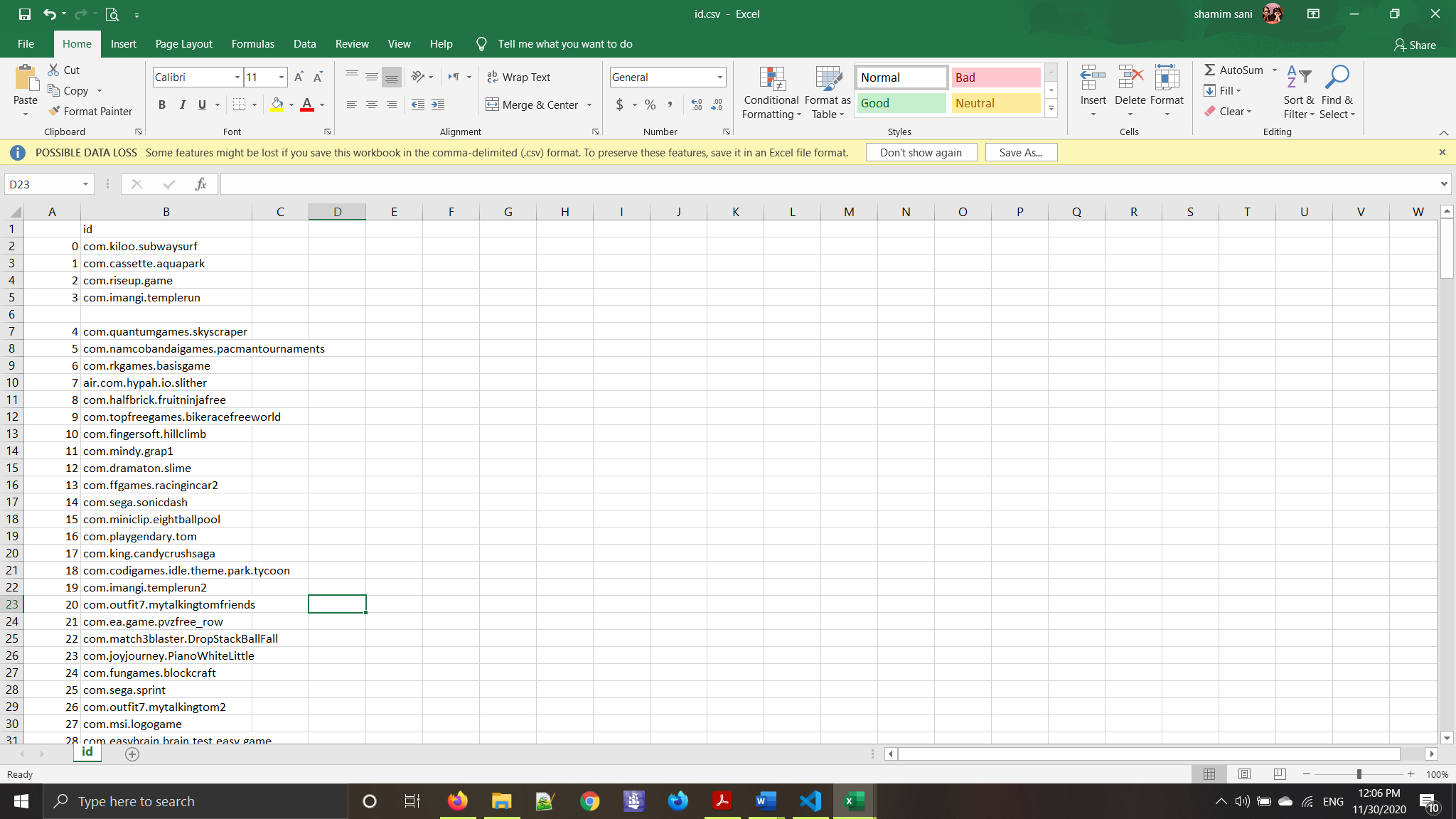
* 1. **استخراج داده‌ها:**

پیش از پیاده‌سازی سیستم اصلی، نیازمند داشتن داده به میزان لازم برای مدل‌سازی سیستم هستیم. در این سامانه دیتاهای مقصد را از گوگل پلی استخراج می‌کنیم.

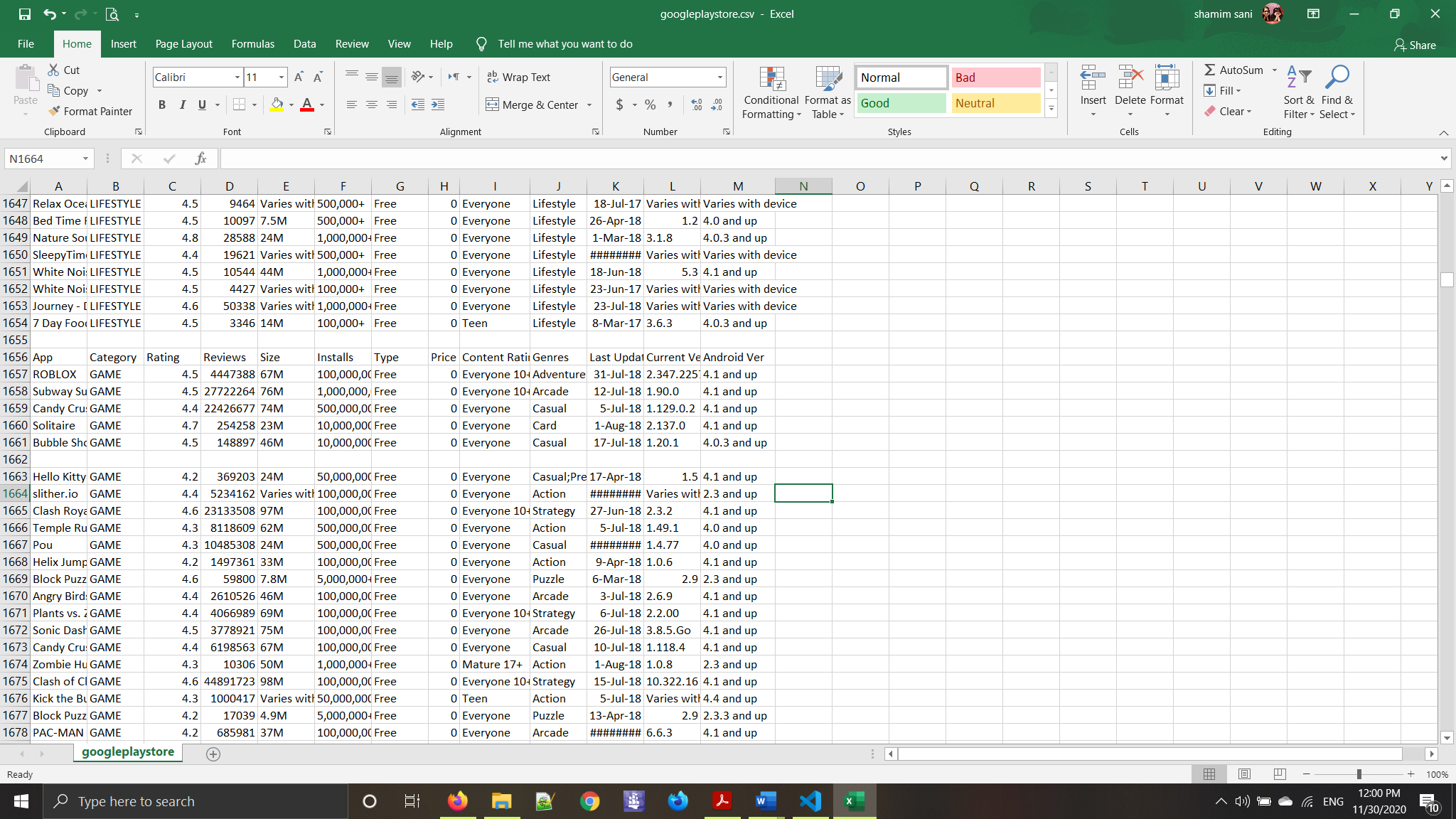
کرال کردن [[10]](#footnote-10)به معنای استخراج دیتاهای موردنیاز از یک سایت با استفاده از کتابخانه‌های مختلف تعبیه شده برای این کار می‌باشد. به این منظور در این سامانه از کتابخانه‌های play\_scraper و PlayStoreScraper در زبان پایتون استفاده شده‌است. در ابتدا همه دسته‌بندی‌های[[11]](#footnote-11) گوگل پلی استخراج و در یک فایل قرار می‌گیرند. سپس به ازای هر یک از دسته‌بندی‌های مربوط به بازی شناسه تمام بازی‌های مرتبط با آن شناسه استخراج شده و ذخیره می‌گردد. در آخر به ازای همه‌ی شناسه‌ها اطلاعات هر بازی که به شکل جدول 3 است ذخیره می‌شود.



**جدول 1- دسته‌بندی**



**جدول 2- نمونه شناسه بازی‌ها**

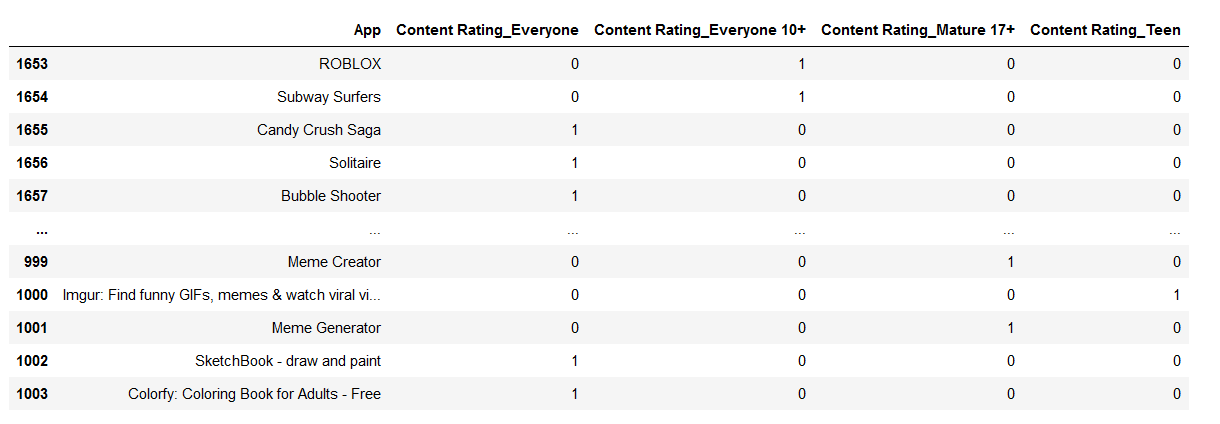


**جدول 3- جزئیات داده‌ها**

* 1. **تمیزکردن و مرتب‌سازی داده‌ها:**

برای پیاده‌سازی این سامانه و برای ساده‌تر شدن کار به جای استفاده از داده‌های استخراج شده از دیتاست آماده‌ای که از گوگل‌پلی استخراج شده، استفاده گردیده ‌است. این کار به‌دلیل صرفه‌جویی در زمان و حافظه سیستم برای ذخیره‌سازی و استخراج تعدادی زیادی انجام می‌شود[[12]](#footnote-12). پیش از مدل‌سازی دیتاها لازم است داده‌های ناموجود[[13]](#footnote-13) و صفر حذف یا اصلاح شوند. پس از حذف و اصلاح داده‌های ناقص، ستون‌هایی که برای ما در این سامانه مورد نیاز نیستند نیز حذف[[14]](#footnote-14) می‌گردند. در ادامه ویژگی‌های مورد نیاز در این سامانه مانند امتیازات، رده‌بندی سنی، دسته‌بندی و مدل پرداختی بازی و ... رمزگذاری[[15]](#footnote-15) می‌شوند. رمزگذاری به شکل زیر صورت می‌گیرد:

* در ابتدا تعداد سطرهای متمایز ستون موردنظر مشخص می‌شود.
* به ازای تعداد سطرهای موردنظر ستون به جدول اضافه می‌گردد.
* در هر سطر به‌ازای هر ویژگی مورد‌نظر در ستون اضافه شده با همان ویژگی، 1 قرار داده و در غیر این‌صورت 0 قرار می‌دهیم.
* ستون اصلی از جدول حذف می‌شود.



**شکل 3- Content Rating Encoding**

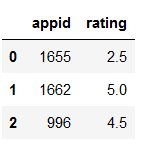
پس از اعمال ‌رمزگذاری ستون‌های مورد نیاز، باقی ستون‌ها از جدول حذف شده و مرحله مدل‌سازی آغاز می‌شود.]2[

* 1. **پیاده‌سازی سیستم‌توصیه‌گر محتوا محور:**

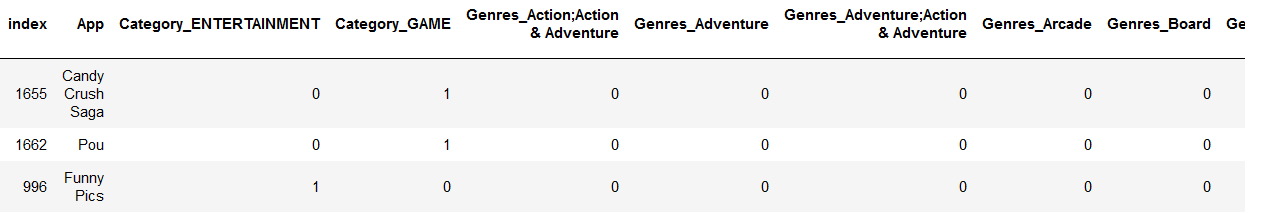
همان‌طور که در بخش‌های پیشین گفته‌شد، مدل پیاده‌سازی در این سامانه مبتنی‌بر محتوا می‌باشد. به‌این‌صورت که به جای ارتباط بین کاربران از ارتباط بین داده‌ها برای پیش‌بینی استفاده می‌گردد.

این پروسه طی نه مرحله به صورت زیر انجام می‌شود:

1. در ابتدا یک جدول به نام Input user rating برای نمایش امتیازات کاربر به هر آیتم تشکیل می‌دهیم.
2. جدول رمزگذاری شده برای آیتم‌ها را تشکیل می‌دهیم.
3. با استفاده از فیلتر، جدولی برای هر کاربر شامل آیدی اپلیکیشن موردنظر و ویژگی‌های رمزگذاری شده آن تشکیل می‌دهیم.

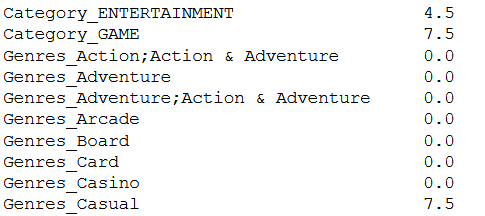


**جدول 4- Input user rating**



**جدول 5- جدول رمزگذاری شده هر کاربر**

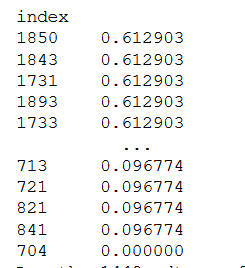
1. جداول مرحله 1 و 3 را در هم ضرب کرده و وزن هر ویژگی را به دست می‌آوریم.
2. مجموع هر سطر را به‌دست آورده و سپس با استفاده از آن، مجموع وزن هر ویژگی را برای هر کاربر محاسبه می‌کنیم و به اصطلاح در این مرحله پروفایل کاربر[[16]](#footnote-16) را تشکیل می‌دهیم.



**جدول 6- پروفایل کاربر**

1. جدول نهایی در مرحله 5 را نرمال‌سازی می‌کنیم. دلیل نرمال‌سازی داده‌ها این است که رنج امتیازات همه بین 0 تا 1 می‌شود و روند پیشنهاد با دقت بیشتری انجام می‌پذیرد.
2. پروفایل کاربر را در جدول مرحله 2 که شامل همه‌ی برنامه‌ها به صورت رمزگذاری شده‌است، ضرب کرده و سپس ماتریس وزنی[[17]](#footnote-17) برای آیتم‌های جدید را تشکیل می‌دهیم.
3. مجموع امتیازات هر سطر را به‌دست آورده و یک ماتریس توصیه‌گر برای پیشنهادات نهایی می‌سازیم.
4. ماتریس نهایی را بر اساس امتیاز مرتب می‌کنیم. بالاترین امتیازات در این جدول بهترین پیشنهاد برای هر کاربر می‌باشد.

به این ترتیب سامانه توصیه‌گر براساس محتوی موردعلاقه هر کاربر تشکیل می‌شود.]4[



**جدول 7- نتایج پیشنهادی**

1. **نتیجه‌گیری:**

سیستم‌های توصیه‌گر یکی از مهمترین ابزار‌های ایجاد ارزش افزوده در کسب‌‌و‌کار‌ها به حساب می‌آیند. این سیستم‌ها می‌توانند به طور موثر سبک‌های رتبه‌بندی را در شباهت محاسباتی برای بهبود پیش‌بینی اعمال کنند. سیستم‌های مبتنی‌بر داده‌، مانند سیستم فوق نسبت به سیستم‌های کاربرمحور کیفیت بهتری ارائه می‌دهند. داده‌ها و پیشرفت‌های ذکر شده پایان توسعه این سیستم نیست، این سیستم در آینده می‌تواند بر پایه پردازش قابل حمل نیز اجرا گردد؛ ولی در حال حاضر این بخش فقط در چهارچوب مفهومی قابل اجرا است.

تمامی کدهای پیاده‌سازی شده در آدرس اینترنتی <https://github.com/smimahs/GooglePlayRecommendedSystem> قرار دارد.

1. **مراجع:**

*[1]* ***H., Zaw. L. 2016.*** *Online Car Recommendation Using Collaborative Filtering-Based Agoregated Query (CFAg Query). May 2016.*

*[2]* ***SAEED AGHABOZORGI, Joseph Santarcangelo. 2019.*** *Machine Learning with Python. s.l. : IBM, 2019*

*[3]* ***Ng, Andrew. 2016.*** *Machine learning. s.l. : Stanford University, 2016.*

*[4]* ***Khin Mar Cho, Mya Sandar Kyin. 2020.*** *Mobile game applications Recommendation System with item-based Collaborative Filtering. 2020.*

1. Recommender, Recommendation Systems [↑](#footnote-ref-1)
2. Information Filtering Systems [↑](#footnote-ref-2)
3. content based [↑](#footnote-ref-3)
4. collaborative filtering [↑](#footnote-ref-4)
5. hybrid matrix factorization [↑](#footnote-ref-5)
6. Recommended system [↑](#footnote-ref-6)
7. Recommendation [↑](#footnote-ref-7)
8. Content based, Item based [↑](#footnote-ref-8)
9. User based, Collaborate filtering [↑](#footnote-ref-9)
10. crawling [↑](#footnote-ref-10)
11. Categories [↑](#footnote-ref-11)
12. https://www.kaggle.com/lava18/google-play-store-apps [↑](#footnote-ref-12)
13. Null data [↑](#footnote-ref-13)
14. drop [↑](#footnote-ref-14)
15. encode [↑](#footnote-ref-15)
16. User profile [↑](#footnote-ref-16)
17. Weighted matrix [↑](#footnote-ref-17)