

## An Algorithm to Measure the Attention Rate of Mobile Application Users using App Reviews

Mehrdad Razavi Dehkordi<sup>1</sup>, Hamid Rastegari<sup>2\*</sup>, Akbar Nabiollahi Najafabadi<sup>3</sup>, Taghi Javdani Gandomani<sup>4</sup>

1. Department of Computer Engineering, Najafabad Branch, Islamic Azad University, Najafabad, Iran. [mehrdad.razavi68@sco.iaun.ac.ir](mailto:mehrdad.razavi68@sco.iaun.ac.ir)
2. Assistant Professor, Department of Computer Engineering, Najafabad Branch, Islamic Azad University, Najafabad, Iran. \*Corresponding Author, [rastegari@iaun.ac.ir](mailto:rastegari@iaun.ac.ir)
3. Assistant Professor, Department of Computer Engineering, Najafabad Branch, Islamic Azad University, Najafabad, Iran. [a.nabi@pco.iaun.ac.ir](mailto:a.nabi@pco.iaun.ac.ir)
4. Assistant Professor, Department of Computer Science, Shahrekord University, Shahrekord, Iran. [javdani@sku.ac.ir](mailto:javdani@sku.ac.ir)

### Abstract

**Introduction:** Since the emergence of mobile apps, user reviews have been of great importance for app developers. as they contain users' sentiment, bugs and new requests, user feedback has been valuable to app developers. Due to the large number of reviews, analysing them is a difficult, time-consuming and error-prone task. So far, many works have been done in the field of classifying reviews but in none of them, it is not clear which category the users paid more attention to after categorization.

**Method:** In this article, an algorithm called UAR(User Attention Rate) is presented based on the date of reviews and their number to measure the level of attention of users to the category of bug report or feature request. To test the proposed algorithm, the reviews of WhatsApp, NextCloud, Mozilla Firefox and VLC Media Player apps were extracted by the web crawler for 30 days, and then the reviews were given to 4 software experts after classification. In the following, after analyzing the reviews by 4 experts, the level of attention of users to the reviews was calculated by them. Then, using the proposed method, amount of users' attention to reviews was measured.

**Results:** It was found that the proposed algorithm was able to have acceptable results in calculating the amount of users' attention to reviews compared to the values calculated by experts. So that the difference in the attention rate calculated by experts according to the proposed algorithm was 72.2%.

**Discussion:** In this article, an algorithm is presented to calculate the users' attention rate to the categories of feature requests and bug reports. By using the relevant algorithm, the development team will understand that the user wants to fix the bugs related to the application or add new features to the application.

**Keywords:** Mobile application, user attention rate, google play store analysis, user reviews.



## ارائه الگوریتمی جهت اندازه‌گیری میزان توجه کاربران برنامه‌های موبایل با استفاده از نظرات مربوط به برنامه‌ها

دوره پنجم، تابستان ۱۴۰۳  
شماره دوم، صص: ۲۳-۳۱

تاریخ دریافت: ۱۸/۰۱/۱۴۰۳  
تاریخ پذیرش: ۰۳/۰۳/۱۴۰۳

مهرداد رضوی دهکردی<sup>۱</sup>، حمید رستگاری<sup>۲\*</sup>، اکبر نبی‌اللهی نجف‌آبادی<sup>۳</sup>، تقی جاودانی‌گندمانی<sup>۴</sup>

۱. گروه مهندسی کامپیوتر، واحد نجف‌آباد، دانشگاه آزاد اسلامی، نجف‌آباد، ایران. [mehرداد.razavi68@sco.iaun.ac.ir](mailto:mehرداد.razavi68@sco.iaun.ac.ir)

۲. استادیار، گروه مهندسی کامپیوتر، واحد نجف‌آباد، دانشگاه آزاد اسلامی، نجف‌آباد، ایران. (نویسنده مسئول) [rastegari@iaun.ac.ir](mailto:rastegari@iaun.ac.ir)

۳. استادیار، گروه مهندسی کامپیوتر، واحد نجف‌آباد، دانشگاه آزاد اسلامی، نجف‌آباد، ایران. [a.nabi@pco.iaun.ac.ir](mailto:a.nabi@pco.iaun.ac.ir)

۴. استادیار، گروه علوم کامپیوتر، دانشگاه شهرکرد، شهرکرد، ایران. [javdani@sku.ac.ir](mailto:javdani@sku.ac.ir)

**چکیده:** از زمان پیدایش برنامه‌های موبایل، نظرات کاربران برای توسعه‌دهندگان برنامه بسیار ارزشمند بوده‌است. به دلیل حجم بالای نظرات، تحلیل آن‌ها کار بسیار مشکلی است. تاکنون کارهای بسیاری در زمینه دسته‌بندی نظرات انجام شده‌است، اما در هیچ کدام از آن‌ها مشخص نشده‌است که پس از دسته‌بندی، توجه کاربران بیشتر به کدام دسته بوده‌است. در این مقاله الگوریتمی بانام UAR بر اساس تاریخ ثبت نظرات و تعداد آن‌ها برای اندازه‌گیری میزان توجه کاربران به دسته گزارش مشکل یا درخواست ویژگی ارائه شده‌است. برای آزمایش الگوریتم مورد نظر، نظرات برنامه‌های WhatsApp، NextCloud، Mozilla Firefox و VLC Media Player توسط خزنده وب به مدت ۳۰ روز استخراج شده و نظرات پس از دسته‌بندی به ۴ متخصص نرم‌افزار داده‌شد. در ادامه پس از تحلیل نظرات توسط ۴ متخصص میزان توجه کاربران به نظرات توسط آن‌ها مورد محاسبه قرار گرفت. سپس با استفاده از روش پیشنهادی میزان توجه کاربران به نظرات اندازه‌گیری شد. در نهایت مشخص شد که روش پیشنهادی توانسته نتایج قابل قبولی در محاسبه میزان توجه کاربران به نظرات نسبت به مقادیر محاسبه شده توسط متخصصان داشته‌باشد طوری که تفاوت میزان توجه محاسبه شده توسط متخصصان با میزان توجه محاسبه شده توسط الگوریتم ارائه شده ۷۲/۲٪ بوده‌است

**واژه‌های کلیدی:** برنامه‌های موبایل، میزان توجه، آنالیز فروشگاه گوگل، میزان توجه کاربران.

## ۱. مقدمه

کاربران برای هر کدام از برنامه‌ها توسط متخصصین به‌طور میانگین اندازه‌گیری شده و با مقدار محاسبه‌شده توسط الگوریتم پیشنهادی مقایسه و در نهایت مشخص شده است که الگوریتم پیشنهادی نتایج قابل‌قبولی در محاسبه نرخ توجه کاربران به دسته‌های نظرات دارد. ادامه این مقاله به صورت زیر سامان‌دهی شده است: در بخش دوم کارهای انجام‌شده در دسته‌بندی نظرات و نیازمندی‌ها بررسی شده و به چالش‌های آن‌ها پرداخته شده است، در بخش سوم الگوریتم پیشنهادی برای محاسبه میزان توجه کاربران به دسته‌های گزارش مشکل یا درخواست ویژگی آورده شده است، در بخش چهارم در مورد روش انجام آزمایش‌ها و معیارهای مقایسه صحبت شده، در بخش پنجم نتایج آزمایش‌ها عنوان شده و در نهایت در بخش ششم نتیجه‌گیری و پیشنهادهایی برای کارهای آینده آمده است.

## ۲. کارهای مرتبط و چالش‌های آن‌ها

در دسته‌بندی نظرات عنوان‌شده توسط کاربران برنامه‌های موبایل کارهای زیادی انجام شده است که از مهم‌ترین آن‌ها می‌توان به موارد زیر اشاره کرد:

چن و همکاران در سال ۲۰۱۴، ابزاری با نام AR-Miner برای تشخیص ندهای دارای اطلاعات ضروری و غیرضروری برای تیم توسعه‌دهنده ارائه دادند، مدل ارائه‌شده برای آموزش، از نظرات مربوط به دو برنامه SwiftKey و Facebook و دو بازی Temple Run و Tapfish، با استفاده از الگوریتم EMNB و Top Modeling استفاده کرد و توانست با Hit rate با مقدار ۰/۷ نظرات سودمند را از نظرات غیر سودمند مجزا سازد [9].

مالج و نبیل در سال ۲۰۱۶، مطالعه‌ای را برای شناخت بهترین روش برای دسته‌بندی نظرات کاربران در فروشگاه Play بر اساس دسته‌های درخواست ویژگی، گزارش مشکل، تجربه کاربران و امتیازدهی انجام دادند. آن‌ها پایگاه داده‌ای از ۱۱۰۰ برنامه با ۱/۱ میلیون نظر ساختند و در نهایت تئوری بیز را به‌عنوان بهترین روش دسته‌بندی با precision با مقدار ۰/۸۲ معرفی کردند [12].

مقدم در سال ۲۰۱۵، با بررسی ۵۰۰۰۰ نظر از فروشگاه برنامه‌های ebay، توانست الگوهای موجود در نظرات کاربران را با استفاده از برچسب زدن گفتار شناسایی کند. سپس با استفاده از تحلیل احساسات و روش بردار پشتیبان و روش کاهش ویژگی LDA برای دسته‌بندی، نظرات کاربران را به دسته‌های درخواست ویژگی، گزارش اشکال یا گزارش کار با برنامه دسته‌بندی کرد. در نهایت مشخص شد که روش مبتنی بر الگوها توانسته در دسته‌بندی نظرات Precision با مقدار ۸۸٪ داشته باشد [13].

د-جوهان و همکاران، ابزاری با نام SAFE برای استخراج ویژگی‌های نهفته در نظرات کاربران در فروشگاه Play ارائه دادند. آن‌ها به روش دستی، ۱۸ الگوی POS و ۵ الگوی مربوط به جملات را از نظرات استخراج کردند. سپس برای آزمایش از نظرات مربوط به ۱۰

امروزه طبق آمارهای رسمی سایت IDC تا نیمه چهارم سال ۲۰۲۲ حدود ۳۰۰/۳ میلیون تلفن همراه هوشمند تولید شده است که ۷۵/۸٪ از آن به گوشی‌های هوشمند با سیستم‌عامل اندروید تعلق دارد. سیستم‌عامل اندروید برای خود فروشگاه‌های با نام Google PlayStore دارد که تمامی توسعه‌دهندگان نرم‌افزارهای تولیدشده خود را بر روی آن قرار می‌دهند [1][2]. برنامه‌های موجود در فروشگاه توسط کاربران بسیار دانلود شده و کاربران فروشگاه گوگل امکان ثبت نظر برای برنامه مورد نظر خود را دارند. مطالعات نشان داده است که نظرات کاربران حاوی اطلاعات مهمی شامل گزارش مشکل، درخواست ویژگی و تجربه کاربران در کار با برنامه است [3][4][5]. مطالعات پیشین نشان می‌دهد که نظرات ثبت‌شده توسط کاربران می‌توانند فرآیند توسعه نرم‌افزار را هدایت کرده و باعث بهبود نسخه‌های آینده نرم‌افزار شوند [6][7]. به‌علاوه نظرات شامل اطلاعات بسیار مهمی برای آنالیزکننده‌ها و طراحان برنامه نیز هستند [4][8][5]. به‌علاوه اولویت‌بندی و دسته‌بندی به‌موقع نظرات نیز باعث می‌شود تا تیم توسعه‌دهنده در نگهداری نرم‌افزار و به‌روزرسانی‌های به‌موقع موفق‌تر باشد [9][10][11].

در زمینه دسته‌بندی نظرات تاکنون کارهای بسیاری انجام شده است؛ اما در هیچ‌کدام از این روش‌ها یا مدل‌ها بیان نشده که پس از دسته‌بندی یا اولویت‌بندی نظرات کاربران، توجه کاربران بیشتر به دسته گزارش مشکل بوده است یا درخواست ویژگی. به بیان دیگر می‌توان گفت در هیچ‌کدام از روش‌ها، گزارش نشده که پس از اولویت‌بندی یا دسته‌بندی نظرات، کاربر برنامه از توسعه‌دهنده خواسته تا بیشتر مشکلات برنامه را رفع کند یا بیشتر درخواست اضافه‌کردن ویژگی‌های جدید را به توسعه‌دهنده داده است؛ به عبارت دیگر توجه کاربر برنامه بیشتر به سمت گزارش مشکلات بوده است یا درخواست ویژگی‌های جدید. باید توجه داشت که تاکنون هیچ کاری در زمینه اندازه‌گیری میزان توجه کاربران انجام نشده است. با استفاده از الگوریتم پیشنهادی تیم توسعه‌دهنده خواهد توانست میزان توجه کاربر به دسته‌های نظرات را اندازه‌گیری کرده و در به‌روزرسانی بعدی سعی در پاسخ‌دادن به آن‌ها داشته باشد. اگر اشکالات گزارش شده توسط کاربران یا ویژگی‌های مورد نیاز آن‌ها به‌موقع پاسخ داده شوند، برنامه نه تنها موفق‌تر است بلکه نزد کاربر نیز دارای محبوبیت بیشتری خواهد بود [10][11].

هدف از این تحقیق ارائه الگوریتمی جهت محاسبه میزان توجه کاربران برنامه‌های موبایل به دسته‌های درخواست ویژگی یا گزارش مشکل با استفاده از تاریخ ثبت نظرات و تعداد آن‌ها است. برای مشاهده و اندازه‌گیری کارایی روش پیشنهادی با استفاده از خزنده وب، نظرات مربوط به چهار برنامه WhatsApp، Next Cloud، Mozilla Firefox و VLC Media Player به مدت ۳۰ روز استخراج شده و نظرات پس از دسته‌بندی به دسته‌های درخواست ویژگی و گزارش مشکل به ۴ متخصص در حوزه نرم‌افزار بوده‌اند، داده‌شده. در ادامه میزان توجه

برنامه استفاده و مشخص شد ابزار ارائه شده می‌تواند ویژگی‌های موجود در نظرات کاربران را با Precision با مقدار ۰/۵۵ و recall با مقدار ۰/۴۳ تشخیص دهد [14].

اسکالابری و همکاران در سال ۲۰۱۹، از روش‌های یادگیری ماشین استفاده کرده و ابزاری با عنوان CLAP را برای دسته‌بندی، خوشه‌بندی نظرات و اولویت‌بندی خوشه‌های که باید در آینده توسط تیم توسعه‌دهنده پیاده‌سازی شوند، ارائه دادند. دسته‌های آن‌ها شامل گزارش مشکل عملیاتی، پیشنهاد برای ویژگی جدید و نیازمندی‌های غیرعملکردی بود. در نهایت آن‌ها با استفاده از الگوریتم‌های درخت تصادفی و DBScan و استفاده از پایگاه داده ابزار [9] Ar-Miner به precision با مقدار ۰/۸۷ در دسته‌بندی نظرات دست یافتند [15].

مارتنز و همکاران در سال ۲۰۱۹، با تحقیق بر روی نظرات جعلی در فروشگاه Apple App Store و مشورت با ۴۳ فرد تولیدکننده نظرات جعلی توانستند پایگاه داده‌ای شامل ۶۰۰۰۰ نظر جعلی و ۶۲ میلیون نظر تولیدکننده. سپس پایگاه داده جمع‌آوری‌شده را به ۷ ابزار دسته‌بندی دادند و در نهایت مشخص شد که ابزار درخت تصادفی توانسته با ۹۸٪ مقدار Recall نظرات جعلی را از نظرات واقعی دسته‌بندی کند [16].

اسلم و همکاران در سال ۲۰۲۰ با استفاده از پایگاه داده ارائه شده در مطالعه مالچ و همکاران در سال ۲۰۱۶ [12] و استخراج ۴۴۰۰ نظر از نظرات این پایگاه داده، سیستمی را با استفاده از شبکه‌های عصبی کونولوشنی (CNN) و یادگیری عمیق برای دسته‌بندی نظرات ارائه دادند. سیستم پیشنهادی توانست F-Measure با مقدار ۹۴/۷٪ در دسته‌بندی نظرات داشته باشد [17].

کیانو و وانگ در سال ۲۰۲۰، مدلی را بر پایه یادگیری عمیق با معماری LSTM برای دسته‌بندی نظرات کاربران در دسته بازی‌های موجود در فروشگاه Apple App Store ارائه دادند. مدل آن‌ها با بررسی 261,515,118 نظر برای بازی ساخته شد. سپس از ۲۶ دانشجوی کارشناسی برای برچسب زدن دستی نظرات به سه دسته درخواست ویژگی، گزارش مشکل و سایر استفاده و در نهایت مشخص شد که ابزار پیشنهادی با استفاده از یادگیری عمیق دارای F-Measure با مقدار ۰/۷۶۶، دارای Precision با مقدار ۰/۷۶۹ و دارای Recall با مقدار ۰/۷۶۸ در دسته‌بندی نظرات بوده است [18].

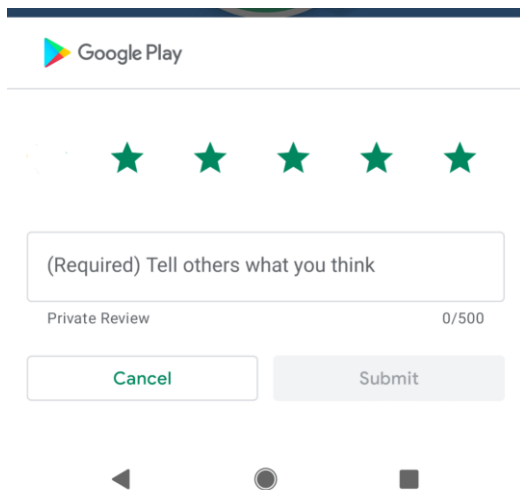
در جدول ۱ مهم‌ترین کارها در دسته‌بندی نظرات با هم مقایسه شده‌اند.

جدول ۱: مهم‌ترین کارها در زمینه دسته‌بندی نظرات کاربران

ارائه‌دهنده / سال نشر	دستاوردها	روش مورد استفاده	نتایج و معیار ارزیابی
۲۰۱۴ [9] چن و همکاران	ار AR-Miner	EMNB و Top Modeling LDA	مقدار ۰/۷ در دسته‌بندی نظرات سودمند از غیر سودمند
۲۰۱۵ [13] مقدم	شناسایی الگوهای موجود در نظرات کاربران	POS, تحلیل احساسات و SVM	مقدار ۰/۸۸ در شناسایی نظرات کاربران
۲۰۱۶ [14] مالچ و همکار	شناسایی بهترین الگوریتم در دسته‌بندی نظرات کاربران	ات	تئوری بیز با precision با مقدار ۰/۸۲ بهترین روش در دسته‌بندی نظرات است
۲۰۱۷ [15] جوهان و همکاران	ار SAFE	Pos Tag & Pattern Recognition	میزان ۰/۵۵ و recall با مقدار ۰/۴۳ در شناسایی ویژگی‌های مورد درخواست کاربران
۲۰۱۹ [15] اسکالابری و همکاران	ار CLAP	Random Forest و DBSCAN	مقدار ۰/۸۷ در دسته‌بندی اولویت‌بندی نظرات کاربران
۲۰۲۰ [17] اسلم و همکاران	اه Play	NLP و	میزان ۹۴/۷ در دسته‌بندی نظرات کاربران
۲۰۲۰ [18] کیانو و همکار	اه Apple App Store	CNN و LSTM	ار ۰/۷۶۶، دار ۰/۷۶۹ و رات

طبق جدول ۱ کارهای انجام‌شده دیگران دارای چالش‌ها و مشکلات زیر است:

- در ابزار ارائه‌شده امکان استخراج نظرات به‌روز برای برنامه، دسته‌بندی، خلاصه‌سازی، تعیین قطبیت و اولویت‌بندی آن‌ها در هر لحظه و با اراده تیم توسعه‌دهنده وجود ندارد؛ بلکه پایگاه داده‌ای ساخته‌شده سیستم با آن آموزش‌دیده و عمل دسته‌بندی، خوشه‌بندی و اولویت‌بندی روی آن پایگاه داده آزمایش‌شده است.
- برای یادگیری سیستم دسته‌بندی آن‌ها از روش‌های قدیمی یادگیری ماشین استفاده شده است.
- در دسته‌بندی و اولویت‌بندی از ویژگی‌هایی استفاده شده که اکنون دیگر توسط فروشگاه گوگل ارائه نمی‌شوند و اکنون امکان استفاده از این ویژگی‌های نمی‌باشد.



شکل ۲: نحوه ثبت امتیاز و نظر برای برنامه در فروشگاه Play

هنگام مراجعه به صفحه مربوط به برنامه مشاهده می‌شود که تعدادی از نظرات تاریخ ثبت قدیمی تری نسبت به نظرات جدید دارند اما در بالای نظرات جدید نمایش داده می‌شوند. به این نظرات، بسیار مرتبط (Most Relevant) گفته می‌شود. فروشگاه گوگل با استفاده از فاکتورهایی مانند الف- طول نظر ثبت شده، ب- اشاره نظر ثبت شده به یک ویژگی یا قسمت خاص برنامه، ج- نظر ثبت شده توسط منتقد خاص یا مورد تأیید Google، د- استفاده از کلمات خاص، ه- استفاده از عکس در ثبت نظرات، و- نظرات دارای احساس منفی، نظرات بسیار مرتبط را از دیگر نظرات متمایز کرده و به کاربران نشان می‌دهد [20][19]. ممکن است تعداد نظرات بسیار مرتبط نسبت به سایر نظرات کمتر باشند اما در بسیاری از موارد این نظرات دغدغه‌های دیگر کاربران استفاده‌کننده از برنامه هستند.

هنگام ثبت نظر تاریخ آن نیز توسط فروشگاه ثبت می‌شود. زمانی که با استفاده از خزنده وب نظرات مربوط به برنامه خاص استخراج می‌شود، تاریخ ثبت نظر و نوع نظر شامل جدید یا بسیار مرتبط بودن آن نیز توسط خزنده استخراج می‌شود.

در نتیجه کاربران هنگام ثبت نظر که شامل گزارش مشکل یا درخواست ویژگی یا شرح کارکرد در کار با برنامه است، در صورت مشاهده دغدغه خود که توسط کاربر دیگر مطرح شده، دیگر آن را ثبت نکرده و برای تصدیق و تأیید نظر دیگران از دکمه ThumbsUp با علامت استفاده می‌کنند. در بسیاری از موارد نیز ممکن است با پیام Did you find this Helpful یا دکمه‌های Yes و No مواجه شوند. در نتیجه تعدادی از نظرات در مورد برنامه دارای محبوبیت بسیار از طرف کاربران وجود دارند.

- برای ساخت پایگاه داده برای آموزش مدل ارائه شده، ویژگی‌ها به صورت دستی توسط عوامل انسانی از نظرات استخراج شده‌اند؛ اما امروزه با استفاده از یادگیری عمیق دیگر عمل استخراج ویژگی کاربرد آن چنانی نخواهد داشت.

به علاوه با توجه به مشکلات مورد اشاره موجود در روش‌های دیگران در هیچ‌کدام از ابزارها مشخص نشده است که پس از دسته‌بندی یا اولویت‌بندی نظرات، کاربران برنامه‌های موبایل از تیم توسعه‌دهنده بیشتر درخواست اضافه کردن ویژگی‌های جدید را داشته یا بیشتر مشکلات مربوط به برنامه را گزارش داده و خواستار رفع آن‌ها بوده‌اند.

با توجه به چالش‌های مذکور، ارائه روش یا الگوریتمی برای محاسبه میزان درخواست کاربران یا میزان توجه آن‌ها مورد نیاز است و پرسش زیر در تحقیق مطرح می‌شود:

آیا روش پیشنهادی می‌تواند به خوبی روش‌های دستی (روش بررسی نظرات توسط متخصصین) میزان توجه کاربران به دسته‌های گزارش مشکل یا درخواست ویژگی را محاسبه کند؟

### ۳. روش محاسبه میزان توجه کاربران

روزانه نظرات بسیاری برای برنامه‌ها ثبت می‌شود. هدف ما محاسبه میزان توجه کاربران به نظرات است. در این تحقیق بر پایه ویژگی‌هایی که در زیر بیان می‌شود الگوریتمی برای محاسبه نرخ توجه کاربران به دسته‌های گزارش مشکل و درخواست ویژگی با عنوان *UAR (User Attention Rate)* ارائه شده است. باید توجه داشت که نظرات برنامه به وسیله خزنده‌های تحت وب پیاده‌سازی شده با استفاده از زبان برنامه‌سازی پایتون با دادن نام Package برنامه در فروشگاه Play قابل-استخراج هستند. برای محاسبه نرخ توجه از ویژگی‌های تعداد نظرات و تاریخ ثبت نظرات استفاده شده است. در ادامه به بررسی هر کدام از ویژگی‌ها به‌طور جداگانه پرداخته شده است.

#### ۱.۳. تاریخ ثبت نظرات

هنگامی که کاربر در فروشگاه Play اقدام به ثبت نظر کند، گزینه‌ای برای امتیازدهی نمایش داده می‌شود که می‌تواند امتیازی بین ۱ تا ۵ برحسب رضایت به برنامه بدهد. به علاوه در زیر فیلد مربوط به امتیاز، محلی برای ثبت نظر تعبیه شده است (شکل ۱).

#### Rate this app

Tell others what you think



Write a review

#### شکل ۱: ثبت نظر و امتیاز مربوط به برنامه در فروشگاه Play

مطابق شکل ۱، هنگام کلیک بر روی گزینه Write a review پنجره‌ای برای نوشتن نظر به کاربر نشان داده می‌شود. در شکل ۲ نحوه امتیازدهی و ثبت نظر کاربر آمده است.

در ادامه از رابطه ۱ و ۲ برای ارائه الگوریتم محاسبه میزان توجه استفاده خواهد شد.

<b>Algorithm : Calculate Attention Rate</b>
<b>Input:</b> Sets of Reviews with their Metadata (R)
<b>Output:</b> $AR_r$ (Attention Rate R)
1. $R_d \leftarrow \text{Sort by Date Descending}(r)$
2. Calculate Dates from $R_d$
3. $AR_r \leftarrow \frac{ R }{\text{Dates}}$
4. Return $AR_r$
5. End

الگوریتم ۱: الگوریتم محاسبه نرخ توجه به دسته خاص از نظرات

همان‌طور که در الگوریتم ۱ مشاهده می‌شود، ورودی الگوریتم دسته نظرات به صورت خام (استخراج شده توسط خزنده) است. در مرحله ۱ نظرات به صورت نزولی بر اساس تاریخ مرتب می‌شوند تا به راحتی بتوان تفاوت تاریخ ثبت اولین نظر تا آخرین نظر دارای ThumbsUp یا محبوبیت بالا را مشخص کرد. در مرحله ۲ میزان تفاوت تاریخ محاسبه شده، در مرحله ۳ با استفاده از رابطه ۲ نرخ توجه کاربر به دسته نظرات محاسبه می‌شود. سپس در مرحله ۴ نرخ توجه به هر دو دسته بندی گزارش مشکل یا درخواست ویژگی به عنوان خروجی منتشر می‌شود. نرخ توجه عددی مثبت با مقدار حداقل صفر و به صورت اعشاری می‌باشد. در ادامه می‌توان برای دسته بندی شامل نظراتی که توجه کاربر به آن‌ها بیشتر است، اولویت بندی را انجام داد. با استفاده از این الگوریتم وقت تیم توسعه دهنده کمتر تلف شده و تیم توسعه دهنده مستقیماً به سراغ اولویت بندی نظراتی خواهد رفت که بیشتر مورد توجه کاربر هستند.

#### ۴. ارزیابی تجربی

در این قسمت به توضیح در مورد محیط انجام آزمایش‌ها، نحوه انجام آن‌ها، معیارهای مقایسه و نتایج آزمایش پرداخته می‌شود.

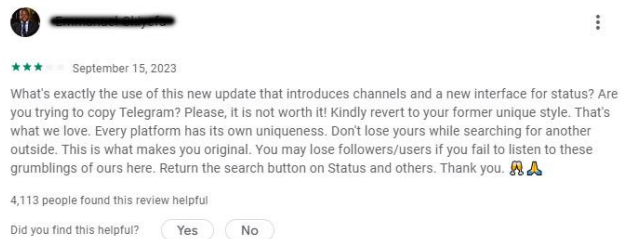
##### ۴.۱. محیط آزمایش

برای آموزش مدل از زبان برنامه نویسی پایتون نسخه 3.10.11 و محیط برنامه نویسی Visual Studio Code نسخه 1.78.2 استفاده شده است. دلیل استفاده از این محیط راحتی Debug و هم شکل بودن با Microsoft Visual Studio است.

برای ارزیابی نتایج از کامپیوتر با پردازنده Intel Xeon E5-2650 v4، 16 گیگابایت Ram از نوع DDR4، بدون کارت گرافیک و ویندوز 10 استفاده شده است.

##### ۴.۲. برنامه‌ها و نظرات مورد استفاده

برای آزمایش الگوریتم پیشنهادی در محاسبه میزان توجه کاربران، از تاریخ 3, 2023 October تا تاریخ 2, 2023 November نظرات بسیار



#### شکل ۳: نظرات ثبت شده در فروشگاه Play برای نرم افزار WhatsApp

همان‌طور که در شکل ۳ مشاهده می‌شود در ابتدا نام ثبت کننده نظر به همراه عکس پروفایل، در زیر آن تاریخ ثبت نظر مشاهده می‌شود. تاریخ ثبت نظر مربوط به چند روز پیش است اما به دلیل Thumbsup بالا (4113 People found this review helpful) نظر در بالای دیگر نظرات نشان داده شده است.

برای محاسبه تاریخ ثبت نظرات، پس از آن که نظرات در روز با استفاده از خزنده استخراج شدند، تاریخ ثبت اولین نظر با تاریخ ثبت آخرین نظر، مقایسه شده و فاصله آن‌ها بر حسب روز طبق رابطه ۱ به دست می‌آید.

$$\text{Dates} = \text{LastDate} - \text{FirstDate} \quad (1)$$

در رابطه ۱، Dates تفاوت تاریخ ثبت اولین نظر با آخرین نظر است. در حقیقت، می‌توان گفت که با استفاده از رابطه بالا می‌خواهیم محاسبه کنیم که نظرات موجود در این دسته بندی (دسته بندی درخواست ویژگی یا گزارش مشکل) در چند روز توسط کاربران ثبت شده اند زیرا هر چه این تعداد کوچک تر باشد نشان دهنده میزان اهمیت کاربر به دسته بندی مربوطه است.

#### ۲.۳. تعداد نظرات

برای محاسبه تعداد نظرات موجود در هر دسته بندی، هنگام استفاده از خزنده جهت استخراج نظرات و پس از دسته بندی آن‌ها توسط یکی ابزارها مورد اشاره در قسمت ۲، تعداد نظرات را با  $|R|$  نشان می‌دهیم که بیانگر تعداد نظرات ثبت شده توسط کاربر در آن دسته بندی خاص می‌باشد.

#### ۳.۳. الگوریتم پیشنهادی محاسبه میزان توجه کاربران (UAR)

برای محاسبه میزان توجه کاربران به دسته بندی‌های انجام شده توسط ابزار دسته بندی، از رابطه ۲ استفاده می‌شود.

$$AR_r \leftarrow \frac{|R|}{\text{Dates}} \quad (2)$$

در رابطه ۲  $|R|$  نشان دهنده تعداد نظرات، Dates برابر تعداد روزی است که در طی آن نظرات توسط کاربران ثبت شده اند و  $AR_r$  برابر میزان یا نرخ توجه کاربران به دسته بندی مورد نظر است.

۶۱/۶ درخواست ویژگی	
۹۶/۹ گزارش مشکل	VLC Media Player
۳/۱ درخواست ویژگی	

در جدول ۳ میانگین نتایج درصد توجه کاربران به دسته‌های درخواست مشکل یا گزارش مشکل که توسط متخصصان اندازه‌گیری شده، آمده‌است. با توجه به اینکه امکان اختلاف نظر بین متخصصان اختلاف نظر نیز وجود داشت، در مجموع رأی اکثریت مورد قبول واقع شد. در ادامه برای بررسی توانایی روش پیشنهادی UAR در محاسبه میزان توجه کاربران به دسته‌های نظرات، در ابتدا برای هر دسته‌بندی با استفاده از الگوریتم ۱ ارائه شده و سپس نرخ یا میزان توجه کاربران به دسته‌های نظرات محاسبه شده‌است و نتایج در جدول ۴ آورده شده‌است.

#### جدول ۴: محاسبه میزان یا نرخ توجه کاربران به دسته‌های نظرات

##### محاسبه شده توسط الگوریتم UAR

نرخ توجّه	درصد توجّه	تفاوت نرخ توجّه	نرم‌افزار
توسط الگوریتم پیشنهادی	کاربران محاسبه شده توسط الگوریتم پیشنهادی UAR	محاسبه شده توسط متخصصین و الگوریتم پیشنهادی	
۷۶/۲۳ گزارش مشکل ۲۰/۷۶ درخواست ویژگی	۷۲/۷۶ گزارش مشکل ۲۷/۲۴ درخواست ویژگی	۵/۸۴ گزارش مشکل ۵/۸۴ درخواست ویژگی	WhatsApp
۲۸/۴ گزارش مشکل ۴/۳ گزارش مشکل	۸۴/۸۵ گزارش مشکل ۱۵/۱۵ گزارش مشکل	۲/۰۵ گزارش مشکل ۲/۰۵ گزارش مشکل	NextCloud
۵۶/۵ گزارش مشکل ۴ درخواست ویژگی	۹۲/۹ گزارش مشکل ۷/۰۱ درخواست ویژگی	۰/۱۵ گزارش مشکل ۰/۴۱ درخواست ویژگی	Mozilla Firefox
۹۲/۵ گزارش مشکل ۵/۲۳ درخواست ویژگی	۹۴/۳۴ گزارش مشکل ۵/۶۶ درخواست ویژگی	۲/۵۶ گزارش مشکل ۲/۵۶ درخواست ویژگی	VLC Media Player

همان‌طور که در جدول ۴ مشاهده می‌شود، میزان توجه نظرات با توجه تعداد نظرات موجود در دسته‌بندی تقسیم بر تعداد روزی که نظرات در آن وارد شده‌است، مشاهده می‌شود. سپس در ستون مقابل میزان توجه به درصد برای فهم بهتر تبدیل شده‌است. به‌علاوه برای فهم بهتر میزان تفاوت نرخ توجه محاسبه شده توسط الگوریتم و نرخ توجه محاسبه شده نیز در کنار آن‌ها آمده‌است.

مرتبط\* مربوط به برنامه‌های WhatsApp، NextCloud، Firefox و VLC Media Player توسط خزنده وب طی ۳۰ روز استخراج شد. برای مجموع نظرات استخراج شده، ویژگی‌های نام ثبت‌کننده نظر، عکس مربوط به او در صورت وجود، نام کاربری وی، امتیاز او به برنامه (عدد بین ۱ تا ۵)، متن نظر، ThumbsUp مربوط به آن، نسخه برنامه، تاریخ ثبت نظر و نوع نظر (Most Relevant یا New) موجود است. برای هر برنامه نظرات با ویژگی‌های یادشده استخراج شد.

#### جدول ۲: تعداد نظرات استخراج شده برای هر برنامه طی ۳۰ روز

برنامه	تعداد نظرات در ۳۰ روز
WhatsApp	۲۹۱۰
NextCloud	۹۸۱
Mozilla Firefox	۱۸۱۵
VLC Media Player	۲۸۶۵

طبق جدول ۲، تعداد نظراتی که طی ۳۰ روز برای هر یک از برنامه‌ها جمع‌آوری شده‌اند، متفاوت است؛ به این دلیل که هر برنامه ممکن است دارای میزان محبوبیت بالا یا محبوبیت پایین با توجه به کاربران آن باشد. به‌علاوه باید در نظر داشت که در مدت ۳۰ روز جمع‌آوری نظرات، نظرات تکراری حذف شده و تنها نظرات متمایز و غیرتکراری به همراه ویژگی‌ها ثبت شده و به کارشناس نرم‌افزار داده شده‌اند.

#### 3.4. نتایج ارزیابی

در ادامه برای پاسخ به پرسش تحقیق، برای هر برنامه مجموع نظرات استخراج شده به ابزار دسته‌بندی اسلم و همکاران [17] - به دلیل داشتن بالاترین دقت - داده شد.

پس از دسته‌بندی نظرات با توجه به روش ارائه شده در [21]، از ۴ متخصص حوزه نرم‌افزار دارای حداقل ۳ سال سابقه کاری استفاده شد. متخصصان سوابق بسیاری در پروژه‌های تجاری داشته و به‌عنوان افراد مجرب فنی شناخته می‌شوند.

نظرات به متخصصین برای تحلیل داده شد و در مدت زمان ۷ روز هر ۴ متخصص جداگانه نظرات را بررسی و مشخص کردند که توجه کاربر چند درصد به گزارش مشکل و چند درصد به درخواست ویژگی در هر نرم‌افزار بوده‌است. میانگین نتایج بررسی در جدول ۳ آمده‌است.

#### جدول ۳: نتایج محاسبه میانگین درصد توجه کاربران به هر کدام از دسته‌ها

نرم‌افزار	میانگین درصد توجه کاربران استخراج شده توسط ۴ متخصص
WhatsApp	۷۸/۶ گزارش مشکل ۲۱/۴ درخواست ویژگی
NextCloud	۸۶/۹ درخواست ویژگی ۱۳/۱ گزارش مشکل
Mozilla Firefox	۹۳/۴ گزارش مشکل

- 590, 2020, doi: 10.1007/s40747-020-00154-3.
- [2] W. Martin, F. Sarro, Y. Jia, Y. Zhang, and M. Harman, "A survey of app store analysis for software engineering," *IEEE Trans. Softw. Eng.*, vol. 43, no. 9, pp. 817–847, 2017, doi: 10.1109/TSE.2016.2630689.
- [3] E. Guzman and W. Maalej, "How Do Users Like This Feature? A Fine Grained Sentiment Analysis of App Reviews," in *2014 IEEE 22nd International Requirements Engineering Conference (RE)*, Aug. 2014, pp. 153–162, doi: 10.1109/RE.2014.6912257.
- [4] D. Pagano and W. Maalej, "User feedback in the appstore: An empirical study," in *2013 21st IEEE International Requirements Engineering Conference (RE)*, Jul. 2013, pp. 125–134, doi: 10.1109/RE.2013.6636712.
- [5] L. V. G. Carreno and K. Winbladh, "Analysis of user comments: An approach for software requirements evolution," *Proc. - Int. Conf. Softw. Eng.*, pp. 582–591, 2013, doi: 10.1109/ICSE.2013.6606604.
- [6] W. Maalej and D. Pagano, "On the socialness of software," *Proc. - IEEE 9th Int. Conf. Dependable, Auton. Secur. Comput. DASC 2011*, pp. 864–871, 2011, doi: 10.1109/DASC.2011.146.
- [7] N. Seyff, F. Graf, and N. Maiden, "Using mobile RE tools to give end-users their own voice," *Proc. 2010 18th IEEE Int. Requir. Eng. Conf. RE2010*, pp. 37–46, 2010, doi: 10.1109/RE.2010.15.
- [8] A. Al-Subaihini et al., "App store mining and analysis," in *Proceedings of the 3rd International Workshop on Software Development Lifecycle for Mobile*, Aug. 2015, pp. 1–2, doi: 10.1145/2804345.2804346.
- [9] N. Chen, J. Lin, S. C. H. Hoi, X. Xiao, and B. Zhang, "AR-miner: Mining informative reviews for developers from mobile app marketplace," in *Proceedings - International Conference on Software Engineering*, May 2014, no. 1, pp. 767–778, doi: 10.1145/2568225.2568263.
- [10] E. C. Groen, J. Doerr, and S. Adam, "Towards Crowd-Based Requirements Engineering A Research Preview," in *Requirements Engineering: Foundation for Software Quality*, 2015, pp. 247–253.
- [11] S. A. Licorish, B. T. R. Savarimuthu, and S. Keertipati, "Attributes that predict which features to fix: Lessons for app store mining," *ACM Int. Conf. Proceeding Ser.*, vol. Part F1286, pp. 108–117, 2017, doi: 10.1145/3084226.3084246.
- [12] W. Maalej, Z. Kurtanović, H. Nabil, and C. Stanik, "On the automatic classification of app reviews," *Requir. Eng.*, vol. 21, no. 3, pp. 311–331, Sep. 2016, doi: 10.1007/s00766-016-0251-9.
- [13] S. Moghaddam, "Beyond Sentiment Analysis: Mining Defects and Improvements from Customer Feedback," 2015, pp. 400–410.
- [14] T. Johann, C. Stanik, A. M. B. Alizadeh, and W. Maalej, "SAFE: A Simple Approach for Feature Extraction from App Descriptions and App Reviews," in *Proceedings - 2017 IEEE 25th International Requirements Engineering Conference, RE 2017*, Sep. 2017, pp. 21–30, doi: 10.1109/RE.2017.71.
- [15] S. Scalabrino, G. Bavota, B. Russo, M. Di Penta, and R. Oliveto, "Listening to the Crowd for the Release Planning of Mobile Apps," *IEEE Trans. Softw. Eng.*, vol. 45, no. 1, pp. 68–86, 2019, doi: 10.1109/TSE.2017.2759112.
- [16] D. Martens and W. Maalej, "Towards understanding and detecting fake reviews in app stores," *Empir. Softw. Eng.*, vol. 24, no. 6, pp. 3316–3355, Dec. 2019.
- الگوریتم UAR پیشنهادی در مقایسه با روش دستی محاسبه میزان توجه کاربران در دسته گزارش مشکل توانسته به‌طور میانگین با اختلاف ۲/۷۳٪ میزان توجه کاربران را محاسبه کند.
  - به‌علاوه الگوریتم UAR پیشنهادی توانسته در مقایسه با روش دستی، به‌طور میانگین با اختلاف ۲/۷۱٪ میزان توجه نظرات کاربران به دسته درخواست ویژگی را محاسبه کند.
  - باید توجه داشت که تاکنون هیچ روش یا ابزار یا مدلی برای محاسبه میزان توجه کاربران و گرایش آن‌ها به دسته‌های خاص نظرات ارائه نشده است. در نتیجه مقایسه الگوریتم ارائه شده با دیگر روش‌ها اشتباه است.
- برای مقایسه روش پیشنهادی با روش دستی تنها راه، استفاده از ۴ کارشناس خبره در زمینه برنامه‌نویسی نرم‌افزار بود. زیرا هیچ ابزار یا روشی برای مقایسه با روش پیشنهادی وجود نداشت.
- ### ۵. نتیجه‌گیری و کارهای آینده
- روزانه نظرات بسیاری توسط کاربران برای برنامه‌های موبایل ثبت می‌شود. با توجه به آن که حجم نظرات ثبت شده بسیار بالا بوده و تحلیل آن‌ها برای تیم توسعه‌دهنده امری دشوار است، ابزار یا روشی مورد نیاز است تا میزان یا نرخ توجه کاربران به نظرات برنامه مورد نظر با توجه به دسته‌های گزارش مشکل یا درخواست ویژگی بررسی کند. در صورت داشتن چنین ابزاری و اندازه‌گیری میزان توجه کاربران، تیم توسعه‌دهنده مستقیماً و بدون اتلاف وقت به دسته مورد توجه کاربر مراجعه کرده و علاوه بر عدم اتلاف زمان، به راحتی ویژگی‌های مورد درخواست کاربر یا مشکلات گزارش شده توسط او را در نسخه‌های بعدی پیاده‌سازی کرده یا پاسخ می‌دهد. بدین ترتیب، نرم‌افزار در بین رقبای محبوبیت خود را حفظ کرده و همیشه به‌روز خواهد بود.
- در این تحقیق الگوریتمی با عنوان UAR (User Attention Rate) برای محاسبه میزان توجه کاربران به دسته‌های گزارش مشکل یا درخواست ویژگی ارائه شده است. برای آزمایش الگوریتم یا روش پیشنهادی از ۴ متخصص کمک گرفته شد و از آن‌ها خواسته شد تا میزان توجه کاربران به نظرات مربوط به ۴ برنامه از فروشگاه Play را بررسی کنند. در نهایت با مقایسه میزان توجه تعیین شده توسط متخصصین و میزان یا نرخ توجه محاسبه شده توسط روش پیشنهادی، مشخص شد که الگوریتم پیشنهادی توانسته با اختلاف بسیار کم میزان توجه کاربران به دسته‌های گزارش مشکل یا درخواست ویژگی یا تشخیص دهد.
- با استفاده از ویژگی‌های بیشتر، می‌توان در کارهای آینده، میزان توجه کاربران به دسته‌بندی‌های بیشتر را بررسی کرد.

### مراجع

- [1] M. R. Dehkordi, H. Seifzadeh, G. Beydoun, and M. H. Nadimi-Shahraki, "Success prediction of android applications in a novel repository using neural networks," *Complex Intell. Syst.*, vol. 6, no. 3, pp. 573–



- doi: 10.1007/s10664-019-09706-9.
- [17] N. Aslam, W. Y. Ramay, K. Xia, and N. Sarwar, "Convolutional neural network based classification of app reviews," *IEEE Access*, vol. 8. Institute of Electrical and Electronics Engineers Inc., pp. 185619–185628, 2020, doi: 10.1109/ACCESS.2020.3029634.
- [18] Z. Qiao, A. Wang, A. Abrahams, and W. Fan, "Deep Learning-Based User Feedback Classification in Mobile App Reviews," 2020. [Online]. Available: <https://aisel.aisnet.org/sigdsa2020>.
- [19] "What Criteria are Used by Google to Rank Reviews Most Relevant? | LaPram Insights." <https://lapraim.com/insights/criteria-for-relevant-google-reviews> (accessed Sep. 18, 2023).
- [20] "How Google Decides What Reviews Are Most Relevant." <https://www.widewail.com/blog/how-google-decides-what-reviews-are-most-relevant> (accessed Sep. 18, 2023).
- [21] Z. Wan, X. Xia, D. Lo, and G. C. Murphy, "How does machine learning change software development practices?," *IEEE Trans. Softw. Eng.*, vol. 47, no. 9, pp. 1857–1871, 2021, doi: 10.1109/TSE.2019.2937083.