

The main components of evaluating the credibility of users according to organizational goals in the life cycle of big data

Sogand Dehghan^{*}, Shahriyar Mohammadi^{**}, Rojyar Pirmohammadiani^{***}

^{*}Master's student, Faculty of Industrial Engineering, Khajeh Nasiruddin Toosi University of Technology, Iran

^{**}Associate Professor, Faculty of Industrial Engineering, Khajeh Nasiruddin Toosi University of Technology, Iran

^{***}Assistant Professor, Faculty of Computer Engineering, University of Kurdistan, Iran

Abstract

Social networks have become one of the most important decision-making factors in organizations due to the speed of publishing events and the large amount of information. For this reason, they are one of the most important factors in the decision-making process of information validity. The accuracy, reliability and value of the information are clarified by these networks. For this purpose, it is possible to check the validity of information with the features of these networks at the three levels of user, content and event. Checking the user level is the most reliable level in this field, because a valid user usually publishes valid content. Despite the importance of this topic and the various researches conducted in this field, important components in the process of evaluating the validity of social network information have received less attention. Hence, this research identifies, collects and examines the related components with the narrative method that it does on 30 important and original articles in this field. Usually, the articles in this field are comparable from three dimensions to the description of credit analysis approaches, content topic detection, feature selection methods. Therefore, these dimensions have been investigated and divided. In the end, an initial framework was presented focusing on evaluating the credibility of users as information sources. This article is a suitable guide for calculating the amount of credit of users in the decision-making process.

Keywords: information validity, valid users, valid information sources, big data life cycle, organizational goals, social networks

مولفه‌های اصلی ارزیابی اعتبار کاربران با توجه به اهداف سازمانی در چرخه حیات کلان داده

سوگند دهقان^{*}، شهریار محمدی^{**۱}، روجیار پیرمحمدیانی^{***}

^{*} دانشجوی کارشناسی ارشد، دانشگاه صنعتی خواجه نصیرالدین طوسی، ایران

^{**} دانشیار، دانشکده مهندسی صنایع، دانشگاه صنعتی خواجه نصیرالدین طوسی، ایران

^{***} استادیار، دانشکده مهندسی کامپیوتر، دانشگاه کردستان، ایران

تاریخ پذیرش: ۱۴۰۱/۰۶/۲۴

تاریخ دریافت: ۱۴۰۱/۰۴/۳۱

نوع مقاله: پژوهشی

چکیده

شبکه‌های اجتماعی به دلیل سرعت انتشار رویدادها و نیز حجم زیاد اطلاعات، به یکی از مهم‌ترین منابع تصمیم‌گیری در سازمان‌ها تبدیل شده‌اند. ولی پیش از استفاده، صحت، قابلیت اطمینان و ارزش اطلاعات تولید شده توسط آن‌ها، باید مورد ارزیابی قرار گیرد. به این منظور، بررسی اعتبار اطلاعات با توجه به ویژگی‌های شبکه‌های اجتماعی در سه سطح کاربر، محتوا و رویداد امکان‌پذیر می‌باشد. سطح کاربر، قابل اطمینان‌ترین سطح این حوزه می‌باشد، زیرا کاربر معتبر، معمولاً اقدام به انتشار محتوای معتبر می‌نماید. از این رو، ارزیابی سطح کاربر مورد توجه این پژوهش می‌باشد.

بیش تر مقالات مرتبط به حوزه ارزیابی اعتبار کاربران شبکه‌های اجتماعی به ارزیابی اعتبار کاربران در حالت کلی پرداخته‌اند و از اهداف سازمانی مانند ارزیابی اعتبار کاربران به منظور یافتن افراد خلاق چشم پوشی نموده‌اند. همچنین چرخه حیات کلان داده و مولفه‌های مهم در فرآیند ارزیابی اعتبار کاربران کمتر مورد توجه قرار گرفته‌اند. از این رو، این تحقیق با بررسی ۵۰ مقاله مهم در این حوزه، مولفه‌های مهم را به سه مولفه اصلی (تعیین موضوع محتوا، انتخاب ویژگی‌ها و ارزیابی اعتبار) دسته بندی می‌نماید و روش‌ها و ویژگی‌های مربوط به هر یک را مورد بحث قرار می‌دهد. نهایتاً یک چارچوب اولیه ارزیابی اعتبار کاربران شبکه‌های اجتماعی با توجه به اهداف سازمانی و چرخه حیات کلان داده ارائه گردید. هدف این چارچوب، ارائه یک راهنما مناسب به سازمان‌ها، برای محاسبه میزان اعتبار کاربران در فرآیند تصمیم‌گیری می‌باشد.

واژگان کلیدی: اعتبار اطلاعات، کاربران معتبر، منابع اطلاعاتی معتبر، چرخه حیات کلان داده، اهداف سازمانی، شبکه‌های اجتماعی

^۱ نویسنده مسئول: شهریار محمدی mohammadi@kntu.ac.ir

۱. مقدمه

مقالات این حوزه به دلایل ذیل بسیار متنوع و پراکنده می‌باشند:

- اهداف متفاوت ارزیابی اعتبار اطلاعات مانند اهداف سیاسی و رخدادهای طبیعی که منجر به تنوع در انتخاب و اولویت‌بندی ویژگی‌های می‌شود.

- ارزیابی اعتبار اطلاعات در سه سطح کاربر، رویداد و محتوا امکان‌پذیر می‌باشد.

- از رویکردهای متنوعی، برای تحلیل اعتبار مانند سیستم تصمیم‌گیری چند معیاره^۲ و یادگیری ماشین استفاده می‌شود.

- هر شبکه اجتماعی شامل ویژگی‌های منحصر به فردی می‌باشد.

- ساختارهای شبکه‌های اجتماعی مانند فیس‌بوک، اینستاگرام و توئیتر متفاوت می‌باشد.

- با توجه به سرعت رشد شبکه‌های اجتماعی، عدم وجود یک راه حل جامع برای ارزیابی اعتبار اطلاعات و تغییر سریع استراتژی کاربران اسپم در مقابل راه‌حل‌های متنوع ارزیابی اعتبار اطلاعات، تعداد مقالات مرتبط به حوزه ارزیابی اعتبار اطلاعات شبکه‌های اجتماعی روبه افزایش می‌باشد [۱۲].

با توجه به تنوع زیاد مقالات مربوط به این حوزه، تعیین مولفه‌های اصلی مربوط به فرآیند ارزیابی کاربران برای سازمان‌ها دشوار می‌باشد. علاوه بر مورد مذکور، اهمیت اهداف سازمانی در ارزیابی اعتبار اطلاعات نادیده گرفته شده است و معمولاً مقالات این حوزه به بررسی اعتبار اطلاعات براساس اهداف سیاسی، رخدادهای طبیعی، بازاریابی و یا به صورت عمومی (بدون توجه به موضوع) پرداخته‌اند، این در حالی می‌باشد، که سازمان‌ها براساس اهداف و مأموریت‌هایی که در نظر می‌گیرند، مانند بازاریابی، افزایش شهرت تجاری، تشویق به پذیرش خلاقیت، تغییر عقیده، افزایش خلاقیت، حل مشکلات در برابر پارادوکس، بهبود وفاداری مشتری، استخدام و ارتقا کارکنان و مدیریت

شبکه‌های اجتماعی بستری بر روی اینترنت می‌باشند، که افراد با طرز فکر مشابه یا افراد علاقمند به موضوعات خاص به یکدیگر مرتبط می‌شوند و ارتباطات را برای کاربران ارزان و سریع می‌سازند [۱] و [۲]. این شبکه‌ها براساس عملکردشان به دسته‌های متفاوت تقسیم می‌شوند. نمونه‌ای از محبوب‌ترین آن‌ها فیس‌بوک، اینستاگرام و توئیتر می‌باشند، که با سرعت زیاد اطلاعات را منتشر می‌سازند. به طور مثال در توئیتر، در هر ثانیه ۳۵۰ هزار پیام، تولید می‌گردد [۳]. از این رو، حجم عظیم و روبه‌رشد اطلاعات، اهمیت این شبکه‌ها را برای سازمان‌ها افزایش داده است و تعداد کاربران هر شبکه، بیانگر ارزش تجاری آن‌ها می‌باشد [۱]. بنابراین، این شبکه‌ها به یکی از مهم‌ترین منابع تصمیم‌گیری در سازمان‌ها تبدیل شده‌اند و سازمان‌ها برای اهدافی متفاوتی از این شبکه‌ها استفاده می‌نمایند [۴]، [۵]، [۶]، [۷]، [۸] و [۹].

ارزیابی اعتبار اطلاعات، عامل مهم دیگری در فرایند تصمیم‌گیری می‌باشد، که بیان‌کننده میزان صحت و ارزش اطلاعات می‌باشد [۴]، [۵]. اعتبار کاربران نیز، میزان قابلیت اطمینان به آن‌ها را به عنوان منابع اطلاعاتی نشان می‌دهد. ارزیابی اعتبار اطلاعات شبکه‌های اجتماعی، با بررسی ویژگی‌های آن‌ها، در سه سطح محتوا، کاربر و رویداد امکان‌پذیر شده است. در سطح رویداد، ویژگی‌های مرتبط به یک رویداد مانند رخدادهای طبیعی، بررسی می‌گردند. در سطح محتوا، ویژگی‌های مرتبط به محتوا ارزیابی می‌شوند. در سطح کاربر، ویژگی‌های یک کاربر به عنوان منبع اطلاعاتی و محتوای تولید شده توسط آن نیز بررسی می‌گردد. از این رو، در سطح کاربر، علاوه بر ویژگی‌های مرتبط به رویداد و محتوا، ویژگی‌های مرتبط به کاربر نیز بررسی می‌شود [۵]. بنابراین بین سطح کاربر و محتوا همبستگی وجود دارد و یک کاربر معتبر، معمولاً محتوا معتبر ارائه می‌دهد. به این منظور، بررسی اعتبار کاربران به عنوان منابع اطلاعاتی، بهترین گزینه جهت ارزیابی اعتبار اطلاعات شبکه‌های اجتماعی می‌باشد [۱۰]، [۱۱]، [۱۲].

² Multiple-criteria decision-making (MCDM)

۳. ارائه یک ساختار سلسله مراتبی برای انتخاب ویژگی‌های شبکه‌های اجتماعی متناسب با اهداف سازمانی.

سازمان‌ها با اهداف متفاوت، با استفاده از چارچوب اولیه این پژوهش می‌توانند، اعتبار کاربران شبکه‌های اجتماعی را اندازه‌گیری نمایند. نمونه این چارچوب در کارهای پیشین دیده نشده است. مولفه‌های پیشنهادی این مقاله جهت ارزیابی اعتبار کاربران متمرکز می‌باشد و این پژوهش با طبقه‌بندی آکادمیک مقالات گذشته و استخراج مولفه‌های اصلی ارزیابی اعتبار اطلاعات سعی دارد تا به این پرسش پاسخ دهد، که کاربران معتبر شبکه‌های اجتماعی براساس اهداف سازمانی چه نوع کاربرانی خواهند بود؟ و فرآیند صحیح برای ارزیابی اعتبار کاربران در جهت اهداف سازمانی به چه شکل می‌باشد؟

این پژوهش در ۶ بخش سازماندهی شد. بخش ۲ مبانی پژوهش می‌باشد. بخش ۳ مقالات پیشین را براساس مولفه‌ها بررسی می‌نماید. در بخش ۴ چارچوب اولیه پیشنهادی ارائه گردید. بخش ۵ به بحث در ارتباط با چارچوب پیشنهادی پرداخته است، بخش ۶ نتیجه پژوهش را ارائه می‌دهد.

۲. مبانی پژوهش

به دلیل تنوع زیاد مقالات حوزه ارزیابی اعتبار اطلاعات، این پژوهش زیرمولفه‌های ارزیابی اعتبار کاربران را به سه مولفه اصلی طبقه‌بندی نموده است، که شامل کل فرآیند ارزیابی اعتبار کاربران از ابتدا تا انتها می‌باشد. این مولفه‌ها و فرآیند کلی ارزیابی اعتبار اطلاعات در شکل ۱ نمایش داده شده است.

تعیین موضوع محتوا در شبکه اجتماعی، جهت ارزیابی اعتبار اطلاعات بسیار مهم می‌باشد، به خصوص، اگر بررسی اعتبار کاربران هدف ارزیابی اعتبار اطلاعات باشد. زیرا کاربران به عنوان منابع اطلاعاتی، تنها در موضوعات خاصی توانایی کسب اعتبار را دارند [۴]. [۱۵]. به این منظور در برخی از مقالات، با توجه به موضوع فعالیت کاربران، اعتبار

بحران از شبکه‌های اجتماعی بهره‌مند می‌شوند و براین اساس، تعاریف متفاوتی از کاربران معتبر ارائه می‌دهند. از این رو، فقدان یک چارچوب برای ارزیابی اعتبار کاربران با توجه به اهداف سازمانی، وجود دارد. در این راستا، این پژوهش به بررسی مقالات حوزه ارزیابی اعتبار اطلاعات شبکه‌های اجتماعی در ژورنال‌های مختلف پرداخته است. از مقالات بررسی شده، ۵۰ مقاله مهم و اصلی با تعداد استناد بالا انتخاب شده‌اند و بر این اساس در قدم اول مولفه‌های مرتبط به این حوزه شناسایی، جمع‌آوری و بررسی می‌شوند و در قالب ۳ مولفه اصلی: ۱. تعیین موضوع محتوا، ۲. انتخاب ویژگی و ۳. ارزیابی اعتبار معرفی شده‌اند و مقالات مرتبط به این حوزه براساس این زیرمولفه آن‌ها بررسی می‌شوند. در نهایت پس از شناسایی مولفه‌های مهم این حوزه، یک چارچوب اولیه ارزیابی اعتبار کاربران طراحی گردید و با توجه به این که شبکه‌های اجتماعی پویا می‌باشند و با سرعت و حجم بالا داده تولید می‌نمایند و تحلیل اعتبار کاربر حساس به زمان و گذارا است [۳]، [۱۲]، از این رو، این چارچوب براساس چرخه حیات کلان داده^۳ طراحی گردیده است. در این چارچوب، یک ساختار برای انتخاب ویژگی‌ها پیشنهاد شد، زیرا انتخاب و اولویت‌بندی ویژگی‌ها باید با اهداف متعدد سازمانی سازگار باشد. سهم اصلی این تحقیق براساس مرور روایتی ادبیات حوزه ارزیابی اعتبار اطلاعات شبکه‌های اجتماعی به شرح ذیل می‌باشد:

۱. دسته بندی زیرمولفه‌های حوزه ارزیابی اعتبار اطلاعات شبکه‌های اجتماعی در ۳ مولفه تعیین موضوع، انتخاب ویژگی و ارزیابی اعتبار

۲. زیرمولفه‌های مهم ارزیابی اعتبار شامل، تشخیص موضوع صریح، تشخیص موضوع ضمنی، سیستم انتخاب ویژگی بر اساس هدف ارزیابی اعتبار، پیش‌بینی اعتبار کاربران در آینده، بررسی اعتبار کاربران در گذر زمان و قابلیت فهم آسان نتایج می‌باشند و براین اساس چارچوب اولیه طراحی گردیده است.

³ Big Data (BD)

نسبت به مقالات عمومی برتری دارند، اما مقالات عمومی فاقد ارزش نیستند و پایه و اساس این حوزه می‌باشند [۲]. موضوع فعالیت کاربران، براساس محتوا منتشر شده توسط آن‌ها، قابل تشخیص می‌باشند. از این رو، پست‌ها و پاسخ‌ها منابع ارزشمندی جهت تعیین موضوع فعالیت در شبکه‌های اجتماعی می‌باشند [۹].

مولفه تعیین موضوع محتوا، شامل دو زیرمولفه صریح (C1) و ضمنی (C2) می‌باشد. موضوعاتی که کاربران در آن‌ها به صورت آشکار فعالیت دارند، به گروه صریح تعلق دارند و موضوعات که کاربران تنها به آن‌ها علاقمند می‌باشند و یا غیرمستقیم در آن‌ها فعالیت دارند، از گروه ضمنی می‌باشند [۴۱]. بیش‌تر مقالات به زیرمولفه تعیین موضوع به صورت صریح توجه نموده‌اند و از زیرمولفه تعیین موضوع ضمنی چشم‌پوشی شده است. این پژوهش در شکل ۲ با استفاده از مرور ادبیات به طبقه‌بندی روش‌های مولفه صریح پرداخته است.

۲.۲ انتخاب ویژگی

شبکه‌های اجتماعی شامل ویژگی‌های متنوعی می‌باشند، تعدادی از آن‌ها در بین انواع مختلف این شبکه‌ها، مشترک هستند و برخی دیگر منحصر به یک نوع از شبکه‌های اجتماعی می‌باشند. معمولاً این ویژگی‌ها به دو دسته فردی و رابطه‌ای تقسیم می‌شوند، ویژگی‌های فردی، به طور مستقل قابل محاسبه می‌باشند، مانند تعداد کلمات پست‌ها و ویژگی‌های رابطه‌ای از تعامل گره‌ها (گره‌ها می‌توانند کاربران، محتوا و یا رویدادها باشند). در شبکه‌ها محاسبه می‌شوند، مانند تعیین مرکزیت کاربران. به طور کلی با استفاده از ۴ روش به شرح اندازه‌گیری محلی، محاسبات بازگشتی، محاسبه براساس هسته و مبتنی بر مسیر کوتاه، ویژگی‌های رابطه‌ای محاسبه می‌گردند [۲].

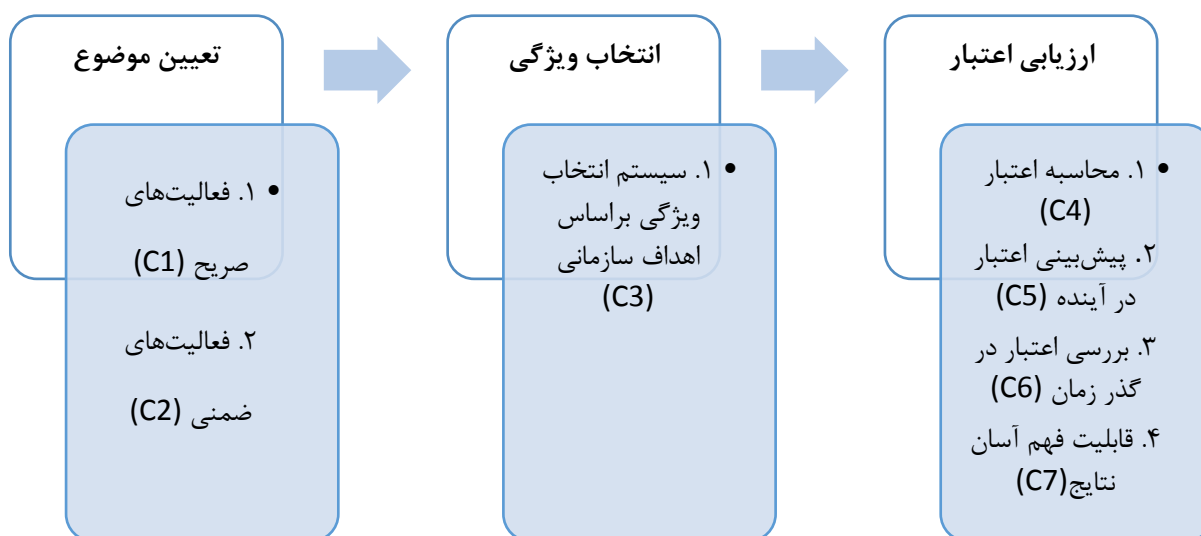
این پژوهش، براساس اهداف ارزیابی اعتبار، ویژگی‌ها را به ۷ گروه به شرح ذیل تقسیم نموده است و ویژگی‌های هر گروه در پیوست الف معرفی شده‌اند.

آن‌ها مورد بررسی قرار گرفته است. اما در برخی دیگر، به صورت عمومی اعتبار کاربران بررسی شده است. آن‌ها به اشتباه تصور کرده‌اند، که یک کاربر توانایی کسب اعتبار را در تمام موضوعات دارد [۴]. در ادامه، این مقالات تعدادی از ویژگی‌های شبکه اجتماعی را انتخاب می‌نمایند و با استفاده از رویکردهای متنوع، اعتبار اطلاعات را محاسبه می‌نمایند. هر یک از این ۳ مولفه (تعیین موضوع محتوا، انتخاب ویژگی، رویکردهای تحلیل اعتبار) در این بخش توضیح داده می‌شوند.

۱.۲ تعیین موضوع محتوا

مقالات حوزه ارزیابی اعتبار اطلاعات براساس مولفه تعیین موضوع محتوا، به دو دسته عام و خاص قابل تقسیم می‌باشند، در مقالات عام، اعتبار بدون توجه به موضوع محتوا و به صورت کلی بررسی می‌گردد، این مقالات از زیرمولفه‌های تعیین موضوع محتوا چشم‌پوشی نموده‌اند [۱۰]، [۱۱]، [۱۴]، [۱۶]، [۱۷]، [۱۸]، [۱۹]، [۲۰]، [۲۱]. اما در مقالات خاص اعتبار با توجه به موضوع محتوا، بررسی می‌شود [۱۵]، [۲۲]، [۲۳]، [۲۴]، [۲۵]، [۲۶]، [۲۷]، [۲۸]، [۲۹]، [۳۰]، [۳۱]، [۳۲]، [۳۳]، [۳۴]، [۳۵]، [۳۶]، [۳۷]، [۳۸].

کاربران تنها در موضوعاتی که به آن علاقمند باشند و یا در آن فعالیت داشته باشند، امکان دریافت اعتبار را دارند. هیچ کاربری قادر به دریافت اعتبار در همه موضوعات نیست [۴]، [۱۵]. همچنین سازمان‌ها به جهت اهداف متنوع، پیش از بهره‌برداری از اطلاعات شبکه‌های اجتماعی و ارزیابی اعتبار آن‌ها، باید موضوع فعالیت کاربران را تشخیص دهند، به عنوان مثال شرکت‌های فعال در طراحی و ساخت تلفن‌های همراه، جهت افزایش خلاقیت در محصولات، به بررسی کاربران فعال در این زمینه پرداخته‌اند [۸]. از این رو، یکی از مهم‌ترین مولفه‌ها در ارزیابی اعتبار کاربران شبکه اجتماعی، تعیین موضوع فعالیت آن‌ها می‌باشد و مقالات خاص به منظور توجه به روش‌های تشخیص موضوع فعالیت



شکل ۱ مولفه‌های اصلی و فرآیند کلی ارزیابی اعتبار کاربران

لزوما ویژگی‌های انتخاب شده در این مرحله، برای ارزیابی اعتبار اطلاعات مناسب نمی‌باشند. از این رو، باید پیش از تحلیل اعتبار اطلاعات، تاثیر آن‌ها با زیرمولفه سیستم انتخاب ویژگی (C3) بررسی گردد. بنابراین انتخاب و اولویت‌بندی ویژگی‌ها از دیگر چالش‌های مهم این حوزه می‌باشد. انتخاب ناصحیح این ویژگی‌ها، موجب بررسی اعتبار اطلاعات با دقت پایین می‌گردد [۲۰]، [۲۶]، [۳۳]. از طرف دیگر، بیش‌تر مقالات این حوزه به وسیله مرور ادبیات و بدون تحلیل تاثیر ویژگی‌ها بر نتایج ارزیابی اعتبار اطلاعات، آن‌ها را انتخاب نموده‌اند و تعداد اندکی از مقالات، قبل از ارزیابی اعتبار اطلاعات، به بررسی تاثیر ویژگی‌ها با استفاده از MCDM، آمار و یا یادگیری ماشین، پرداخته‌اند [۱۰]، [۱۱]، [۲۰]، [۲۴]، [۲۶]، [۲۸]، [۳۶]، [۳۷]، [۳۸].

متدهای انتخاب ویژگی یادگیری ماشین: این متدها شامل ۳ نوع متد، متد پوشه^۴، فیلتر و تعبیه شده^۵ می‌باشد. این

- محبوبیت (F1): ویژگی‌هایی که توانایی کاربران را در تحت تاثیر قرار دادن، سایر افراد نمایش می‌دهد.
- محتوا (F2): ویژگی‌هایی که نشان‌دهنده کیفیت محتوا می‌باشند.
- اطلاعات کلی (F3): ویژگی‌هایی که کاربر را به صورت کلی معرفی می‌نمایند.
- انتشار (F4): کاربر با استفاده از این ویژگی‌ها، بازدید از حساب خود را افزایش می‌دهد.
- فعالیت (F5): میزان فعالیت کاربران در شبکه‌های اجتماعی را نشان می‌دهد.
- پشتیبانی (F6): نشان‌دهنده حمایت کاربر از سایر افراد می‌باشد.
- تخصص (F7): میزان تخصص افراد در موضوعات مختلف را مشخص می‌نماید.

⁵ Embedded method

⁴ wrap

اتوماتیک به دو گروه مستقل و وابسته به داده قابل تقسیم می‌باشد.

رویکردهای انسانی: با استفاده از نظرسنجی و روش‌های آماری، مقدار نهایی اعتبار اطلاعات محاسبه می‌گردد. این رویکردها بخاطر دخالت‌های انسانی، معمولاً با خطا همراه می‌باشند و شامل ناسازگاری و عدم اطمینان در قضاوت و ارزیابی می‌باشند [۱].

مستقل از داده: به داده‌ها وابستگی ندارند، با استفاده از عملگرهای تجمیع و یا روش‌های رتبه بندی MCDM مقدار نهایی اعتبار اطلاعات محاسبه می‌گردد [۱۰]، [۱۴]، [۱۵]، [۲۰]، [۳۵]، [۳۶]، [۳۷]. برخی از روش‌های این رویکرد شامل تاپسیس^۷، انتگرال فازی چکیت^۸، ویکور^۹ و ویکور فازی می‌باشد.

وابسته به داده: همان روش‌های یادگیری ماشین و یادگیری عمیق می‌باشند که به پیش‌بینی می‌پردازند و جهت ارائه نتایج دقیق، نیازمند ویژگی‌های قدرتمندی می‌باشند. به سه شیوه یادگیری باناظر، بدون ناظر و نیمه‌نظارت تقسیم می‌شوند.

باناظرات: به نمونه‌ها با حجم زیاد و برجسب خورده نیاز دارند و بیشتر مقالات از این رویکرد استفاده نموده‌اند [۱۶]، [۱۷]، [۱۸]، [۲۲]، [۲۳]، [۲۴]، [۲۵]، [۲۶]، [۲۷]، [۲۸]، [۳۳]. برخی از روش‌های این دسته شامل: ماشین‌بردار پشتیبانی (SVM)^{۱۰}، بیز ساده (NB)^{۱۱}، جنگل تصادفی (RF)^{۱۲}، رگرسیون لجستیک^{۱۳}، درخت تصمیم‌گیری (DT)^{۱۴}، K نزدیکترین همسایه (KNN)^{۱۵}، اکس جی باست^{۱۶}، رگرسیون خطی (LR)^{۱۷}، پرسپترون چند لایه^{۱۸}

بدون نظارت: ناتوانی در تشخیص خروجی مورد انتظار برای داده‌های ورودی معین دارند [۲۹]، [۳۰]، [۳۱]، [۳۲]، [۳۴]. روش‌های این دسته شامل: آیتپی توئیتر^{۱۹}

متد انتخاب ویژگی، وابسته به داده می‌باشد و هریک شامل چندین الگوریتم می‌باشند.

متدهای انتخاب ویژگی MCDM: این متدها، به داده‌ها وابستگی ندارند و تعامل بین ویژگی‌ها را بهتر محاسبه می‌نمایند [۳۷]. نمونه‌ای از آن‌ها شامل فرآیند تحلیل سلسله مراتبی (AHP)^۶، فازی و آنتروپی می‌باشد.

۳.۲ ارزیابی اعتبار

پس از انتخاب ویژگی‌های شبکه اجتماعی، با استفاده از رویکردهای تحلیلی معرفی شده در این بخش، میزان اعتبار اطلاعات محاسبه می‌گردد. این مجموعه شامل ۴ زیرمولفه به شرح ذیل می‌باشد:

- محاسبه اعتبار (C4): این مولفه با توجه به ویژگی‌های شبکه اجتماعی به محاسبه اعتبار کاربران می‌پردازد.

- پیش‌بینی اعتبار در آینده (C5): با توجه به این که ممکن است، در آینده ارتباطات جدیدی بین کاربران شکل بگیرد و یا به فعالیت در موضوع جدیدی علاقمند شوند، به پیش‌بینی میزان اعتبار کاربران در آینده می‌پردازند.

- بررسی اعتبار در گذر زمان (C6): به دلیل ماهیت ناشناس و سریع ارتباطات آنلاین، اعتماد به کاربران پویا و گذرا می‌باشد [۱۲]. به این منظور، این زیرمولفه با در نظر گرفتن پارامتر زمان، اعتبار کاربران را در بازه‌های زمانی متفاوت محاسبه می‌نماید [۱۵].

- قابلیت فهم آسان نتایج (C7): این مولفه با تعیین نقطه آستانه و گروه‌بندی کاربران براساس این نقطه، نتایج ارزیابی اعتبار اطلاعات را برای عموم قابل فهم می‌نمایند.

در این پژوهش با بررسی ادبیات رویکردهای حوزه ارزیابی اعتبار اطلاعات شبکه‌های اجتماعی، مطابق شکل ۲، به دو گروه اتوماتیک و انسانی تقسیم گردیده است. گروه

¹⁴ Decision Tree (DT)

¹⁵ k-nearest neighbors algorithm (k-NN)

¹⁶ XGBoost

¹⁷ Linear Regression (LR)

¹⁸ Multi-layer perceptron (MLP)

¹⁹ ITPTWITTER

⁶ Analytic hierarchy process (AHP)

⁷ Topsis

⁸ choquet

⁹ Vikor

¹⁰ Support vector machines

¹¹ Naive Bayes (NB)

¹² Random Forest (RF)

¹³ Logistic regression

در سازمان‌ها بسیار مهم می‌باشد، که معمولاً از اهداف سازمانی چشم‌پوشی شده است.

- با توجه به این که سازمان‌ها براساس اهداف و ماموریت‌هایی که در نظر می‌گیرند. تعاریف متفاوتی از کاربران معتبر ارائه می‌دهند. بنابراین، ویژگی‌های شبکه‌های اجتماعی باید براساس این اهداف و ماموریت‌ها، انتخاب و اولویت‌بندی شوند. اما معمولاً یک روش مناسب جهت اولویت‌بندی ویژگی‌های صحیح، متناسب با هدف ارزیابی اعتباراطلاعات در این حوزه ارائه نگردیده است و بیش‌تر کارها، ویژگی‌ها را براساس مرور ادبیات و بدون تحلیل تاثیر آن‌ها بر نتایج ارزیابی اعتباراطلاعات انتخاب کرده‌اند.

- علی‌رغم آن که، شبکه‌های اجتماعی پویا می‌باشند و با سرعت و حجم بالا داده تولید می‌نمایند و تحلیل اعتبار کاربر حساس به زمان و گذار است [۳]، [۱۲]. تعداد مقالات کمی به زیرمولفه C6 و چرخه حیات BD توجه نموده‌اند.

رتبه‌بندی صفحه^{۲۰}، فاکتورسازی ماتریسی، خوشه‌بندی، قوانین تداعی^{۲۱}، تخصیص دیریکله نهفته^{۲۲}، حافظه کوتاه مدت بلند مدت^{۲۳}، نمایش رمزگذار دو طرفه از ترانسفورماتورها^{۲۴}

نیمه‌نظارت: این روش‌ها به دو گروه نیمه‌نظارت و نیمه‌نظارتی مبتنی بر گراف^{۲۵} قابل تقسیم می‌باشند [۱۹]، [۳۳]. روش‌های این دسته شامل: الگوریتم‌های طبقه‌بندی تکراری^{۲۶} و به کارگیری روش‌های یادگیری ماشین بانظارت به شیوه خاص

۳. پیشینه تحقیق با توجه به مولفه‌ها

کارها زیادی در حوزه ارزیابی اعتباراطلاعات انجام شده است. از ۵۰ مقاله بازنگری شده، ۲۸ مقاله انتخاب شد و به‌صورت کامل براساس زیرمولفه‌ها، اهداف و توجه آن‌ها به چرخه حیات BD در جدول ۱ بررسی شده‌اند. معیار انتخاب مقالات این پژوهش به شرح ذیل می‌باشد:

- مقالات منتشر شده در ژورنال‌ها و کنفرانس‌های مربوط به هوش تجاری، علوم کامپیوتر، تحلیل شبکه‌های اجتماعی و مدیریت مشتری

- مقالاتی که به‌طور واضح مولفه‌ها و تکنیک‌های به‌کار گرفته شده را، بیان کرده‌اند.

-مقالاتی که تعداد استناد بالایی دارند.

براساس جدول ۱، شکاف‌های تحقیقاتی به شرح ذیل می‌باشد:

- به دلیل تنوع و نیز پراکندگی مقالات این حوزه، کارهای پیشین به زیرمولفه‌های مهم تحلیل اعتباراطلاعات توجه کم‌تری داشته‌اند. و نگاه جامع و ترکیبی از کلیه زیرمولفه‌ها در نظر گرفته نشده است.

- اعتباراطلاعات معمولاً در زمینه‌های سیاسی، رخدادهای طبیعی، بازاریابی و یا به‌صورت عمومی ارزیابی شده است. از طرف دیگر، ارزیابی اعتباراطلاعات شبکه‌های اجتماعی

²⁴ Bidirectional Encoder Representations from Transformers (BERT)

²⁵ Graph-based semi-supervised learning (G-SSL)

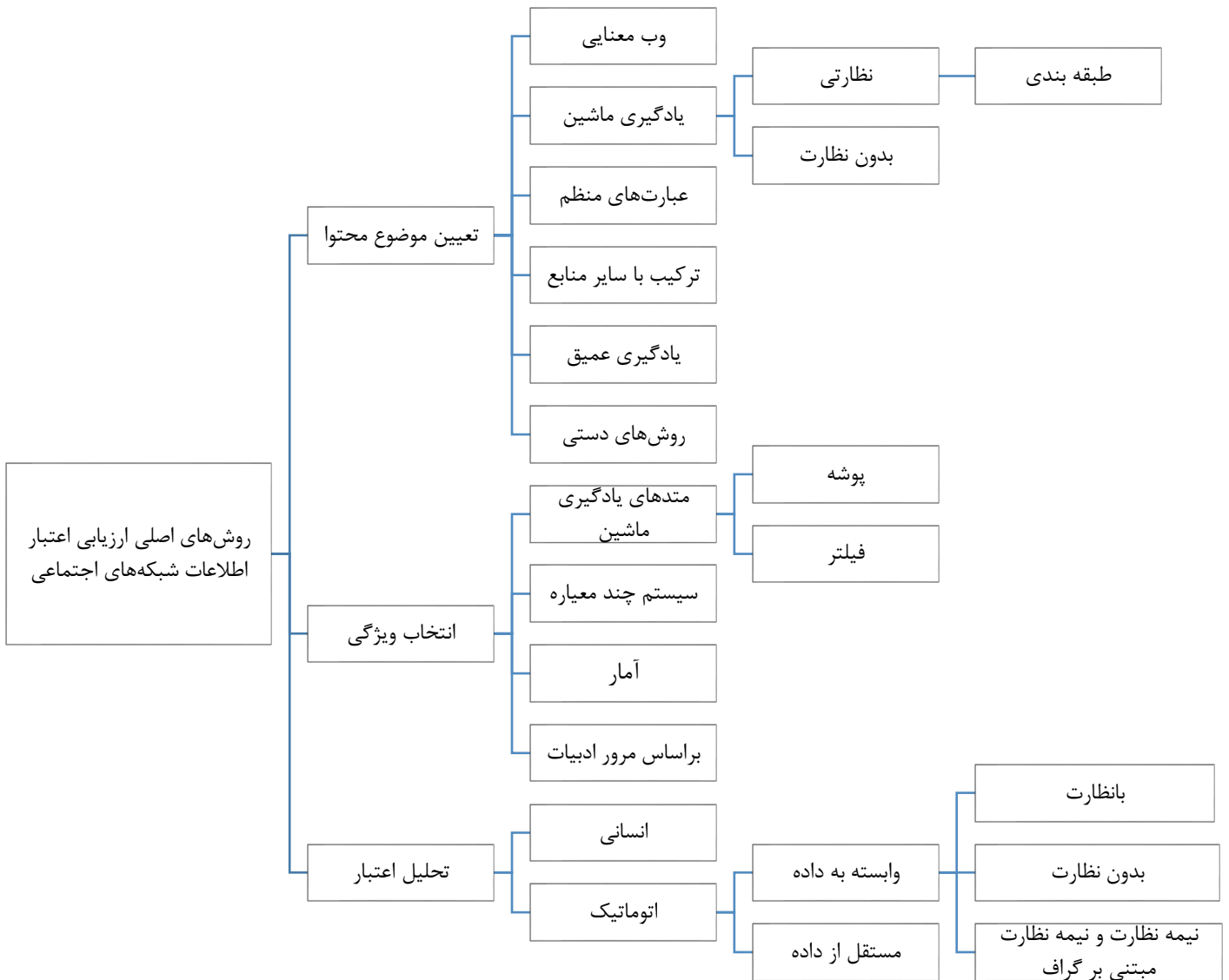
²⁶ Iterative Classification Algorithm (ICA)

²⁰ Page Rank

²¹ association rules

²² Latent Dirichlet Allocation (LDA)

²³ Long short-term memory (LSTM)



شکل ۲. طبقه بندی ۳ مولفه ارزیابی اعتبار اطلاعات در شبکه های اجتماعی

جدول ۱ بررسی مقالات مرتبط به ارزیابی اعتبار اطلاعات شبکه‌های اجتماعی

مرجع	شبکه اجتماعی	زیر مولفه‌ها	ویژگی‌ها	رویکردها	BD	اهداف مقالات
[۱۱]	توییتر	C3, C4	F3, F7	C3 : آمار – C4: انسانی	×	تاثیر اهمیت ویژگی‌ها در حوزه ارزیابی اعتبار محتوا
[۱۶]	فیس‌بوک	C7, C4	F1, F7	C4 / C7 : بانظارت (SVM)	×	برای درک آسان‌تر اعتبار محتوا، به محتوا امتیاز بین ۱ تا ۱۰ اختصاص می‌دهد.
[۲۲]	فیس‌بوک	C1, C4, C7	F1, F2, F3, F5, F7	C1: با استفاده از هشتگ‌های محتوا – C4 / C7 : رویکردهای بانظارت (NB, SVM, RF رگرسیون لجستیک)	×	ارزیابی اعتبار محتوای خبری
[۱۷]	فیس‌بوک	C4, C7	F1, F2, F3, F4, F5, F6, F7	C4 / C7 : بانظارت (SVM)	×	ارزیابی اعتبار کاربران
[۲۳]	توییتر	C1, C4, C7	F2	C1: با استفاده از هشتگ‌های محتوا – C4 / C7 : رویکردهای بانظارت (NB, SVM, RF)	×	بررسی اعتبار محتوا
[۱۸]	توییتر، فیس‌بوک و اینستاگرام	C4, C7	F1, F2, F7	C4 / C7 : بانظارت (LR, SVR, KNN)	×	با استفاده از نتایج ارزیابی کاربران بانفوذ با روش‌های مختلف، تاثیر ویژگی‌ها را در شبکه‌های اجتماعی مختلف محاسبه می‌نماید.
[۲۴]	توییتر و فیس‌بوک	C1, C3, C4	F1, F2, F3, F4, F5	C1: الگوریتم‌های طبقه‌بندی – C3 : افزودن ویژگی‌ها یکی یکی به صورت دستی – C4: الگوریتم‌های بانظارت (NB, SVM, DT رگرسیون لجستیک)	×	بررسی اعتبار محتوا
[۲۵]	توییتر	C1, C4	F1, F2, F4, F7	C1: بدون نظارت (LDA) – C4: روش‌های بانظارت (رگرسیون لجستیک)	×	آن‌ها به بررسی چگونگی تاثیر اعتبار پروفایل بر اعتبار محتوا و باز نشر در هنگام رخدادهای طبیعی پرداخته‌اند.
[۲۶]	توییتر	C1, C3, C4, C7	F3, F5, F7, F2, F1, F6, F4	C1: بررسی دستی – C3 : یادگیری ماشین (فیلتر، پوشه) – C4/C7: بانظارت (DT, NB, LR, RF, KNN)	×	بررسی اعتبار محتوا در حوزه بورس و اوراق بهادار
[۲۷]	توییتر	C1, C4	F2, F7	C1: بررسی دستی – C4: بانظارت (DT)	×	تعیین اعتبار کاربران در حوزه خبری

اهداف مقالات	BD	رویکردها	ویژگی‌ها	زیر مولفه‌ها	شبکه اجتماعی	مرجع
تعیین اعتبار محتوا عربی	×	C1: بدون نظارت (TF-IDF) C3: یادگیری ماشین (پوشه‌هریس هاکس ²⁷) C4: بانظارت (اکس‌جی‌باست، SVM، RF، NB، KNN، DT، LR)	F1, F2, F3, F4, F5, F6, F7	C1, C3, C4	توییتر	[۲۸]
رتبه‌بندی کاربران بانفوذ	×	C1: ابزار وب معنایی آماده C4: بدون نظارت (مشابه رتبه‌بندی صفحه)	F1, F6	C1, C4	توییتر	[۲۹]
رتبه‌بندی کاربران بانفوذ	×	C1: مدل موضوع بیترم (BTM) ²⁸ C4 – رتبه‌بندی صفحه – C2/ C5، FM، CRT	F1, F6	C1, C2, C4, C5	توییتر	[۳۰]
ارزیابی اعتبار کاربران و محتوا خبری	×	C1: ترکیب منابع C4: بدون نظارت مبتنی بر گراف (آی‌تی‌پی توییتر)	F1, F6, F7	C1, C4	توییتر	[۳۱]
بررسی اعتبار کاربران در حوزه کرونا و انتخابات ایالت متحده	✓	C1: ورد۲وک – C4: بدون نظارت (LDA، خوشه‌بندی، قوانین تداعی)	F1, F3, F6, F7	C1, C4	توییتر	[۳۲]
هدف این مقاله کاهش زمان محاسبه ارزیابی اعتبار کاربران و محتوا مرتبط به رخداد‌های طبیعی می‌باشد.	×	C1: هشتگ C4/C7: بانظارت (لاجستیک رگرسیون، SVM، RF، KNN، DT)، بدون نظارت (ابتکاری)	F1, F2, F3, F4, F5, F6, F7	C1, C4, C7	توییتر	[۳۳]
بررسی اعتبار محتوا، کاربران و لینک‌ها	×	C1: عبارات‌های منظم – C4: G- (ICA) SSL	F1, F2, F4, F5, F6, F7	C1, C4	توییتر	[۳۴]
بررسی اعتبار کاربر	×	C3: آماری (ضریب پیرسون)، C4: بانظارت (، KNN، RF، NB، SVM)، بدون نظارت (، LSTM، BERT)	F1, F3, F5, F7	C3, C4, C7	توییتر	[۳۹]
ارزیابی اعتبار محتوا چند رسانه‌ای	×	C4/ C7: نیمه نظارت (رگرسیون لجستیک، DT)	F1, F2, F3, F4, F5, F7	C4, C7	توییتر	[۱۹]

²⁷ Harris Hawks

²⁸ Biterm Topic Model (BTM)

مرجع	شبکه اجتماعی	زیر مولفه‌ها	ویژگی‌ها	رویکردها	BD	اهداف مقالات
[۴۰]	توثیتر	C4, C7	F1, F2, F3, F5	C4/ C7: نیمه نظارت (SVM, RF), MLP	×	ارزیابی اعتبار کاربران در حوزه سیاست
[۳۵]	توثیتر	C1, C4, C5	F1, F4, F7	C1: عبارت‌های منظم - C4: مستقل از داده (میانگین وزنی ویژگی‌ها) - C5: توزیع نمایی	×	ارزیابی اعتبار کاربران در حوزه ورزشی
[۱۴]	وبلاگ	C4, C6	F1, F2, F5, F7	C4: مستقل از داده (میانگین ویژگی‌ها)، C6: روش‌های آماری	×	بررسی اعتبار کاربران
[۳۶]	وبلاگ	C1, C3, C4	F1, F5	C1: دستی C3: مستقل از داده (AHP) C4: مستقل از داده (تاپسیس)	×	تعیین اعتبار کاربران در حوزه گردشگری
[۱۵]	توثیتر	C1, C4, C6, C7	F1, F2, F5, F6, F7	C1: ابزار وب معنایی آماده C4/ C6: مستقل از داده (جمع‌وزنی ویژگی‌ها)	✓	تعیین اعتبار کاربران در ۷ گروه
[۳۷]	یلپ	C1, C3, C4	F1, F2, F3, F4, F6, F7	C1: بررسی دستی - C3/ C4: مستقل از داده (انتگرال فازی چکیت)	×	ارزیابی اعتبار محتوا
[۱۰]	اپینیز ^{۲۹}	C3, C4	F1, F3, F5, F7	C3: AHP فازی C4: ویکور فازی	×	ارزیابی اعتبار کاربران
[۲۰]	توثیتر	C1, C3, C4, C7	F1, F2, F5	C1: هشتم‌ها - C3: براساس مرور ادبیات، بانظارت (RF) - C4/ C7: مستقل از داده (عملگر تجمیع میانگین وزنی مرتب شده (OWA) ^{۳۰})	×	بررسی اعتبار کاربران
[۳۸]	توثیتر	C1, C3, C4	F1, F2, F4	C1: هشتم - C3: روش‌های آماری C4: وابسته به داده (ابتکاری)	×	ارزیابی اعتبار کاربران

با توجه به شکاف‌های تحقیقاتی مذکور، در این پژوهش سعی شده است، از طریق ارائه یک چارچوب اولیه برای اهداف سازمانی، مولفه‌های اصلی ارزیابی اعتبار کاربران (منابع اطلاعاتی) معرفی شوند و فرآیند صحیح ارزیابی اعتبار کاربران روشن گردد. به این منظور، در قدم اول، با استفاده از مرور ادبیات، مولفه‌های ارزیابی اعتبار اطلاعات شناسایی،

جمع‌آوری و بررسی گردیده است. در قدم دوم، چارچوب پیشنهادی براساس زیرمولفه‌ها، چرخه BD و اهداف سازمانی شکل گرفته است، زیرا این شبکه‌ها پویا می‌باشند و با سرعت و حجم بالا، داده تولید می‌نمایند و تحلیل اعتبار اطلاعات حساس به زمان می‌باشد. همچنین، در کنار چارچوب

²⁹ Epinions

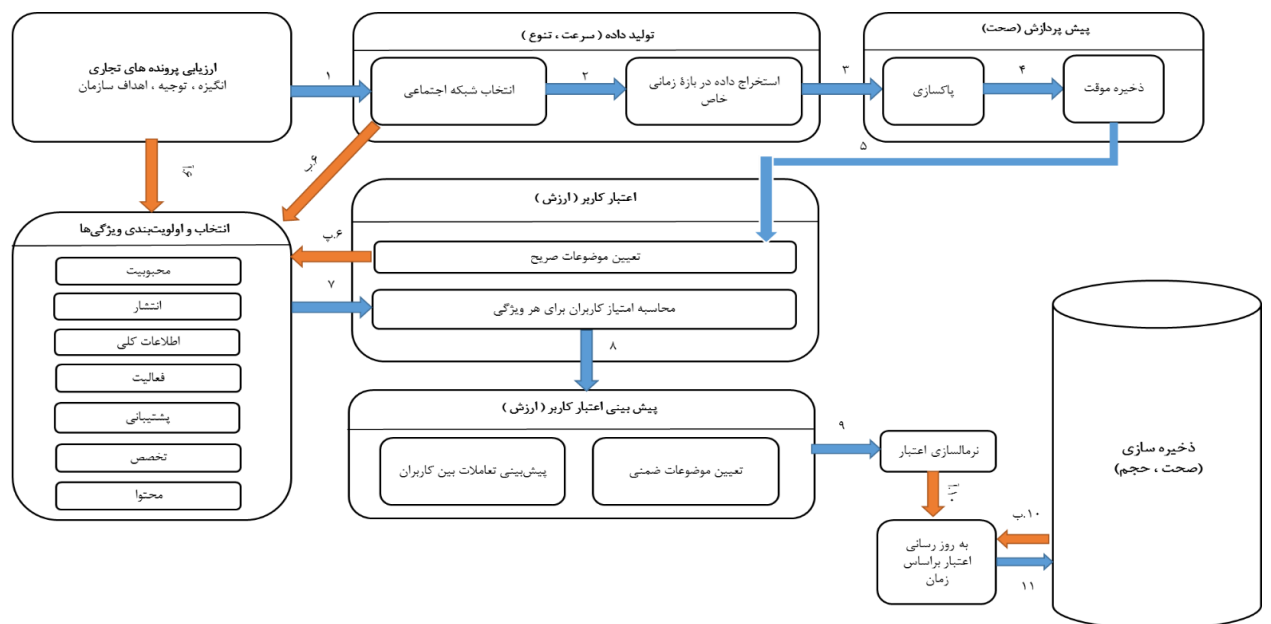
³⁰ Ordered Weighted Averaging (OWA)

پیشنهادی، یک ساختار جهت انتخاب ویژگی‌ها براساس اهداف متفاوت ارزیابی اعتبار کاربران ارائه شد.

۴. چارچوب اولیه پیشنهادی جهت تحلیل اعتبار کاربران شبکه‌های اجتماعی بر اساس اهداف سازمانی

همان‌طور که اشاره شد، علی‌رغم اهمیت ارزیابی اعتبار اطلاعات شبکه‌های اجتماعی، به دلیل تنوع و پراکندگی زیاد این حوزه، به مولفه‌های مهم توجه کمی داشته‌اند. از این رو، این پژوهش براساس مرور ادبیات صورت گرفته، به شناسایی مولفه‌های مهم در این حوزه می‌پردازد. در نهایت براساس زیرمولفه‌ها (موضوع فعالیت صریح، موضوع فعالیت ضمنی، سیستم انتخاب ویژگی، محاسبه اعتبار، پیش‌بینی اعتبار در آینده، بررسی اعتبار در گذر زمان، قابلیت فهم آسان نتایج) و اهداف متفاوت سازمانی یک چارچوب اولیه جهت ارزیابی اعتبار کاربران پیشنهاد می‌شود و با توجه به این‌که، تحلیل اعتبار کاربران به دلیل سرعت و حجم بالا اطلاعات، حساس به

زمان می‌باشند، این چارچوب براساس چرخه حیات BD، در شکل ۳ در ۹ مرحله طراحی گردیده است. براساس کتاب ایرل و همکاران (۲۰۱۶)، چرخه حیات BD شامل ۹ گام ارزیابی پرونده تجاری، شناسایی داده‌ها، جمع‌آوری و فیلتر کردن داده‌ها، استخراج داده‌ها، اعتبارسنجی و پاکسازی داده‌ها، جمع‌آوری و نمایش داده‌ها، تجزیه و تحلیل داده‌ها، تجسم داده‌ها و استفاده از نتایج تجزیه و تحلیل می‌باشد [۳]. چارچوب اولیه این مقاله برای ارزیابی اعتبار کاربران شبکه‌های اجتماعی به عنوان منابع اطلاعاتی، در فرآیند تصمیم‌گیری سازمان‌ها مورد استفاده قرار می‌گیرد. از طرف دیگر، در این حوزه روش مناسبی جهت انتخاب ویژگی شبکه‌های اجتماعی ارائه نگردیده است. از این رو، در کنار این چارچوب، یک ساختار سلسله‌مراتب در ۴ سطح برای انتخاب ویژگی، با توجه به هدف ارزیابی اعتبار کاربران در سازمان‌ها ارائه گردید.



شکل ۳. چارچوب پیشنهادی تحلیل اعتبار کاربران در شبکه‌های اجتماعی براساس اهداف سازمانی

۱.۴ ارزیابی اعتبار براساس اهداف متنوع

یک سازمان براساس اهدافش، کاربران معتبر را تعریف می‌نماید، به این منظور، پیش از ارزیابی اعتبار کاربران، باید اهداف ارزیابی به روشنی بیان شود. پنج سوال در فاز برنامه‌ریزی استراتژیک سازمان مطرح می‌باشد [۳]. هر سازمان باید در مراحل متفاوت از این چارچوب، به این سوالات پاسخ دهد. تا براساس اهداف سازمانی، کاربران معتبر را انتخاب نماید. در هر بخش از این چارچوب، راهنمایی جهت پاسخ به این سوالات ارائه شده است.

Q1: هدف از ارزیابی اعتبار کاربران شبکه‌های اجتماعی چیست؟

Q2: نتیجه این ارزیابی تا چه اندازه برای سازمان مفید می‌باشد؟ آیا سود آن بیشتر از هزینه‌ها می‌باشد؟

Q3: کدام یک از شبکه‌های اجتماعی متناسب با اهداف سازمان می‌باشند؟

Q4: چه نوع کاربران معتبری (بانفوذ، پشتیبان، متخصص) برای رسیدن به هدف سازمان مناسب می‌باشند؟

Q5: چه ویژگی‌هایی از شبکه‌های اجتماعی براساس نوع کاربر مناسب می‌باشند؟

در صورتی که جواب سوال Q1 و Q2 قانع‌کننده باشد، فرآیند ارزیابی اعتبار کاربران در شبکه‌های اجتماعی شروع خواهد شد.

مقالات ارزیابی اعتبار کاربران شبکه‌های اجتماعی، براساس اهداف متفاوت، تعاریف متنوعی از کاربران معتبر ارائه نموده‌اند. این پژوهش به استناد تعاریف موجود در این حوزه، کاربران معتبر را به سه نوع مختلف تقسیم نموده است:

بانفوذ: در بین سایر کاربران محبوبیت فراوانی دارند و دیگران را تحت تاثیر قرار می‌دهند [۱۲]، [۱۳]، [۲۹]، [۳۰]. جهت اهدافی مانند بازاریابی، افزایش شهرت تجاری سازمان و بهبود وفاداری به برند مناسب می‌باشند.

پشتیبان: مصرف‌کننده واقعی خدمات می‌باشند [۱۲]، [۱۳]، [۱۴]. این کاربران با ارائه بازخورد نسبت به موضوعات مورد نظر سازمان،

جهت اهدافی مانند حل مشکلات سازمان در برابر پارادوکس‌ها و بهبود خدمات، مناسب می‌باشند.

متخصص: کاربرانی که در موضوعات خاص متخصص می‌باشند [۱۲]، [۲۶]، [۳۱]، [۳۲]. آن‌ها جهت اهدافی مانند استخدام نیرو، افزایش آگاهی سازمان و مدیریت بحران مناسب هستند.

پاسخ به سوال Q4 براساس این سه نوع کاربر می‌باشد و احتمالاً ترکیبی از آن‌ها، برای دستیابی به اهداف سازمان مناسب باشد، به طور مثال، جهت افزایش خلاقیت، هر سه نوع کاربر مورد توجه سازمان می‌باشد، اما میزان اهمیت هر یک از آن‌ها متفاوت می‌باشد [۸].

۲.۴ شناسایی و اکتساب داده

با توجه به شبکه اجتماعی انتخاب شده در سوال Q3 با رابط برنامه نویسی کاربردی^{۳۱} آن شبکه، داده‌ها در یک بازه زمانی خاص استخراج استخراج می‌شوند. این مرحله نشان دهنده ویژگی سرعت و تنوع BD می‌باشد. از این رو، در یک بازه زمانی کوتاه حجم زیادی از داده‌ها با فرمت‌های متفاوت (تصویر، فیلم، محتوا و صوت) جمع‌آوری می‌شود [۳].

۳.۴ پیش پردازش

به ویژگی صحت BD اشاره دارد. داده‌ها به دو دسته سیگنال و نویز قابل تقسیم می‌باشند. نویزها توانایی تبدیل به اطلاعات را ندارند. از این رو، برای ایجاد ارزش بیشتر و افزایش صحت اطلاعات، نویزها باید حذف شوند [۳]. حذف کلمات توقف (ضمایر، حروف ربطی و ...)، حذف زبان‌هایی که مورد مطالعه نیستند، ریشه‌یابی کلمات^{۳۲} و ریشه‌یابی کلمات عطف^{۳۳} جز این مراحل می‌باشند [۱۷].

۴.۴ اعتبار کاربر

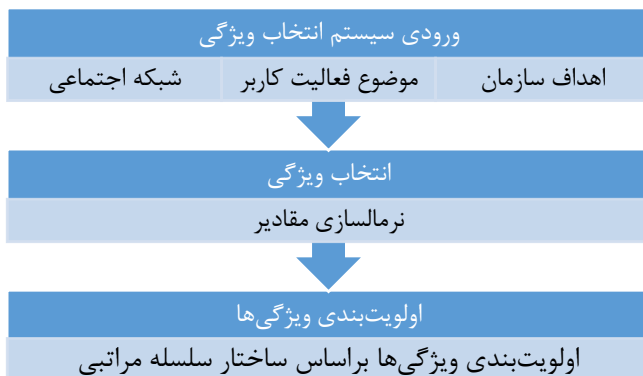
پس از جمع‌آوری، پیش‌پردازش و ذخیره‌سازی موقت داده‌ها، نوبت به تجزیه و تحلیل آن‌ها برای محاسبه اعتبار کاربر می‌رسد. این مرحله به ویژگی ارزش BD اشاره دارد، یعنی تبدیل داده‌ها به اطلاعات

³³ Lemmatizer

³¹ Application Programming Interface (API)

³² Stemmer

۴.۵ سیستم انتخاب ویژگی



شکل ۴. معماری سیستم انتخاب و اولویت‌بندی به ویژگی‌ها

انتخاب ویژگی از الزامات پایه این چارچوب می‌باشد، عدم انتخاب صحیح ویژگی‌های شبکه‌های اجتماعی، مانع از ارزیابی اعتبار کاربران با دقت بالا می‌گردد. علاوه بر این، جعل نیز امکان پذیر می‌شود و حتی ممکن است، برخی ویژگی‌ها در موضوعات متفاوت نشان‌دهنده اعتبار یا اسپم باشند [۲۵]. براساس مرور ادبیات، سیستم انتخاب ویژگی، جهت ارائه بهترین ویژگی‌ها به سه ورودی به شرح اهداف سازمان، شبکه اجتماعی و موضوع فعالیت کاربران نیاز دارد. در شکل ۴ معماری این سیستم مشاهده می‌شود. ورودی اول، اهداف سازمان می‌باشد، که براساس پاسخ به سوال Q1، مشخص می‌شود. ورودی دوم، شبکه اجتماعی می‌باشد. ویژگی‌ها به نوع شبکه اجتماعی وابسته می‌باشند، یعنی ویژگی‌های اعتبار در توئیتر با ویژگی‌های اعتبار در فیس‌بوک یکسان نمی‌باشند [۱۸]. ورودی سوم موضوع فعالیت کاربران می‌باشد. به طور مثال در برخی از موضوعات نرخ بازنشر پست‌ها، ویژگی قدرتمندی برای ارزیابی اعتبار کاربران می‌باشد. اما در سایر موضوعات این ویژگی فاقد اهمیت می‌باشد [۲۹].

با توجه به تنوع بالای ویژگی‌های شبکه‌های اجتماعی، برای انتخاب و اولویت‌بندی ساده‌تر آن‌ها، یک ساختار سلسله مراتب در شکل ۵ براساس ۴ سطح پیشنهاد گردیده است. در این ساختار، امتیاز هر سطح با توجه به سطح بالاتر، محاسبه می‌شود. سطح صفر، هدف این سیستم یعنی انتخاب بهترین ویژگی‌ها می‌باشد. سطح اول به بررسی هدف ارزیابی اعتبار کاربران می‌پردازد. سطح دوم شامل سه نوع کاربر معتبر (بانفوذ، پشتیبان، متخصص) می‌باشد و امتیاز هر کاربر براساس سطح اول تعیین می‌گردد. سطح سوم، شامل گروه‌های ویژگی شبکه

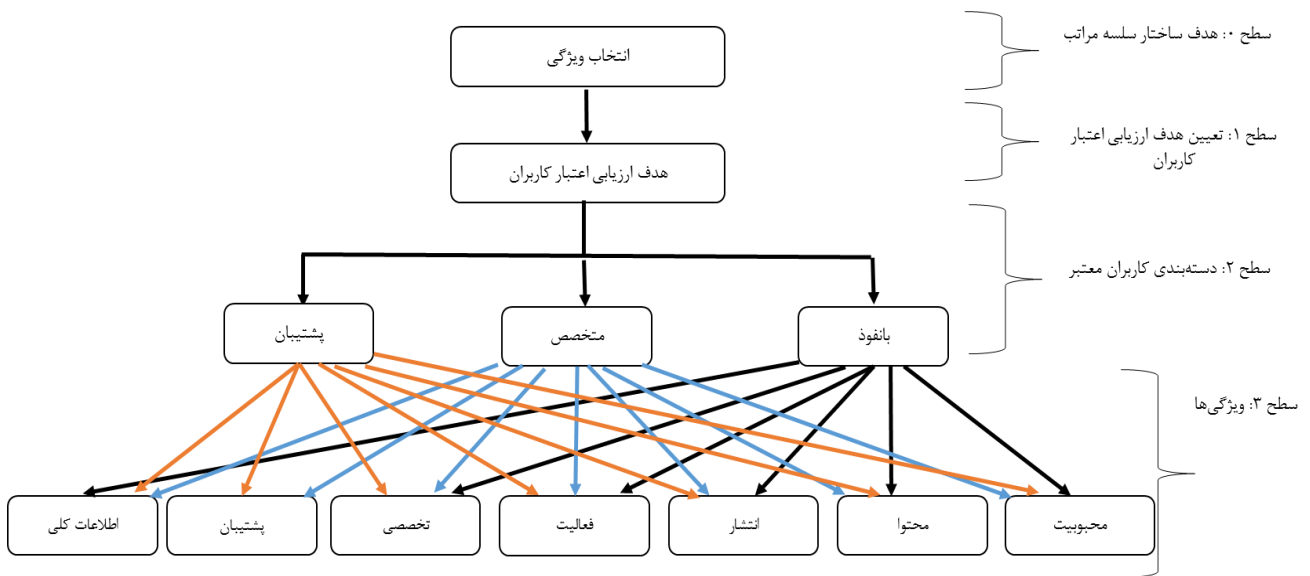
معنادار [۳]. اعتبار کاربر در این بخش طی ۲ مرحله، تعیین فعالیت صریح و امتیاز در هر ویژگی، محاسبه می‌گردد.

۴.۴.۱ تعیین موضوعات صریح

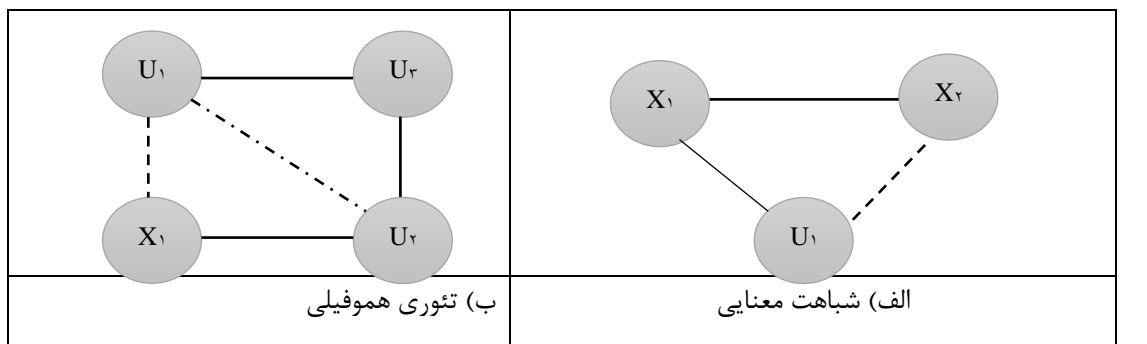
اعتبار کاربران شبکه‌های اجتماعی، باید با توجه به موضوع فعالیت آن‌ها بررسی گردد. زیرا کاربران، در تعداد محدودی از موضوعات معتبر می‌باشند [۴]. از طرف دیگر، موضوع فعالیت کاربران در این شبکه‌ها نامشخص می‌باشد. به این دلیل، پیش از ارزیابی اعتبار کاربران باید موضوع فعالیت آن‌ها مشخص گردد. تعیین موضوع فعالیت کاربران پایه و اساس این چارچوب می‌باشد و به دو زیرمولفه صریح و ضمنی قابل تقسیم می‌باشد. جهت تشخیص صحیح موضوع فعالیت کاربران، باید هر دو گروه بررسی شوند. موضوع فعالیت کاربران با بررسی محتوای تولید شده توسط آن‌ها، مشخص می‌گردد. روش‌های تشخیص موضوعات فعالیت صریح براساس مرور ادبیات، در شکل ۲ معرفی گردید، که انتخاب این روش‌ها به نوع شبکه اجتماعی وابسته می‌باشد [۵].

۴.۴.۲ محاسبه امتیاز کاربران برای هر ویژگی

قبل از شروع عملیات این بخش، تعدادی از ویژگی‌های شبکه اجتماعی جهت ارزیابی اعتبار کاربران انتخاب می‌شوند و امتیاز کاربران برای هر یک از این ویژگی‌ها محاسبه می‌گردد. از طرف دیگر، جعل برخی از این ویژگی‌ها برای کاربران اسپم راحت می‌باشد، اما جعل تعداد زیادی از آن‌ها تقریباً غیرممکن است [۲۵]. از این رو، انتخاب صحیح تعداد زیادی از ویژگی‌های موثر، متناسب با هدف ارزیابی اعتبار کاربران بسیار مهم می‌باشد. بنابراین قبل از انجام عملیات این بخش، سیستم دقیقی جهت انتخاب و اولویت‌بندی ویژگی‌های شبکه اجتماعی ارائه گردیده است.



شکل ۵. ساختار سلسله مراتبی انتخاب و اولویت بندی به ویژگی ها



شکل ۶. پیش بینی اعتبار کاربران براساس شباهت معنایی و تئوری هموفیلی

۷.۴ نرمالسازی اعتبار

نرمالسازی اعتبار، به معنی تبدیل میزان اعتبار کاربران به مقادیر قابل فهم، برای عموم افراد می باشد. به این منظور، یک سازمان برای تحلیل نتایج به نیروی متخصص نیاز نخواهد داشت [۳۵].

۸.۴ ذخیره سازی

ذخیره سازی از الزامات پایه ای چارچوب می باشد، که جهت حفظ صحت و ذخیره حجم عظیمی از داده ها باید انبار داده مناسب استفاده شود.

۹.۴ به روزرسانی اعتبار

به دلیل ماهیت ناشناس و سریع ارتباطات آنلاین، اعتماد به کاربران پویا و گذرا می باشد [۱۲]. همچنین، معمولاً کاربران اسپم، در بازه زمانی کوتاهی، خود را معتبر معرفی می نمایند. به این منظور، باید اطلاعات کاربران در بازه های زمانی متفاوت، جمع آوری و تحلیل گردد و براین اساس نتایج ارزیابی آنان، به روزرسانی خواهد شد [۱۵].

۵. بحث

تجاری، شناسایی داده‌ها، جمع آوری و فیلتر کردن داده‌ها، استخراج داده‌ها، اعتبارسنجی و پاکسازی داده‌ها، جمع آوری و نمایش داده‌ها، تجزیه و تحلیل داده‌ها، تجسم داده‌ها و استفاده از نتایج تجزیه و تحلیل می‌باشد [۳].

- ارائه یک تعریف روشن از هدف ارزیابی اعتبار اطلاعات، بخصوص در هنگام اعتبارسنجی کاربران به عنوان منابع اطلاعاتی، مهم می‌باشد. زیرا براساس هدف ارزیابی، تعاریف متفاوتی از کاربران معتبر در ادبیات این حوزه ارائه شده است. این تعاریف منجر به انتخاب و اولویت‌بندی متفاوت ویژگی‌های شبکه‌های اجتماعی می‌شود. در تعدادی از مقالات، بررسی اعتبار کاربران و میزان نفوذ آن‌ها، به‌عنوان دو حوزه تحقیقاتی متفاوت از یکدیگر معرفی شده‌اند. از طرفی دیگر، در برخی از مقالات، کاربران معتبر به دو گروه بانفوذ و پشتیبان تقسیم می‌شوند و تعدادی دیگر، کاربران معتبر را افراد متخصص یک حرفه، می‌دانند [۱۲]، [۱۳]، [۱۵]، [۲۵]، [۳۲]، [۴۰]. در این راستا، پژوهش فعلی معتقد است، که کاربران معتبر براساس اهداف متنوع ارزیابی اعتبار کاربران در سازمان‌ها تعریف می‌شود. به این منظور کاربران معتبر به سه گروه بانفوذ، پشتیبان و متخصص قابل تقسیم می‌باشند.

- با توجه به حجم زیاد اطلاعات، انتخاب یک انبارداده مناسب، جهت حفظ صحت داده‌ها ضروری می‌باشد [۴]، [۱۵].

- کاربران در تعداد محدودی از موضوعات معتبر می‌باشند [۴]، [۱۵]. همچنین برخی از کاربران معتبر، علاقه به تولید محتوا در این شبکه‌ها ندارند [۴۱]. به این منظور، استفاده از یک روش مناسب، جهت تعیین موضوع فعالیت آنان به صورت صریح و ضمنی، مهم می‌باشد.

- شبکه‌های اجتماعی شامل ویژگی‌های متنوعی مانند تعداد بازخوردهای مثبت به محتوا، سن حساب کاربری و تعداد فالورها می‌باشند. انتخاب و اولویت‌بندی این ویژگی‌ها با توجه به اهداف سازمانی به سه عامل، هدف ارزیابی اعتبار کاربران، موضوع فعالیت آن‌ها و نوع شبکه‌های اجتماعی مرتبط می‌باشد [۱۸]، [۲۰]، [۲۸]، [۲۹]، [۳۵]. بنابراین،

شبکه‌های اجتماعی، به دلیل سرعت انتشار و نیز حجم بالا اطلاعات به عنوان یکی از مهم‌ترین منابع اطلاعاتی خارجی در سازمان‌ها استفاده می‌شوند. این شبکه‌ها به دلیل عدم نظارت، احتمالاً شامل اطلاعات ناصحیح و نامعتبر می‌باشند. به این منظور پیش از بهره‌برداری از اطلاعات شبکه‌های اجتماعی، ارزیابی اعتبار آن‌ها بسیار پراهمیت می‌باشد. اعتبار اطلاعات در این شبکه‌ها با بررسی ویژگی‌های آن‌ها مانند تعداد فالورها و سن حساب کاربری قابل ارزیابی می‌باشد. بهترین نوع ارزیابی اعتبار اطلاعات در این شبکه‌ها، بررسی اعتبار کاربران است، زیرا علاوه بر ویژگی‌های کاربران، ویژگی‌های محتوای منتشر شده نیز، بررسی می‌گردد.

به دلیل اهمیت روز افزون شبکه‌های اجتماعی و نیز تمایل سریع کاربران نامعتبر به تغییر استراتژی‌ها در پاسخ به اقدامات دفاعی، تحقیقات مرتبط به حوزه ارزیابی اعتبار اطلاعات شبکه‌های اجتماعی در حال افزایش می‌باشد، که این تنوع تحقیقاتی، موجب پراکندگی در مقالات حوزه ارزیابی اعتبار اطلاعات گردیده است. از این رو، مقالات این حوزه از برخی از مولفه‌های اساسی ارزیابی اعتبار اطلاعات، چشم پوشی نموده‌اند. همچنین فرآیند درک اعتبار اطلاعات در این شبکه‌ها مبهم می‌باشند [۱۲]. براساس مرور ادبیات صورت گرفته در این پژوهش، این مولفه‌ها به ۳ مولفه اصلی تعیین موضوع محتوا، روش‌های انتخاب ویژگی، ارزیابی اعتبار قابل تقسیم می‌باشند. براساس یافته‌های این پژوهش نکاتی که باید در فرآیند ارزیابی اعتبار اطلاعات شبکه‌های اجتماعی توجه کرد، به شرح ذیل می‌باشد:

- چارچوب ارزیابی اعتبار کاربران باید براساس چرخه حیات BD باشد. زیرا تحلیل اعتبار اطلاعات به صورت آنلاین جهت تحقق اهداف سازمانی، امکان‌پذیر می‌گردد [۴]، [۱۲]، [۱۵]، [۳۲]. چارچوب پیشنهادی، این پژوهش براساس چرخه حیات کلان داده، [۳] طراحی شده است. به صورت کلی این چرخه، شامل ۹ گام ارزیابی پرونده

می‌باشد [۱۴]، [۱۸]، [۲۰]، [۲۸]، [۲۹]، [۳۵]. از این رو، در کنار چارچوب اولیه، یک ساختار جهت انتخاب و اولویت‌بندی ویژگی‌ها پیشنهاد گردید.

به این ترتیب، منبع استخراج مولفه‌های چارچوب پیشنهادی و روش توسعه آن معرفی گردید. این مقاله با بازنگری مختصری از مقالات حوزه ارزیابی اعتبار اطلاعات شبکه‌های اجتماعی به یک چارچوب اولیه جهت ارزیابی اعتبار کاربران در جهت اهداف سازمانی دست یافت. به این منظور، این پژوهش راهنمای مناسب برای اندازه‌گیری اعتبار کاربران (منابع اطلاعاتی) در فرآیند تصمیم‌گیری سازمان‌ها می‌باشد.



شکل ۷. اهمیت زیرمولفه‌ها در مقالات

۶. نتیجه‌گیری

در این پژوهش تلاش شد، که براساس روش‌ها و مفاهیم موجود در ادبیات حوزه ارزیابی اعتبار اطلاعات شبکه‌های اجتماعی در سطوح متفاوت، یک چارچوب اولیه جهت ارزیابی اعتبار کاربران براساس چرخه حیات کلان داده و اهداف متنوع سازمانی ارائه شود، که نمونه این چارچوب در کارهای پیشین دیده نشده است. برای این منظور ۵۰ مقاله مرتبط به حوزه ارزیابی اعتبار اطلاعات شبکه‌های اجتماعی در سطوح متفاوت، بررسی شد. براساس یافته‌های این پژوهش:

- فرآیند ارزیابی اعتبار کاربران شبکه‌های اجتماعی، تابعی از مولفه‌های مهم می‌باشند، که این مولفه‌ها به ۳ مولفه اصلی (تعیین موضوع محتوا، انتخاب ویژگی و ارزیابی اعتبار

در این پژوهش با توجه به این ۳ عامل یک ساختار سلسله‌مراتبی، برای انتخاب و اولویت‌بندی ویژگی‌ها پیشنهاد گردیده است. عدم انتخاب صحیح ویژگی‌ها، موجب ارزیابی نادرست کاربران شبکه‌های اجتماعی می‌شود [۲۵].

- زمان عامل مهمی در فرآیند ارزیابی اعتبار کاربران می‌باشد. معمولاً کاربران اسپم، در یک بازه زمانی محدود، معتبر می‌باشند [۲۵]. به این منظور اعتبار کاربران باید در طی گذر زمان محاسبه گردد [۱۲]، [۱۵]، [۳۵].

- برخی از کاربران، در حال حاضر در تعدادی از موضوعات فعالیتی ندارند، اما در آینده به افراد تاثیرگذاری در آن موضوعات تبدیل خواهند شد [۳۰]، [۴۱]. در این راستا، پیش‌بینی اعتبار کاربران در آینده، در موضوعات متنوع مهم می‌باشد.

- نتایج ارزیابی اعتبار اطلاعات در شبکه‌های اجتماعی، باید برای عموم قابل فهم باشد. زیرا معمولاً در سازمان‌ها افراد غیرمتخصص، نتایج را مورد بررسی قرار می‌دهند [۳۵].

چارچوب پیشنهادی این مقاله، با تمرکز بر ارزیابی اعتبار کاربران، براساس نکات مذکور ارائه گردیده است. شکل ۷ میزان استفاده از زیر مولفه‌های را نمایش می‌دهد. براساس این شکل، مقالات این حوزه، به زیرمولفه تعیین موضوع فعالیت کاربران به صورت صریح، بیش‌ترین توجه را داشته‌اند. به این منظور نتیجه‌گیری می‌شود، که بیش‌تر مقالات این حوزه، به این آگاهی دست یافته‌اند، که تعیین اعتبار کاربران بدون در نظر گرفتن موضوع فعالیت آنان فاقد ارزش می‌باشد. همچنین این مقالات به پیش‌بینی اعتبار کاربر در آینده، تعیین موضوع فعالیت کاربر به صورت ضمنی و بررسی اعتبار در گذر زمان توجه کم‌تری داشته‌اند. همچنین، انتخاب ویژگی یکی از مولفه‌های مهم ارزیابی اعتبار اطلاعات می‌باشد. از طرف دیگر، بیش‌تر مقالات این حوزه، با استفاده از مرور ادبیات و بدون تحلیل تاثیر ویژگی‌ها بر نتایج ارزیابی اعتبار اطلاعات، ویژگی‌ها را انتخاب نموده‌اند و تعداد مقالات کمی اشاره داشته‌اند، که انتخاب ویژگی به موضوع فعالیت کاربران، هدف سازمان از ارزیابی اعتبار اطلاعات و نوع شبکه اجتماعی وابسته

و گذارا است. از این رو، چارچوب ارزیابی اعتبار کاربران باید براساس چرخه حیات کلان داده طراحی شود.

از آن جهت که هر کار پژوهشی محدودیتهایی دارد، در این مقاله نیز امکان پیاده‌سازی تمام مولفه‌ها در سازمان‌ها با اهداف متفاوت غیرممکن می‌باشد، بنابراین تنها مفروضات براساس مرورادبیات در نظر گرفته شده است. اما پیشنهاد می‌گردد تا در پژوهش‌های آتی چارچوب پیشنهادی بر روی یک شبکه اجتماعی پویا در دنیای واقعی با تمرکز بر روی یک هدف سازمانی پیاده‌سازی شود. همچنین با استفاده از ترکیب ابزارهای هوش مصنوعی و سیستم تصمیم‌گیری چند معیاره یک سیستم انتخاب ویژگی ارزیابی اعتبار کاربران براساس اهداف متفاوت سازمانی پیاده‌سازی گردد. نتایج این چارچوب پیشنهادی می‌تواند، به سازمان‌ها کمک نماید از طریق شناسایی و مدیریت کاربران معتبر، به منابع معتبر اطلاعاتی قابل اطمینان دست یابند و تصمیمات بهتری اتخاذ نمایند. همچنین نتایج این مقاله به عنوان پایه‌ای برای تحقیقات آتی می‌باشد و فرآیند درک اعتبار اطلاعات در این شبکه‌ها را روشن می‌نماید.

Chan, K.Y., Alhabashneh, M., Albtoush, T., Alqahtani, S., Alqahtani, A. and Alahmari, M., 2019, March. Social credibility incorporating semantic analysis and machine learning: a survey of the state-of-the-art and future research directions. In Workshops of the International Conference on Advanced Information Networking and Applications (pp. 887-896). Springer, Cham.

[5] Alrubaian, M., Al-Qurishi, M., Alamri, A., Al-Rakhmi, M., Hassan, M.M. and Fortino, G., 2018. Credibility in online social networks: A survey. *IEEE Access*, 7, pp.2828-2855.

[6] Al-Garadi, M.A., Varathan, K.D., Ravana, S.D., Ahmed, E., Mujtaba, G., Khan, M.U.S. and Khan, S.U., 2018. Analysis of online social network connections for identification of influential users: Survey and open research

اطلاعات) قابل تقسیم می‌باشند، که شامل کل فرآیند ارزیابی اعتبار کاربران از ابتدا تا انتها می‌باشد.

- مولفه تعیین موضوع شامل صریح و ضمنی می‌باشد. مولفه انتخاب ویژگی شامل زیرمولفه، سیستم انتخاب ویژگی است و مولفه ارزیابی اعتبار شامل زیرمولفه‌های، محاسبه اعتبار، پیش‌بینی اعتبار در آینده، بررسی اعتبار در گذر زمان، قابلیت فهم آسان نتایج می‌باشد.

- کاربران معتبر شبکه‌های اجتماعی، متناسب با اهداف سازمانی شامل کاربران متخصص، پشتیبان و بانفوذ می‌باشند و ویژگی‌های آن‌ها براساس این اهداف به ۷ گروه به شرح اطلاعات کلی، فعالیت، محتوا، انتشار، پشتیبان، محبوبیت و تخصص قابل تقسیم می‌باشد. از این رو، در این چارچوب، یک ساختار برای انتخاب ویژگی‌ها پیشنهاد شده است، زیرا انتخاب و اولویت‌بندی ویژگی‌ها باید با اهداف متعدد سازمانی سازگار باشد.

- شبکه‌های اجتماعی پویا می‌باشند و با سرعت و حجم بالا داده تولید می‌نمایند و تحلیل اعتبار کاربر حساس به زمان

مراجع

[1] روجیار پیرمحمدیانی، شهریار محمدی، معیارهای ارزیابی ارزش اثرگذاری کاربران رسانه‌های اجتماعی چارچوبی براساس کاوش رسانه‌های اجتماعی، فناوری اطلاعات و ارتباطات ایران، سال یازدهم شماره ۳۹ و ۴۰، صفحه ۱۰۹-۱۲۵، ۱۳۹۸.

[2] Peng, S., Zhou, Y., Cao, L., Yu, S., Niu, J. and Jia, W., 2018. Influence analysis in social networks: A survey. *Journal of Network and Computer Applications*, 106, pp.17-32.

[3] Erl, T., Khattak, W. and Buhler, P., 2016. Big data fundamentals: concepts, drivers & techniques (Vol. 1). Boston: Prentice Hall.

[4] Abu-Salih, B., Bremie, B., Wongthongtham, P., Duan, K., Issa, T.,

- [14] Weerkamp, W. and de Rijke, M., 2012. Credibility-inspired ranking for blog post retrieval. *Information retrieval*, 15(3-4), pp.243-277.
- [15] Abu-Salih, B., Wongthongtham, P., Chan, K.Y. and Zhu, D., 2019. CredSaT: Credibility ranking of users in big social data incorporating semantic analysis and temporal factor. *Journal of Information Science*, 45(2), pp.259-280.
- [16] Saikaew, K.R. and Noyunsan, C., 2015. Features for measuring credibility on facebook information. *International Scholarly and Scientific Research & Innovation*, 9(1), pp.174-177.
- [17] Afify, E., Sharaf Eldin, A., E Khedr, A. and Kamal Alsheref, F., 2019. User-generated content (UGC) credibility on social media using sentiment classification. *النشرة المعلوماتية في الحاسبات والمعلومات*, 1(1), pp.1-19.
- [18] Arora, A., Bansal, S., Kandpal, C., Aswani, R. and Dwivedi, Y., 2019. Measuring social media influencer index-insights from Facebook, Twitter and Instagram. *Journal of Retailing and Consumer Services*, 49, pp.86-101.
- [19] Boididou, C., Papadopoulos, S., Zampoglou, M., Apostolidis, L., Papadopoulou, O. and Kompatsiaris, Y., 2018. Detection and visualization of misleading content on Twitter. *International Journal of Multimedia Information Retrieval*, 7(1), pp.71-86.
- [20] Pasi, G., De Grandis, M. and Viviani, M., 2020, July. Decision making over multiple criteria to assess news credibility in microblogging sites. In *2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)* (pp. 1-8). IEEE.
- [21] Hershkovitz, A. and Hayat, Z., 2020. The role of tie strength in assessing issues. *ACM Computing Surveys (CSUR)*, 51(1), pp.1-37.
- [7] Keller, J., Wong, S.S. and Liou, S., 2020. How social networks facilitate collective responses to organizational paradoxes. *Human Relations*, 73(3), pp.401-428.
- [8] Olszak, C.M., Bartuś, T. and Lorek, P., 2018. A comprehensive framework of information system design to provide organizational creativity support. *Information & Management*, 55(1), pp.94-108.
- [9] Siciliano, M.D. and Thompson, J.R., 2018. If you are committed, then so am I: The role of social networks and social influence on organizational commitment. *Administration & Society*, 50(7), pp.916-946.
- [10] Abbasimehr, H., Nourani, E. and Shabani, M., 2020. A hybrid framework for ranking reviewers based on interval type-2 fuzzy AHP and VIKOR. *International Journal of Intelligent Engineering Informatics*, 8(2), pp.95-116.
- [11] Morris, M.R., Counts, S., Roseway, A., Hoff, A. and Schwarz, J., 2012, February. Tweeting is believing? Understanding microblog credibility perceptions. In *Proceedings of the ACM 2012 conference on computer supported cooperative work* (pp. 441-450).
- [12] Alkhamees, M., Alsaleem, S., Al-Qurishi, M., Al-Rubaian, M. and Hussain, A., 2021. User trustworthiness in online social networks: A systematic review. *Applied Soft Computing*, 103, p.107159.
- [13] Tabassum, S., Pereira, F.S., Fernandes, S. and Gama, J., 2018. Social network analysis: An overview. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(5), p.e1256.

analysis in microblogging platforms—a topic-sensitive based approach. *Semantic Web*, 5(5), pp.357-372.

[30] Hamzehei, A., Wong, R.K., Koutra, D. and Chen, F., 2019. Collaborative topic regression for predicting topic-based social influence. *Machine Learning*, 108(10), pp.1831-1850

[31] Zhao, L., Hua, T., Lu, C.T. and Chen, R., 2016. A topic-focused trust model for Twitter. *Computer Communications*, 76, pp.1-11.

[32] Diaz-Garcia, J.A., Ruiz, M.D. and Martin-Bautista, M.J., 2022. NOFACE: A new framework for irrelevant content filtering in social media according to credibility and expertise. *Expert Systems with Applications*, 208, p.118063.

[33] Assery, N., Xiaohong, Y.D., Xiuli, Q., Kaushik, R. and Almalki, S., 2021. Evaluating disaster-related tweet credibility using content-based and user-based features. *Information Discovery and Delivery*, 50(1), pp.45-53.

[34] O'Brien, K., Simek, O. and Waugh, F., 2019, January. Collective classification for social media credibility estimation. In *Proceedings of the 52nd Hawaii International Conference on System Sciences*.

[35] Embar, V.R., Bhattacharya, I., Pandit, V. and Vaculin, R., 2015, August. Online topic-based social influence analysis for the wimbledon championships. In *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 1759-1768).

[36] bin Baharudin, B. and bin Md Said, A., 2018. Trust blog ranking using multi-criteria decision analysis AHP and TOPSIS. In *IT Convergence and Security 2017* (pp. 68-76). Springer, Singapore.

[37] Pasi, G., Viviani, M. and Carton, A., 2019. A Multi-Criteria Decision Making

credibility of scientific content on facebook. *Technology in Society*, 61, p.101261

[22] Gupta, S., Sachdeva, S., Dewan, P. and Kumaraguru, P., 2018, December. CbI: Improving Credibility of User-Generated Content on Facebook. In *International Conference on Big Data Analytics* (pp. 170-187). Springer, Cham.

[23] Devi, P.S. and Karthika, S., 2019. # CycloneGaja-rank based credibility analysis system in social media during the crisis. *Procedia Computer Science*, 165, pp.684-690.

[24] Setiawan, E.B., Widyantoro, D.H. and Surendro, K., 2020. Measuring information credibility in social media using combination of user profile and message content dimensions. *International Journal of Electrical & Computer Engineering* (2088-8708), 10(4).

[25] Son, J., Lee, J., Oh, O., Lee, H.K. and Woo, J., 2020. Using a Heuristic-Systematic Model to assess the Twitter user profile's impact on disaster tweet credibility. *International Journal of Information Management*, 54, p.102176.

[26] Evans, L., Owda, M., Crockett, K. and Vilas, A.F., 2021. Credibility assessment of financial stock tweets. *Expert Systems with Applications*, 168, p.114351.

[27] Momen Bhuiyan, M., Horning, M., Lee, S.W. and Mitra, T., 2021. NudgeCred: Supporting News Credibility Assessment on Social Media Through Nudges. arXiv e-prints, pp.arXiv-2108.

[28] Thaher, T., Saheb, M., Turabieh, H. and Chantar, H., 2021. Intelligent detection of false information in arabic tweets utilizing hybrid harris hawks based feature selection and machine learning models. *Symmetry*, 13(4), p.556.

[29] Cano, A.E., Mazumdar, S. and Ciravegna, F., 2014. Social influence

Theory Driven Framework, and Recommendation. *IEEE Access*, 9, pp.137744-137781.

[46] Jain, P.K., Pamula, R. and Ansari, S., 2021. A supervised machine learning approach for the credibility assessment of user-generated content. *Wireless Personal Communications*, 118(4), pp.2469-2485.

[47] Zaslavsky, A. and Paris, C., 2012. Context-aware twitter validator (CATVal): A system to validate credibility and authenticity of twitter content for use in decision support systems. In *Fusing Decision Support Systems into the Fabric of the Context* (pp. 323-334). IOS Press.

[48] Sikdar, S., Adali, S., Amin, M., Abdelzaher, T., Chan, K., Cho, J.H., Kang, B. and O'Donovan, J., 2014, July. Finding true and credible information on Twitter. In *17th International Conference on Information Fusion (FUSION)* (pp. 1-8). IEEE.

[49] Page, J.T. and Duffy, M.E., 2018. What does credibility look like? Tweets and walls in US presidential candidates' visual storytelling. *Journal of political Marketing*, 17(1), pp.3-31.

[50] Yin, C. and Zhang, X., 2020. Incorporating message format into user evaluation of microblog information credibility: A nonlinear perspective. *Information processing & management*, 57(6), p.102345.

[51] Kitsios, F., Mitsopoulou, E., Moustaka, E. and Kamariotou, M., 2022. User-Generated Content behavior and digital tourism services: A SEM-neural network model for information trust in social networking sites. *International Journal of Information Management Data Insights*, 2(1), p.100056.

approach based on the Choquet integral for assessing the credibility of User-Generated Content. *Information Sciences*, 503, pp.574-588.

[38] Alrubaian, M., Al- Qurishi, M., Al-Rakhami, M., Hassan, M.M. and Alamri, A., 2017. Reputation- based credibility analysis of Twitter social network users. *Concurrency and Computation: Practice and Experience*, 29(7), p.e3873.

[39] Verma, P.K., Agrawal, P., Madaan, V. and Gupta, C., 2022. UCred: fusion of machine learning and deep learning methods for user credibility on social media. *Social Network Analysis and Mining*, 12(1), pp.1-10.

[40] Khan, T. and Michalas, A., 2021. Seeing and Believing: Evaluating the Trustworthiness of Twitter Users. *IEEE Access*, 9, pp.110505-110516.

[41] Zarrinkalam F, Kahani M, Bagheri E. Mining user interests over active topics on social networks. *Information Processing & Management*. 2018 Mar 1;54(2):339-57.

[42] Viviani, M. and Pasi, G., 2017. Credibility in social media: opinions, news, and health information—a survey. *Wiley interdisciplinary reviews: Data mining and knowledge discovery*, 7(5), p.e1209.

[43] Zhao, K., Xing, L. and Wu, H., 2021. S3UCA: Soft-Margin Support Vector Machine-Based Social Network User Credibility Assessment Method. *Mobile Information Systems*, 2021.

[44] Przybyła, P. and Soto, A.J., 2021. When classification accuracy is not enough: Explaining news credibility assessment. *Information Processing & Management*, 58(5), p.102653.

[45] Qureshi, K.A., Malick, R.A.S. and Sabih, M., 2021. Social Media and Microblogs Credibility: Identification,

پیوست

در جداول الف و ب، مجموعه ویژگی‌های شبکه‌های اجتماعی جهت ارزیابی اعتبار اطلاعات معرفی شده‌اند. که این ویژگی‌ها براساس مرور ادبیات جمع‌آوری گردیده است.

جدول الف. مجموعه اول ویژگی‌های شبکه اجتماعی

گروه	ویژگی‌ها
اطلاعات کلی	داشتن بیوگرافی ([۱۰]، [۱۹]، [۲۸]، [۳۳])، احساسات بیوگرافی ([۲۴])، نام کاربری ([۱۱]، [۲۲]، [۲۴]، [۲۶])، [۲۸]، [۳۳])، داشتن ایمیل ([۲۲])، تصویر پروفایل ([۱۹]، [۲۶]، [۱۱]، [۲۸]، [۳۷])، تعداد گروه‌های عضو شده ([۲۲]، [۲۸]، [۳۲]، [۳۹]، [۴۰])، تعداد لینک‌ها پروفایل ([۲۲]، [۲۸])، نوع ثبت نام ([۲۸])، موقعیت مکانی منتشر شده در پروفایل ([۲۲])، تم پروفایل ([۲۲])، تعداد زبان‌های استفاده شده ([۲۲]، [۳۹])
فعالیت	سن حساب کاربری ([۱۰]، [۱۵]، [۱۷]، [۱۹]، [۲۲]، [۲۴]، [۲۶]، [۲۰]، [۳۳]، [۳۴])، تعداد فالوئینگ ([۱۷]، [۱۹]، [۲۴]، [۲۶]، [۳۳]، [۳۴]، [۳۹]، [۴۰])، نظم در ارسال محتوا ([۱۹]، [۲۶]، [۲۲]، [۱۴]، [۱۰]، [۲۸])، تعداد محتوا و بازنشرها ([۲۸]، [۳۳])، تاریخ انتشار محتوا ([۳۳])، تعداد محتوا ([۱۰]، [۳۶]، [۴۰])
تخصص	تعداد موضوعات مورد علاقه کاربر ([۱۸]، [۳۲]، [۳۵]، [۳۹])، تاییدیه حساب کاربری ([۱۹]، [۲۲]، [۲۶]، [۱۱])، [۲۴]، [۲۸]، [۳۳])، بررسی وب سایت کاربر ([۲۲]، [۲۴]، [۲۶]، [۲۷])، محتوای‌های منحصربه‌فرد ([۱۵]، [۳۴])، لینک‌های منحصربه‌فرد ([۳۴])، شباهت محتوا با یک منبع معتبر ([۳۱]، [۱۴])، نزدیکی مکان محتوا به موقعیت رویداد ([۱۶]، [۳۱])، نزدیکی زمان محتوا با رویداد ([۱۴]، [۳۱])، تعداد لینک‌های هر محتوا ([۱۷]، [۱۹]، [۲۶]، [۲۲]، [۳۴]، [۳۷]، [۱۶]، [۲۴]، [۲۵]، [۲۸]، [۳۳]، [۴۰])، تعداد موضوع هر محتوا ([۲۵])، تعداد دفعاتی که نظر کاربر مورد پسند قرار گرفته است ([۱۸]، [۲۴])، کلمات کلیدی ([۲۸])، نسبت تعداد محتوا منتشر شده توسط کاربر نسبت به سایر کاربران در یک حوزه خاص ([۱۰])، انتشار موقعیت در ارسال محتوا ([۱۶])

جدول ب. مجموعه دوم ویژگی‌های شبکه اجتماعی

گروه	ویژگی‌ها
محتوا	تعداد کلمات بیوگرافی ([۲۲]، [۲۴]، [۲۶])، تعداد جملات در بیوگرافی ([۲۲]، [۲۴])، طول بیوگرافی ([۲۸])، میانگین طول کلمات در بیوگرافی ([۲۲])، تعداد کلمات توقف در بیوگرافی ([۲۲])، تعداد کلمات منحصر به فرد در بیوگرافی ([۲۲])، تعداد کلمات اسپم ([۲۴])، تعداد محتوای اسپم ([۱۴])، طول محتوا ([۱۷]، [۱۹]، [۲۳]، [۲۲]، [۲۶]، [۱۴]، [۲۴]، [۲۸]، [۳۳]، [۳۷])، تعداد علائم نگارشی ([۱۹]، [۲۴]، [۱۴])، محتواهای دارای علامت سوال ([۱۹]، [۲۶]، [۲۸]، [۳۴])، محتواهای دارای علامت تعجب ([۱۹]، [۲۶]، [۲۸]، [۳۴])، محتواهای دارای ایموجی ([۱۹]، [۲۶]، [۲۲]، [۳۴]، [۲۴]، [۱۴])، محتواهای دارای ضمیر اول شخص ([۱۹]، [۲۶]، [۳۴])، محتواهای دارای ضمیر دوم شخص ([۲۶])، محتواهای دارای ضمیر سوم شخص ([۲۶])، تعداد کلمات با حروف بزرگ ([۱۹]، [۲۶]، [۲۲]، [۱۴])، تعداد کلمات توقف ([۲۲])، تعداد غلط‌های املائی ([۱۴])، تعداد کلمات کوتاه عامیانه

<p>([۲۶],[۲۲],[۲۴],[۱۴],[۱۹])، محتواها با افعال زمان گذشته ([۲۶])، محتواها با افعال زمان حال ([۲۶])، تعداد صفات در هر محتوا ([۲۶])، تعداد قید در هر محتوا ([۲۶])، تعداد کلمات مانند لطفاً، تشکر ([۱۹])، میزان خوانایی محتوا ([۱۹])، ارزیابی اعتبار لینک محتوا ([۱۹],[۲۶],[۳۳])، تعداد علامت سوال در نظرات ([۲۷])، احساسات مخاطبین در پاسخ به محتوا کاربر ([۱۸])،تعداد کلمات محتوا ([۱۹],[۲۶],[۲۲],[۲۳],[۳۴],[۲۴],[۲۵],[۲۸])، تعداد علامت هشتگ ([۲۸])،احساسات محتوا ([۱۵],[۱۷],[۳۸],[۱۹],[۲۲],[۲۰],[۲۴],[۲۸],[۴۰])، تعداد تصاویر ([۱۶],[۲۲])، تعداد ویدیوها ([۱۶],[۲۲])، میانگین طول کلمات محتوا ([۲۲])</p>	
<p>تعداد بازنشر محتوا سایر کاربران ([۱۷],[۲۶],[۲۸],[۲۹],[۳۰],[۳۱],[۳۲],[۳۳],[۳۴],[۳۷])، تعداد محتوایی که کاربر پسندیده است ([۳۰])، تعداد نظرات کاربر ([۱۵],[۲۲],[۳۱])</p>	پشتیبان
<p>تعداد فالور ([۱۰],[۱۷],[۱۸],[۳۶],[۳۸],[۱۹],[۲۹],[۳۵],[۲۶],[۲۲],[۳۰],[۳۴],[۲۰],[۲۴],[۳۱],[۳۷],[۲۸],[۳۲],[۳۳],[۳۹],[۴۰])، نسبت فالور به فالوئینگ ([۱۵],[۱۹],[۲۸])، نسبت فالوئینگ به فالور ([۲۸]) تعامل مخاطبین با کاربر در بازه‌های زمانی متفاوت ([۱۸],[۲۵],[۳۵])، تعداد پاسخ برای هر محتوا ([۱۴],[۱۶],[۱۸],[۲۲],[۳۵])،احساسات مخاطبین هنگام ذکر نام کاربر در محتوا ([۱۸],[۲۴])، تعداد دفعاتی که محتواهای کاربر توسط سایرین به اشتراک گذاشته شده است ([۱۵],[۱۷],[۱۸],[۳۸],[۱۷],[۱۸],[۳۵],[۱۶],[۲۴],[۴۰])، تعداد بازخوردهای مثبت محتوا ([۱۵],[۱۷],[۱۸],[۱۶],[۲۴],[۲۵],[۳۳],[۴۰])، تعداد بازدیدکنندگان محتوا ([۱۰],[۳۶])</p>	محبوبیت
<p>تعداد هشتگ‌های هر محتوا ([۱۶],[۱۷],[۱۹],[۲۴],[۲۵],[۲۶],[۲۸],[۳۳])،تعداد دفعاتی که از نام سایر کاربران در محتواها استفاده شده ([۱۹],[۳۸],[۲۹],[۳۵],[۲۶],[۳۴],[۲۵],[۲۴],[۲۸],[۳۷])، هشتگ‌های موجود در بیوگرافی ([۲۸])</p>	انتشار