The 9th International Conference on Web Research
ICWR2023
May 3th-4th, 2023, Tehran, Iran

Fake Accounts Detection on Social Media using Machine Learning: Review

Yasaman Mohammadi Dargah, Chitra Dadkhah*, Niloofar Rezaei

Computer Engineering Faculty, K. N. Toosi University of Technology, Tehran, Iran
yasaman@email.kntu.ac.ir, dadkhah@kntu.ac.ir, niloofarrezaei@email.kntu.ac.ir

Abstract

In today society, social networking websites have drowned a remarkable attention from users ranging from a child to an old aged person all around the world. The community consumes a huge amount of time on online social networks by interacting and exchanging their information with the other people in the globe. As a result, some of the popular websites like Facebook, Twitter, Instagram, and others witnessed an unexpected growth in registered users. Meanwhile, researches exhibits that all registered accounts are not real; there exist a huge number of fake accounts created for a specific purpose. The major purpose of creating fake accounts is to spread spam content, rumor, and other unauthentic messages on the platforms. This leads to a motivation of developing a system that is able to identify and filter fake accounts on the social networks, but it has many challenges. Researchers have proposed several advanced algorithms to recognize fake accounts. In this paper, the development of fake account detection algorithms using various machine learning approach and deep learning algorithms are reviewed, which give an open vision to the future researchers to develop a foundation in this field.

Keywords: Social Network, Spammer, Fake Account, Detection, Machine Learning, Deep Learning.

The 9th International Conference on Web Research
ICWR2023
May 3th-4th, 2023, Tehran, Iran

مروری بر شناسایی حساب‌های جعلی در رسانه‌های اجتماعی با استفاده از یادگیری ماشین

یاسمن محمدی درگاه[۱]، چیترا دادخواه*[۲]، نیلوفر رضایی[۳]

[۱] دانشکده مهندسی کامپیوتر، گروه هوش مصنوعی، دانشگاه صنعتی خواجه نصیرالدین طوسی ، تهران،
yasaman@email.kntu.ac.ir

[۲] استادیار، دانشکده مهندسی کامپیوتر،گروه هوش مصنوعی، دانشگاه صنعتی خواجه نصیرالدین طوسی ، تهران
dadkhah@kntu.ac.ir

[۳] دانشکده مهندسی کامپیوتر، گروه هوش مصنوعی، دانشگاه صنعتی خواجه نصیرالدین طوسی ، تهران،
niloofarrezaei@email.kntu.ac.ir

چکیده

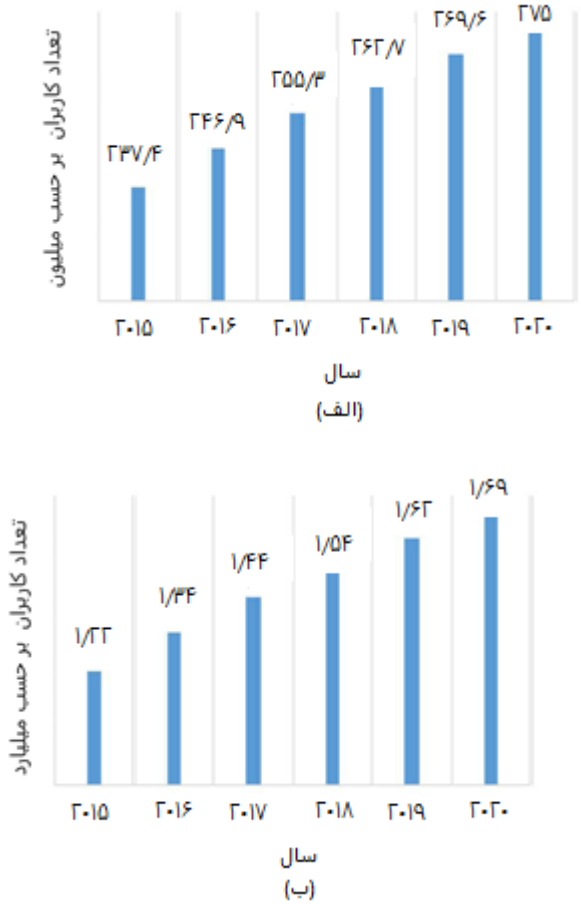
در جامعه‌ی امروزی، وب‌سایت‌های شبکه‌های اجتماعی توجه زیادی از کاربران، کودکان تا سالمندان، را در سراسر جهان به خود جلب کرده‌اند. افراد جامعه معمولاً زمان زیادی را در شبکه‌های اجتماعی برخط از طریق تعامل و تبادل اطلاعات خود با افراد دیگر در جهان صرف می‌کنند، در نتیجه، برخی از وب سایت‌های محبوب مانند فیس‌بوک، توییتر، اینستاگرام و غیره شاهد رشد غیرمنتظره کاربران ثبت‌نام‌شده هستند. در همین حال، تحقیقات نشان می‌دهد که تمام حساب‌های ثبت شده واقعی نیستند. ممکن است تعداد زیادی از حساب جعلی برای یک هدف خاص ایجاد شده باشد. هدف اصلی ایجاد حساب‌های جعلی برای انتشار محتوای هرزنامه، شایعه و سایر پیام‌های غیر معتبر در پلتفرم‌ها است. این موضوع منجر به انگیزه‌ی توسعه سیستم‌می‌شود که قادر به شناسایی و فیلتر کردن حساب‌های جعلی در شبکه‌های اجتماعی باشد اما با چالش‌های زیادی روبه‌روخواهد بود. محققان الگوریتم‌های پیشرفته متعددی برای تشخیص حساب‌های جعلی پیشنهاد کرده‌اند. هدف این مقاله ارائه مروری بر سیستم‌های اخیر جهت شناسایی حساب‌های جعلی با استفاده از الگوریتم‌های مختلف یادگیری ماشین و یادگیری عمیق می‌باشد که چشم‌اندازی باز به محققان آینده برای توسعه در این زمینه می‌دهد.

کلمات کلیدی

شبکه اجتماعی، هرزنامه، شناسایی، حساب جعلی، یادگیری ماشین، یادگیری عمیق.

۱- مقدمه

در سال‌های اخیر، شبکه‌های اجتماعی برخط[۱] کاربران زیادی را از سراسر جهان جذب کرده است. از آنجایی که فناوری به سرعت در حال رشد بوده و دسترسی به اینترنت آسان‌تر و سریع‌تر از قبل شده است، رسانه‌های معروف و متداولی مانند توییتر، فیس‌بوک و اینستاگرام کاربران ثبت‌نام‌شده‌ی متعددی از سراسر دنیا را دربرگرفته اند [۱–۵]. شما می‌توانید به صورت رایگان به همه‌ی این رسانه‌ها بپیوندید و احساسات، علایق و دانش خود را به اشتراک بگذارید. در شبکه‌های اجتماعی برخط افراد حتی می‌توانند بدون پرداخت هزینه با مخاطبین یا دنبال کنندگان خود ارتباط داشته باشند، پیام‌ها، عکس‌ها، فیلم‌ها و غیره را مبادله و به اشتراک بگذارند. آمار نشان می‌دهد که فیس‌بوک در سال ۲۰۱۷ دارای ۲/۱۳ میلیارد کاربر فعال ماهانه است. کاربران ثبت شده در هند حدود ۲۶۰ میلیون نفر هستند که بیشترین تعداد کاربران را تشکیل می‌دهد، آمریکا ۱۹۴ میلیون کاربر ثبت شده و توییتر ماهانه ۳۳۰ میلیون کاربر فعال دارد. کاربران توییتر و فیس‌بوک در سراسر جهان در شکل (۱) نشان داده شده است [۸].



شکل ۱.تعداد کاربران در شبکه‌های اجتماعی(الف) توییتر. (ب) فیس بوک[۸]

کمک به افراد برای ایجاد ارتباط جدید با دیگران و حفظ ارتباط با دوستان قدیمی و جدید یکی از برجسته‌ترین مزایای شبکه‌های اجتماعی برخط است. قابلیت‌های ذکر شده در کنار بسیاری از راه‌های ارتباطی دیگر با دوستان، شبکه‌های اجتماعی را بسیار جذاب می‌کند. در مقابل، این شبکه‌ها دارای اشکالات متعددی مانند امنیت، حریم خصوصی، ارسال هرزنامه، امکان شایعه پراکنی در آن‌ها و حساب‌های جعلی را دارد که با افزایش تعداد کاربران شبکه‌های اجتماعی، ایجاد حساب‌های جعلی نیز افزایش می‌یابد و از آنجایی که محدودیتی برای تعداد حساب‌های ایجاد شده به ازای هر کاربران وجود ندارد، فرصت بزرگی برای کاربران جعلی جهت ایجاد حساب‌های جعلی و سوء استفاده به نفع منافع شخصی خود می‌گردد. به همین دلیل در سال‌های گذشته تعداد جرائم سایبری افزایش یافته است.

طبق آمار، دغدغه‌ی وجود حساب‌های جعلی در شبکه‌های اجتماعی رایج بوده و به عنوان مثال، فیس‌بوک ۱/۳ میلیارد حساب کاربری جعلی دارد. ارسال پیام‌های هرزنامه، شایعه پراکنی، اخبار نادرست، سخنان تنفرآمیز و کسب درآمد غیرقانونی از فعالیت‌هایی است که توسط حساب‌های کاربری جعلی انجام می‌شود [۹].

با توجه به آمار، توییتر معروف‌ترین شبکه‌ی اجتماعی برای اسپم‌ها و انتشار پیام‌های نادرست، شایعات و پیام‌های هرزنامه است [۱۰]. هرزنامه‌ها لینک‌های مخرب خارجی خطرناکی را به پست‌های خود در توییتر اضافه می‌کنند. پیوندهای پیوست شده به پیام‌ها ممکن است حاوی ویروس‌ها یا سایر برنامه‌های مخرب باشد که اطلاعات کاربر را می‌دزدند [۱۱]. پیام‌های هرزنامه‌ای که از طریق ایمیل منتشر می‌شوند بین ۰/۰۰۰۳٪ و ۰/۰۰۰۶ ٪ است درحالی‌که پیام‌های هرزنامه که از طریق توییتر پخش می‌شوند ۰/۱۳٪ میباشد [۱۲].

حساب‌های جعلی در هر جایی مانند شبکه‌های اجتماعی، وب‌سایت‌های دوست‌یابی برخط، وبلاگ‌های گفتگو، وب‌سایت‌های خرید و غیره وجود دارند. انواع مختلفی از پروفایل‌های جعلی در شبکه‌های اجتماعی وجود دارد که برخی از آن‌ها عروسک‌های جورابی[۲] هستند [۱۳]. عروسک‌های جورابی حسابی است که با هدف متقاعد کردن حجم زیادی از افراد در مورد یک محصول خاص یا یک مقصد گردشگری خاص یا یک شرکت خاص و غیره ایجاد می‌شود. از این حساب‌ها برای فریب مردم استفاده می‌شود. ربات‌ها [۱۴] برنامه‌های کامپیوتری هستند که برخی از داده‌ها را از طریق تعامل با انسان تولید کرده و برای خودکارسازی و سرعت بخشیدن به کاراستفاده می‌شوند. این ربات‌ها می‌توانند به عنوان پاسخ دهنده‌ی به سؤالات مشتریان، راهنمای سفر ، کمک‌های پزشکی و غیره بطور خودمختار عمل کنند. این ربات‌ها می‌توانند توسط مهاجمان برای ارسال خودکار درخواست دوستی، ارسال پیام و غیره مورد سوء استفاده قرار گیرند. ربات‌ها انواع مختلفی دارند مانند ربات‌های هرزنامه، ربات‌های اجتماعی، ربات‌های چت و غیره. برخی از حساب‌های کاربری در معرض جعل شدن قرار دارندو حساب‌های کاربری شبیه‌سازی شده، از نوع پروفایل‌های حساب کاربری واقعی و موجود در فضای مجازی هستند. این نوع پروفایل‌ها برای اهداف غیرقانونی متعددی مانند سرقت اطلاعات خصوصی افراد استفاده می‌شود. حساب‌های کاربری جعلی ایجاد شده از روی حساب‌های کاربری واقعی باعث می‌شود که مالک، کنترل کامل یا جزئی بر روی حساب کاربری خود را از دست دهد. این نوع پروفایل‌ها اساساً توسط هکرها برای انتشار محتوای مخرب از طریق یک پروفایل قابل اعتماد استفاده می‌شود. شکل (۲) برخی از چالش‌های موجود در شبکه‌های اجتماعی را نشان می‌دهد که می‌تواند بر کاربران شبکه‌های اجتماعی برخط تأثیر بگذارد [۱۵]. تحقیقات زیادی برای شناسایی پیام‌های اسپم انجام شده است. [۱۵–۲۳]. با این حال، شناسایی پروفایل‌های جعلی در شبکه‌های اجتماعی هنوز یک مشکل بزرگی است که برای حل آن، مطالعات زیادی توسط محققان انجام شده است. در این مقاله‌ی، تحقیقات اخیر در زمینه‌ی شناسایی پروفایل‌های جعلی در شبکه‌های اجتماعی که توسط محققان با استفاده از الگوریتم‌های مختلف یادگیری ماشینی توسعه داده شده است، مورد تحلیل قرار گرفته شده و انواع مختلف ویژگی‌های مربوط به حساب‌های کاربران شبکه‌های اجتماعی مورد بررسی در این رابطه نیز ارائه شده است. همانطور که محققان در حال توسعه تکنیک‌های بیشتر و بیشتری هستند، هکرها نیز به بهره برداری از این تکنیک‌ها ادامه می‌دهند. این امر منجر به نیاز به توسعه‌ی مداوم تکنیک‌های جدید برای دفاع سیستم‌ها و شبکه‌های اجتماعی در برابر مهاجمان می‌شود.

شکل۲. چالش‌های موجود در شبکه‌های اجتماعی برخط[۱۵]

۲- روش‌های شناسایی حساب‌های جعلی

سـیرینواس رووا[۳] و همکاران روشـی مبتنی بر یادگیری ماشـین و پردازش زبان طبیعی برای افزایش دقت در تشخیص حساب‌های جعلی و واقعی پیشنهاد کردند. در این تحقیق از مجموعه داده‌ی فیس‌بوک برای شـناسـایی حسـاب‌های جعلی استفاده شـده و با اسـتفاده از روش ماشـین بردار تصـمیم (SVM [۴]) و الگوریتم ساده‌ی بیزین[۵]، بسیاری از مشکلات در رسانه‌های اجتماعی از جمله مسائل مربوط به حریم خصوصی، آزار و اذیت سایبری، خرابکاری و غیره را از بین بردند [۲۴].

معماری پیشـنهادی توسـط روهیت راتوری[۶] دارای دو رویکرد اسـت که در رویکرد اول مبتنی بر روشـهای پردازش زبان طبیعی و شـناسـه‌های شـبکه برای شـناسـایی جزئیات حسـاب‌های کاربری اسـت. بر اسـاس شـناسـه‌های شـبکه، درصـورتی‌که بیش از یک حسـاب کاربری مورد اتهام قرار گیرد، از کاربران درخواست امنیت می‌شود. در رویکرد دوم، از SVM BOW [۷]: کیسه کلمات در کنار روش ماشین بردار تصمیم برای تعیین تعداد کلمات مضر استفاده می‌کنند. تعداد کلمات مضـر به عنوان مجموعه داده جمع‌آوری و برای آزمایش و آموزش تقسیم می‌شوند و سپس کیسه کلمات از طریق SVM به مدل پیش بینی متصل می‌شود. تعداد کلمات مخرب برای هر حساب را محاسبه کرده و حساب‌های جعلی با توجه به محتوای حساب، شناسایی می‌شوند، مرحله بعدی ارسال یک پیام هشدار دهنده برای ارائه‌ی یک مدرک معتبر برای اجازه‌ی ادامه کار با حسـاب کاربری آن‌ها اسـت. از طریق این فرآیند، تعداد یک کلمه و تعداد تکرار کلمات مضـر مشخص می‌شود بنابراین اگر این مقیاس بزرگ‌تر یا مساوی ۳ باشد، کاربر باید مدارکی جهت ادامه با حساب کاربری خود ارائه دهد و اگر این عدد کوچک‌تر از ۲ باشد کاربر بدون ایراد است. آن‌ها در این تحقیق از مجموعه داده‌های فیسبوک و توییتر و ویژگی‌هایی مانند توییت‌ها، برچسـب‌ها، لایک‌ها و پسـت‌های توییتر و برای شناسایی اهداف از تحلیل معنایی استفاده نموده‌اند[۲۵].

مدلی توسط کالکارنی[۸] و پروفسور ویدیا[۹] بر اساس ساختار برنامه‌ریزی شده پیشـنهاد شـده که از طریق آن واقعی یا غیر واقعی بودن حسـاب برخط کشـف می‌شـود. از چندین رویکرد طبقه بندی مانند SVM و بیز سـاده توسـط این مدل برای شـناسـایی حسـاب‌ها اسـتفاده می‌شـود. ابتدا حسـاب مورد نظر برای آزمایش انتخاب شده، سپس ویژگی‌های مناسب انتخاب میگردند، در سطح انتخاب ویژگی، اقدامات احتیاطی مناسبی انجام شده تا ویژگی‌ها مستقل از سایر ویژگی‌ها باشند، همچنین ویژگی‌هایی که کارایی طبقه‌بندی را افزایش می‌دهند، انتخاب شوند. پس از این مرحله، مجموعه داده‌های حساب‌های واقعی یا جعلی شناسایی شده‌ی قبلی برای آموزش الگوریتم طبقه‌بندی مورد اسـتفاده قرار می‌گیرند. طبقه‌بند، مشـخص می‌کند که حساب کاربری مورد نظر ما اصلی است یا جعلی؛ و اگر با یک حساب جعلی مواجه شـویم، یک اعلان برای آن حسـاب ارسـال می‌شـود تا هویت واقعی مشخص شود و سپس در شرایط واقعی بودن حساب هدف، یک پیام فیسبوک به مدل طبقه‌بند ارسـال می‌شـود که بیان می‌کند این حسـاب جعلی نیسـت. آنها با اسـتفاده از روش‌های پردازش زبان طبیعی حسـاب‌های غیر واقعی را با کارایی بالا شـناسـایی نمودند. این امر منجر به نرخ دقت تقریباً ۹۵٪ برای هر حساب کاربری اجتماعی برخط گشـته که رویکرد ذکر شـده با تلاش کمتر حسـاب‌های جعلی را شناسایی می‌کند[۲۶].

یک مدل ترکیبی و الگوریتم تشـخیص پوسـت توسـط اسـمروتی[۱۰]،هارینی[۱۱] برای شناسایی حساب‌های غیرواقعی در شبکه‌های اجتماعی ارایه شده است[۲۷]. در این مقـالـه از پنج نوع الگوریتم یـادگیری مـاشـیـنی: کـا-نزدیـک ترین همسـایه(KNN) [۱۲]، ماشـین بردار تصـمیم، بیز سـاده، درخت تصـمیم و جنگل تصادفی[۱۳] استفاده شـده است. مجموعه داده‌ی این تحقیق شامل ترکیبی از ۲۰۰ حساب جعلی شناسایی شده به صورت دستی به اضافه ۲۰۰ حساب واقعی است که در مجموع ۴۰۰ حسـاب کاربری را تشـکیل می‌دهد. تمامی‌الگوریتم‌های یادگیری ماشینی با نظارت[۱۴] ذکر شده بر روی مجموعه داده‌ی مذکور اعمال شده است. از طرف دیگر یک مرحله پردازش تصـویر برای یافتن پیکسـل‌ها و نواحی همرنگ پوست در تصاویر یا ویدیوها بکارگرفته شده است. این یک مرحله پیش‌پردازش به منظور تشـخیص نواحی حاوی صـورت انسـان و اندام، در تصـاویر اسـت. الگوریتم تشـخیص پوسـت برای تشـخیص تصـاویر مبتذل در پروفایل کاربران اسـتفاده می‌شـود. یک روش یادگیری عمیق برای تشـخیص این که آیا انسـان در تصـاویر وجود دارد یا خیر نیزبکارگرفته شـده اسـت. تصـاویر حاوی انسـان تحت تشـخیص پوست قرار گرفته و درصد وجود پوست در تصویر محاسبه می‌شود. با توجه به این درصـد پروفایل‌های تقلبی و اصـلی به این صـورت مشـخص می‌شـوند که پروفایل‌های تقلبی متشـکل از درصـد زیادی پوسـت و یا فاقد تصـویر هسـتند. اگر نسبت درصد پوست بیش از ۱۳٪ باشد، می‌توان آن را به عنوان یک پروفایل جعلی در نظر گرفت. با اسـتفاده از الگوریتم‌های یادگیری ماشـین ذکر شـده، اعتبارسـنجی متقابل[۱۵] شامل ۱۰ فولد[۱۶] برروی مجموعه داده‌ها جهت ارزیابی سیستم پیشـنهادی اعمال شـده اسـت. تعداد پسـت‌ها و نظرات، حضـور در سـایر رسـانه‌های اجتماعی، اعلام حضور، اطلاعات شخصی، دوستان، دنبال‌کنندگان و رویدادها از ویژگی‌های شناسایی دستی حساب کاربران هستند که جهت شناسایی حساب های جعلی با دقت بالا انجام شده است. در میان تمام الگوریتم‌های یادگیری ماشین اعمال شده، طبقه‌بند درخت تصمیم و بیز ساده بالاترین دقت یعنی ۸۰ درصد را داشته و سایر

الگوریتم‌های دارای دقت بین ۶۰ تا ۸۰ درصد که میزان دقت الگوریتم KNN، ۶۰ درصد است.

نامبوری سراویا و همکاران[۱۷]، سیستمی برای تعیین فعالیت‌های مخرب در رسانه‌های اجتماعی پیشنهاد داده‌اند[۲۸]. آنها بر روی Sybil و هویت‌های هرزنامه با استفاده از الگوریتم‌های یادگیری ماشین برای غلبه بر هویت‌های جعلی تمرکز کردند. در این روش، مجموعه داده‌ها از یک وبلاگ داده‌ی بزرگ جمع‌آوری و سپس ذخیره می‌شوند که اگر داده‌ها مناسب نباشند، پاک و ذخیره نمی‌شوند. با این اقدام، داده‌ها دقیق‌تر می‌شوند و باعث استفاده از یک تکنیک پیشرفته برای حساب‌های Sybil و Troll می‌شوند. پس از فرآیند پاکسازی داده‌ها، فضاهای داده‌های از دست رفته پر می‌شود. این بدان معنی است که مناطق گمشده افراد جعلی و مناطق پرکننده افراد جعلی پاک شده هستند. داده‌ها قبل از مرحله تمیزی برای مراجعات بعدی در یک پایگاه داده غیر مرتبط ذخیره و به حذف حساب‌های کاربری جعلی کمک می‌کنند. پس از مرحله آماده‌سازی داده‌ها، داده‌ها برای آموزش و آزمایش تقسیم شدند که ۸۰ درصد داده‌ها برای آموزش و ۲۰ درصد داده‌ها برای آزمایش در نظر گرفته شده است. الگوریتم‌های یادگیری ماشین اعمال‌شده بر روی مجموعه‌ی داده‌ها، رگرسیون خطی و الگوریتم‌های نظارتی بیز ساده هستند. سیستم پیشنهادی آنها به ویژگی‌هایی مانند نام، مکان و تصویر حساب کاربری وابسته است. دقت تشخیص حساب کاربری جعلی در سیستم پیشنهادی آنها بیش از ۹۰ درصد است.

ارشاهین[۱۸] و همکاران روشی را برای یافتن حساب‌های جعلی در توییتر پیشنهاد کرده‌اند[۲۹]. مجموعه داده‌ی تهیه شده برای این تحقیق به صورت دستی جمع‌آوری شده است. طبقه‌بند بر اساس ویژگی هایی مانند نام کاربری، حساب کاربری و تصویر پس زمینه، تعداد دنبال کننده‌ها، توضیحات مندرج در حساب کاربری، تعداد توییت‌ها و محتوای توییت‌ها طبقه‌بندی را انجام می‌دهد. پایگاه داده شامل ۵۰۱ حساب جعلی و ۴۹۹ حساب واقعی است، تعداد داده‌ها در هر طبقه/دسته به خاطر کیفیت نتیجه باید متعادل باشد. ۱۶ ویژگی از API اطلاعات توییتر جمع‌آوری و از دورویکرد به منظور طبقه‌بندی حساب‌های جعلی استفاده شده است. درروبکرد اول از الگوریتم بیزین ساده بدون گسسته‌سازی در مجموعه داده‌ی توییتر و در دومی‌از روش بیز ساده پس از گسسته سازی استفاده شده است. میزان موفقیت تشخیص پروفایل جعلی درروبکرد اول ۸۶/۱ ٪و در رویکرد دوم، ۹۰٪/۹ بود.

یک روش مبتنی بر خوشه‌بندی مارکوف (MCL) توسط احمد[۱۹] و ابولایش[۲۰] پیشنهاد شده است که برای تعیین پروفایل‌های جعلی در شبکه‌های اجتماعی برخط است[۳۰]. مجموعه داده‌ی مورد استفاده فیس‌بوک شامل ۳۲۰ پروفایل کاربران، حاوی ۱۶۵ حساب جعلی و ۱۵۵ حساب واقعی است. سه ویژگی اصلی از حساب‌های کاربری استخراج و جهت شناسایی استفاده شده است. ۱) دوست فعال: نشان می‌دهد که کاربر چقدر با دوستان خود ارتباط برقرار می‌کند. ۲) لایک‌ها: هر چند وقت یک‌بار کاربران مطالب به اشتراک گذاشته شده را لایک می‌کنند۳) URL: تعداد دفعاتی که کاربران، URL را به اشتراک می‌گذارند. تجزیه و تحلیل‌ها نشان می‌دهد که افراد جعلی URL را بیشتر از کاربران معتبر به اشتراک می‌گذارند. آن‌ها همچنین با استفاده از کاربران و ارتباطات آن‌ها یک نمودار اجتماعی ایجاد می‌کنند تا از رفتارهای مشابه پروفایل‌های واقعی سوءاستفاده کنند.

زارعی[۲۱] و همکاران مدلی را پیشنهاد کردند که پروفایل‌های جعل هویت سیاستمداران را در شبکه‌های اجتماعی شناسایی می‌کند[۳۱]. مجموعه داده شامل سه سیاستمدار اینستاگرام، از جمله دونالد ترامپ، باراک اوباما و امانوئل مکرون است. مجموعه داده در یک دوره سه ماهه شامل ۸۰ پست، ۳۵۰۰۰۰ نظر، ۹ میلیون لایک و ۱/۵ میلیون پروفایل ساخته شده است. با استفاده از این مجموعه داده، مدل آن‌ها می‌تواند طیف گسترده‌ای از تقلید کنندگان و ربات‌ها را شناسایی کند. نویسندگان ادعا می‌کنند که این اولین مقاله‌ای است که چنین تحلیل داده‌ای را در اینستاگرام انجام می‌دهد. این تحقیق در [۳۲] تعمیم داده شد و در آن، مجموعه داده از سه نوع حساب کاربری: سیاستمدار، خبرگزاری و ستاره ورزشی جمع‌آوری شده است. در مجموع ۵۵۰ پست، ۱/۳ میلیون بازدید، ۲۰ میلیون لایک و ۶ میلیون حساب کاربری در قالب فرمت JSON ذخیره شده است. از تکنیک Tf-Idf[۲۲]برای شناسایی پروفایل‌ها با اطلاعات یکسان استفاده شده است برای تطبیق تصاویر پروفایل از مدل شبکه عصبی پیچشی(CNN) [۲۳] استفاده شده است. ویژگی‌های دیگر مانند تعداد نظرات و لایک‌ها نیز در نظر گرفته شده است. در نهایت از تکنیک‌های خوشه‌بندی برای خوشه‌بندی پروفایل‌ها به منظور مشاهده رفتارهای عجیب و پروفایل‌های غیرعادی استفاده شده است.

رواثی[۲۴] و همکاران از رویکردهای ارتباطی شباهت پروفایل برای شناسایی حساب‌های تکراری در شبکه‌های اجتماعی استفاده کرده‌اند[۳۳]. رکوردهای موجود در مجموعه داده را از سایت‌های مختلف در یک گروه جمع‌آوری و حساب‌های کاربری تکراری را در مجموعه داده‌ها جست‌وجو کرده‌اند. بالاترین دقت ۹۳/۸۷٪ با استفاده از تطبیق ارتباط شباهت گره NSCM [۲۵] به دست آمده است.

تکنیک‌های داده‌کاوی ارایه شده توسط سوریاکالا[۲۶]و رواثی[۲۷] برای شناسایی حساب‌های کاربری تکراری در شبکه‌های اجتماعی استفاده شده است[۳۴]. آن‌ها مجموعه داده این آزمایش را از GitHub جمع‌آوری و از روش‌هایNSCM، SVM، جنگل تصادفی، تکنیک تولید قاعده منطق[۲۸] توصیف شبکه و تکنیک‌های محافظت شده از حریم خصوصی برای شناسایی حساب‌های جعلی استفاده کرده‌اند در بین همه‌ی این روش‌ها، مدل سیستم محافظت شده با حریم خصوصی به بهترین مقدار دقت ۹۷/۳٪ رسیده است.

پورقمی[۲۹] و همکاران مدلی برای تشخیص حساب‌های کاربری غیرواقعی در فیسبوک معرفی کردند[۳۵]. تحقیق آنها شامل ۳ مورد است. ۱) اثربخشی الگوریتم برای شناسایی حساب‌های جعلی فیس‌بوک، ۲) عملکرد یادگیری هوش مصنوعی در فیسبوک برای تمایز بین پروفایل‌های جعلی و واقعی و ۳) پیامدهای اخلاقی تغییرات خط مشی فیسبوک. سیاست‌های امنیتی و حریم خصوصی متعددی توسط این مطالعه پیشنهاد شده است و این تحقیق روشی مؤثر برای یافتن پروفایل‌های جعلی احتمالی ارائه کرده است.

یک مدل یادگیری عمیق به نام Deep Profile توسط , واندا[۳۰] و جی[۳۱] برای تشخیص پروفایل‌های جعلی در شبکه‌های اجتماعی پیشنهاد شده است[۳۶]. مورد جدیدی که آن‌ها در تحقیق خود معرفی کردند، تغییر لایه ادغام CNN بود. مقادیر معیارهای ارزیابی دقت، یادآوری[۳۲] و امتیاز F1 برای مدل پیشنهادی آن‌ها به ترتیب ۹۴/۰۰، ۹۳/۲۱ و ۹۳/۴۲ بود. مقادیر ROC از ۰/۹۵۰ تا ۰/۹۵۹ متغیر بوده و مقدار خطا ۰/۲۱۴ بوده‌است.

آدوول[۳۳] و همکاران به منظور شناسایی پیام‌های هرزنامه و حساب‌های جعلی در شبکه‌های اجتماعی مجموعه داده‌ای از سه منبع جمع‌آوری: مجموعه‌ی

کلکسون پیامک V1، مجموعه نوشته‌های بزرگ پیام کوتاه V.0.1، توییتر و مجموعه نوشته‌های اسپم به ترتیب با مجموع ۵۵۷۴، ۱۳۲۴ و ۱۸۰۰۰ نمونه. از ۲۰ ۹۹۸ حساب توییتر و ۳ ۷۵۵ ۳۶۷ توییت برای شناسایی حساب‌های اسپم جمع‌آوری نمودند. آن‌ها هجده ویژگی، از جمله ویژگی‌های مبتنی بر محتوا/رفتار، تعداد توییت‌ها، تعداد هشتگ‌ها، تعداد کلمات جعلی میانگین زمان بین هر توییت و غیره را در شناسایی اسپمرها در توییتر استخراج کرده‌اند. با استفاده از الگوریتم‌های یادگیری ماشین مقدار دقت ۰/۹۳۳ و مقدار AUC برابر ۰/۹۷۷ به دست آمد. در میان کلیه طبقه‌بندها روش جنگل تصادفی بهترین نتیجه را حاصل نمود[۳۷].

سو[۳۴] و میو[۳۵] برای تشخیص پروفایل جعلی شبکه اجتماعی برخط از یک لیست سیاه به جای لیست کلمات هرزنامه سنتی استفاده نمودند[۳۸]. لیست سیاه با استفاده از مدل سازی موضوع و استخراج کلمات کلیدی ایجاد می‌شود. ارزیابی بر روی مجموعه داده 1KS-10KN وهانی پات اجتماعی انجام شده است. مقدار دقت روش سنتی مبتنی بر فهرست هرزنامه–کلمه در مجموعه داده‌ی 1KS-10KN برابر با ۰/۸۵۴،مقدار معیار یادآوری برابر با ۰/۹۰۴ و مقدار F-measure برابر ۰/۸۷۹ است، در حالی که برای روش‌های پیشنهادی دقت، یادآوری و مقدار F1 به ترتیب ۰/۹۵۸، ۰/۹۵۰و ۰/۹۵۴ است. با استفاده از روش پیشنهادی، درصد یافتن پروفایل جعلی برابر با ۹۵/۴ درصد گشته. نرخ مثبت کاذب رویکرد مبتنی بر فهرست هرزنامه–کلمه ۰/۱۵۴ و نرخ مثبت کاذب رویکرد مبتنی بر لیست سیاه با استفاده از مجموعه داده‌هانی پات اجتماعی ۹۴/۹٪ است. نرخ تشخیص رویکرد مبتنی بر فهرست کلمات هرزنامه با استفاده از رویکرد مبتنی برهانی پات اجتماعی ۹۱/۱٪ گشته. رویکرد مبتنی بر لیست سیاه می‌تواند دقت قابل قبولی را به دست آورد و نرخ مثبت کاذب را کاهش دهد. ویژگی‌های مبتنی بر پروفایل و شبکه برای مدل آن‌ها مورد نیاز نیست و این مدل زمان و هزینه سربار برای حذف این ویژگ‌ها را کاهش می‌دهد.

اعوان[۳۶] و همکاران مدلی را برای پیش بینی حساب‌های کاربران جعلی در رسانه‌های اجتماعی که شامل کتابخانه‌های Spark ML برای طبقه‌بند جنگل تصادفی و ابزار رسم معرفی نمودند[۳۹]. در مرحله اول، آنها داده‌های در دسترس عموم ۴۰۰۰ پروفایل فیسبوک را جمع‌آوری و سپس از کتابخانه Spark ML برای ایجاد چارچوب داده برای پردازش بیشتر استفاده کردند. طبقه بند جنگل تصادفی بر اساس تعداد دنبال کننده‌ها، تعداد علاقه مندی‌ها، تعداد دوستان، تعداد لیست شده و جنسیت، طبقه بندی داده‌ها را با دقت ۹۴٪ انجام داده است.

شن[۳۷] و همکاران روشی را برای تشخیص هرزنامه‌های اجتماعی با استفاده از فاکتورسازی ماتریس غیرمنفی محدب (CNMF [۳۸]) پیشنهاد نمودند[۴۰]. آنها استدلال می‌کنند که رویکرد CNMF می‌تواند به طور موثری رابطه‌ی بین کاربران و ویژگی‌ها را نشان داده و می‌تواند رفتار ارسال‌کنندگان هرزنامه را به تصویر بکشد. روش پیشنهادی ابتدا یک ماتریس کاربر–ویژگی ایجاد می‌کند که فعالیت کاربران را بر روی ویژگی‌های مختلف، مانند تعداد پست‌ها، نظرات و لایک‌ها نشان می‌دهد. سپس، الگوریتم CNMF برای فاکتورسازی ماتریس ویژگی کاربر به دو ماتریس غیرمنفی اعمال می‌شود که رفتار پنهان کاربران و وزن‌های ویژگی را نشان می‌دهد. آنها استدلال می‌کنند که ساختار پراکنده ماتریس کاربر، تشخیص هرزنامه‌ها را به عنوان ناهنجاری در ماتریس ممکن می‌سازد. ارزیابی روش پیشنهادی آنها بر روی چندین مجموعه داده در دنیای واقعی نشان می‌دهند که از دیگر روش‌های پیشرفته برای تشخیص هرزنامه‌های اجتماعی بهتر عمل می‌کند. آن‌ها همچنین آزمایش‌هایی را برای نشان دادن استحکام روش پیشنهادی در برابر انواع مختلف رفتار هرزنامه انجام داده‌اند.

تاندان[۳۹] و همکاران روش پیشنهادی با رویکرد شبکه عصبی کانولوشن گرافیکی (GCNN) برای تشخیص فعالیت هرزنامه‌ها در پلتفرم‌های رسانه‌های اجتماعی را ارائه داده‌اند[۴۱]. روش پیشنهادی از ساختار نموداری، گره‌ها کاربران و لبه‌ها تعاملات بین کاربران، برای نمایش کاربران رسانه‌های اجتماعی و تعاملات آن‌ها استفاده نمودند. توسط مدل GCNN ویژگی‌هایی کاربران را استخراج شده و سپس مدل از این ویژگی‌ها برای طبقه‌بندی کاربران به عنوان هرزنامه یا غیرهرزنامه استفاده می‌کند. آنها روش پیشنهادی را بر روی مجموعه داده‌ی سایت‌های رسانه‌ی اجتماعی در دنیای واقعی ارزیابی می‌کنند. آن‌ها همچنین آزمایش‌هایی را برای نشان دادن اثربخشی اجزای مختلف مدل خود و ارائه بینشی در مورد رفتار ارسال کنندگان هرزنامه در رسانه‌های اجتماعی انجام داده‌اند. بیشترین دقت به دست آمده در این تحقیق ۹۱/۶۷٪ است.

سرینیواس[۴۰] و همکاران روشی را برای شناسایی هرزنامه‌ها در شبکه‌های اجتماعی با ترکیب تکنیک‌های یادگیری ماشین و پردازش زبان طبیعی که دقت تشخیص هرزنامه را با ثبت الگوهای زبانی و رفتاری هرزنامه‌ها بهبود بخشیده شده است، ارائه نمودند[۴۲]. آنها ابتدا داده‌های شبکه اجتماعی را با حذف ویژگی‌های نامربوط و سپس نرمالایز کردن، پیش پردازش نموده، سپس، ویژگی‌های متنی مختلف، مانند ویژگی‌های مبتنی بر فرکانس، ویژگی‌های مبتنی بر احساسات و ویژگی‌های نحوی را از داده‌های پیش پردازش شده استخراج می‌کنند. این ویژگی‌ها برای آموزش یک طبقه‌بند یادگیری ماشین، مانند ماشین‌های بردار پشتیبان یا جنگل‌های تصادفی، برای شناسایی هرزنامه استفاده نمودند. آنها روش پیشنهادی را بر روی چندین مجموعه داده از جمله فیسبوک و توییتر در دنیای واقعی ارزیابی کرده و آن را با سایر روش‌های پیشرفته برای تشخیص هرزنامه مقایسه کردند. آن‌ها نشان دادند که روش ارایه شده از نظر دقت و کارایی بهتر از روش‌های دیگر است. دقت این تحقیق با روش یادگیری بیز ساده ۹۳/۵۹ ٪ است. آنها همچنین آزمایش‌هایی را برای نشان دادن استحکام روش خود در برابر انواع مختلف رفتار هرزنامه انجام می‌دهند..

رامش[۴۱] و همکاران یک رویکرد جدید برای شناسایی هرزنامه‌ها و شناسایی کاربران جعلی در شبکه‌های اجتماعی ارائه داده‌اند [۴۳]. رویکرد پیشنهادی شامل سه مرحله اصلی است. ابتدا، نویسندگان ویژگی‌های مختلفی از جمله ویژگی‌های نمایه کاربر، ویژگی‌های شبکه و ویژگی‌های مبتنی بر محتوا را از داده‌های شبکه اجتماعی استخراج و سپس از تکنیک‌های آماری و یادگیری ماشین برای طبقه‌بندی کاربران به یکی از سه دسته کاربران واقعی، ارسال کنندگان هرزنامه یا کاربران جعلی استفاده می‌کنند. در نهایت، آنها تجزیه و تحلیل دقیقی از هرزنامه‌ها و کاربران جعلی شناسایی شده انجام می‌دهند تا بینشی در مورد رفتار و ویژگی‌های آن‌ها به دست آورند. آنها سیستم پیشنهادی را بر روی چندین مجموعه داده از جمله توییتر در دنیای واقعی ارزیابی کرده و آن را با سایر روش‌های پیشرفته مقایسه نمودندکه رویکرد آن‌ها از نظر دقت و کارایی بهتر از روش‌های دیگر عمل کرده است. آن‌ها همچنین آزمایش‌هایی را برای نشان دادن استحکام رویکرد خود در برابر انواع مختلف هرزنامه‌ها و رفتار کاربر جعلی انجام می‌دهند.

حما[۴۲] یک مدل شامل تکنیک‌های مختلف یادگیری ماشین، از جمله درخت‌های تصمیم، جنگل‌های تصادفی و ماشین‌های بردار پشتیبانی برای تجزیه

و تحلیل ویژگی‌های مختلف پروفایل‌های کاربر و شناسایی پروفایل‌های جعلی در شبکه‌های اجتماعی پیشنهاد نمودند[۴۴]. آنها از مجموعه داده‌ای از پروفایل‌های کاربر جمع‌آوری‌شده از یک سایت شبکه اجتماعی محبوب استفاده کرده و عملکرد مدل را با استفاده از معیارهای مختلف دقت، یادآوری و امتیاز F1 ارزیابی نمودند که نتایج نشان داده که مدل پیشنهادی در شناسایی پروفایل‌های جعلی، با دقت بیش از ۹۰ درصد، به خوبی عمل کرده است. آنها همچنین تجزیه و تحلیل دقیقی از ویژگی‌ها انجام داده که به شناسایی پروفایل‌های جعلی کمک و بینش‌های ارزشمندی را در مورد ویژگی‌های چنین پروفایل‌هایی ارائه می‌دهد.

۳- بحث

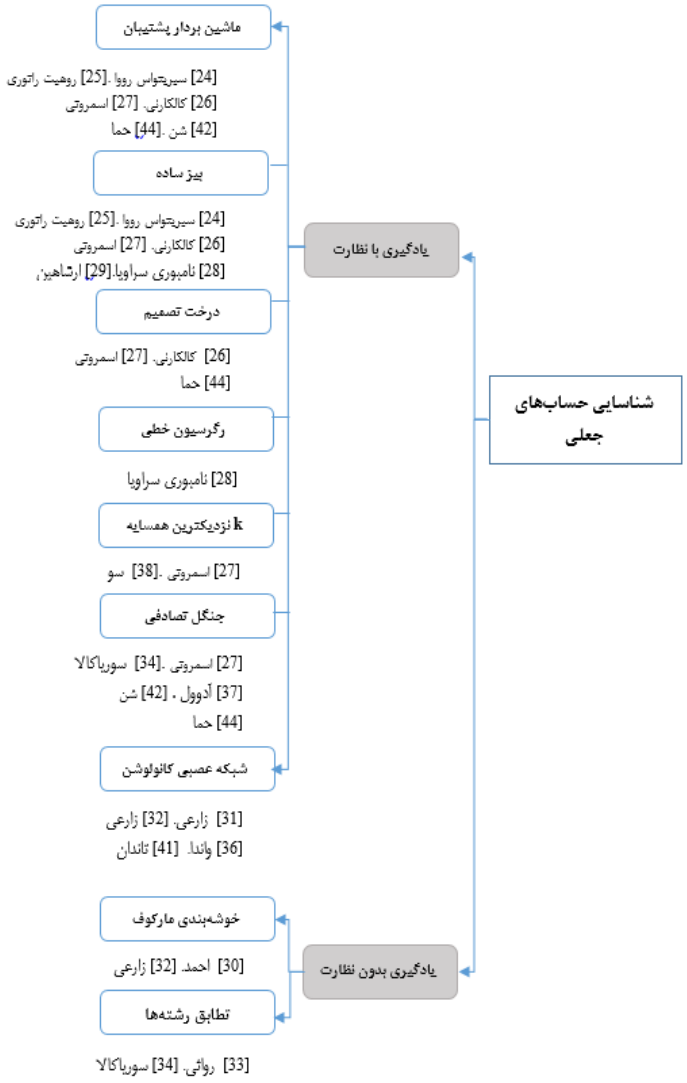
در این مقاله، الگوریتم‌های یادگیری ماشین در دو دسته با نظارت و بدون نظارت بکارگرفته شده توسط محققین جهت شناسایی حساب‌های جعلی در شکل (۳) دسته‌بندی شده و همچنین جدول(۱) پلتفرم‌های رسانه‌های اجتماعی به عنوان مجموعه داده استفاده شده در کارهای محققین ارائه داده شده است. بررسی‌های انجام شده، نشان می‌دهد که در بیشتر کارهای محققین از الگوریتم‌های یادگیری ماشین بانظارت مخصوصا جنگل تصادفی و ماشین بردار پشتیبان به دلیل دقت بالا در دسته‌بندی حساب‌های کاربری استفاده شده است. همچنین اکثریت سیستم‌های شناسایی حساب‌های جعلی از شبکه اجتماعی فیس بوک جهت ارزیابی استفاده نمودند.

۴- نتیجه‌گیری

در این مقاله مروری بر تحقیقات انجام‌شده بر روی شناسایی حساب‌های جعلی در شبکه‌های اجتماعی برخط با استفاده از الگوریتم‌های مختلف یادگیری ماشین صورت گرفته است. در چند سال گذشته، شبکه‌های اجتماعی در سراسر جهان مورد توجه عموم مردم قرار گرفته است، شبکه‌های اجتماعی محبوبی مانند توییتر، فیسبوک، اینستاگرام و غیره به دلیل ثبت نام زیاد کاربران و عدم محدودیت کاربری، براحتی مورد حمله مخربین قرار گرفته و تعداد زیادی از حساب کاربری ایجاد شده در این شبکه‌ها واقعی نبوده و برای اهداف مختلفی مانند انتشار پیام‌های غیر معتبر، شایعات و محتوای هرزنامه و غیره استفاده شده‌اند. از این رو، شناسایی و فیلتر کردن این حساب‌های‌کاربری جعلی مورد نیاز بوده و تحقیقات زیادی در این راستا در سال‌های اخیر صورت گرفته است، بنابراین با بررسی انجام شده و دسته بندی الگوریتم‌های بکار گرفته شده می‌تواند به محققین جهت ادامه روند بررسی و شناسایی افراد و حساب‌های جعلی در شبکه‌های اجتماعی کمک نماید..

جدول ۱. پلتفرم‌های رسانه‌های اجتماعی به عنوان مجموعه داده

پلتفرم‌های شبکه‌های اجتماعی	مقالات
Facebook	[24], [26], [30], [35],[39]
Facebook, Twitter	[25]
همه‌ی شبکه‌های اجتماعی	[27]
Blogs	[28]
Twitter	[29], [37], [38]
Instagram	[31],[32]
صفحات وب سایت	[33], [36]
Github	[34]



شکل۳. الگوریتم‌های یادگیری ماشین در روش‌های شناسایی حساب‌های جعلی

مراجع

[1] P. V. Savyan and S. M. S. Bhanu, “Behaviour profiling of reactions in Facebook posts for anomaly detection,” in Proc. 9th Int. Conf. Adv.Comput., 2017, pp. 220–226.

[2] M. A. Wani, N. Agarwal, S. Jabin, and S. Z. Hussain, “Analyzing real and fake users in Facebook network based on emotions,” in Proc. IEEE 11th Int. Conf. Commun. Syst. Netw., 2019, pp. 110–117.

[3] K. Chakraborty, S. Bhattacharyya, and R. Bag, “A survey of sentiment analysis from social media data,” IEEE Trans. Comput. Social Syst.,vol. 7, no. 2, pp. 450–464, Apr. 2020.

[4] P. Chunaev, “Community detection in node-attributed social networks: A survey,” Comput. Sci. Rev., vol. 37, 2020, Art. no. 100286.

[5] A. El Azab, A. M. Idrees, M. A. Mahmoud, and H. Hefny, “Fake account detection in twitter based on minimum weighted feature set,”Int. Scholarly Sci. Res. Innov., vol. 10, no. 1, pp. 13–18, 2016.

[6] E. Karunakar, V. D. R. Pavani, T. N. I. Priya, M. V. Sri, and K. Tiruvalluru,“Ensemble fake profile detection using machine learning (ML),”J.Inf.Comput. Sci., vol. 10, pp. 1071–1077, 2020.

[7] Roy, P. K., & Chahar, S. “Fake profile detection on social networking websites: a comprehensive review”. IEEE Transactions on Artificial Intelligence 1, no.3, pp.271-285, 2020.

[8] K. S. Adewole, N. B. Anuar, A. Kamsin, K. D. Varathan, and S. A. Razak,“Malicious accounts: Dark of the social networks,” J. Netw. Comput. Appl., vol. 79, pp. 41–67, 2017.

[9] P. Wanda and H. J. Jie, “Deepprofile: Finding fake profile in online social network using dynamic CNN,” J. Inf. Secur. Appl., vol. 52, pp. 1–13, 2020.

[10] S. Lee and J. Kim, “WarningBird: A near real-time detection system for suspicious URLs in twitter stream,” IEEE Trans. Dependable Secure Comput., vol. 10, no. 3, pp. 183–195, May/Jun. 2013.

[11] A. J. Sarode and A. Mishra, “Audit and analysis of impostors: An experimental approach to detect fake profile in online social network,” in Proc. 6th Int. Conf. Comput. Commun. Technol., 2015, pp. 1–8.

[12] T. Solorio, R. Hasan, and M. Mizan, “A Case Study of Sockpuppet Detection in Wikipedia,” Proc. Work. Lang. Anal. Soc. Media, no. Lasm, pp. 59–68, 2013.

[13] A. Wang, “Detecting spam bots in online social networking sites: a machine learning approach,” Data Appl. Secur. Priv. XXIV, pp. 335–342, 2010.

[14] Boparai, R. S., & Bhatia, D. “Detectionof Fake Profiles in Online Social Networks–A Survey”, 2022.

[15] H. Drucker, D. Wu, and V. N. Vapnik, “Support vector machines for spam categorization,” IEEE Trans. Neural Net., vol. 10, no. 5, pp. 1048–1054, Sep. 1999

[16] A. H. Wang, “Don’t follow me: Spam detection in twitter,” in Proc. Int. Conf. Securi. Cryptography, 2010, pp. 1–10.

[17] M. Cha, F. Benevenuto, H. Haddadi, and K. Gummadi, “The world of connections and information flow in twitter,” IEEE Trans. Syst., Man, Cybern. A: Syst. Humans, vol. 42, no. 4, pp. 991–998, Jul. 2012.

[18] Z. Miller, B. Dickinson, W. Deitrick, W. Hu, and A. H. Wang, “Twitter spammer detection using data stream clustering,” Inf. Sci., vol. 260, pp. 64–73, 2014.

[19] C. Chen et al., “A performance evaluation of machine learning-based streaming spam tweets detection,” IEEE Trans. Comput. Social Syst., vol. 2, no. 3, pp. 65–76, Sep. 2015.

[20] C. Chen, Y. Wang, J. Zhang, Y. Xiang, W. Zhou, and G. Min, “Statistical features-based real-time detection of drifted twitter spam,” IEEE Trans. Inf. Forensics Secur., vol. 12, no. 4, pp. 914–925, Apr. 2016.

[21] C. Chen et al., “Investigating the deceptive information in twitter spam,” Future Gener. Comput. Syst., vol. 72, pp. 319–326, 2017.

[22] P. K. Roy, J. P. Singh, and S. Banerjee, “Deep learning to filter SMS spam,” Future Gener. Comput. Syst., vol. 102, pp. 524–533, 2020.

[23] Rao, P. S., J. Gyani, and G. Narsimha. "Fake profiles identification in online social networks using machine learning and NLP." Int. J. Appl. Eng. Res 13.6 (2018): 973-4562.

[24] Raturi, Rohit. "Machine learning implementation for identifying fake accounts in social network." International Journal of Pure and Applied Mathematics 118.20 (2018): 4785-4797. J. Wang, “Fundamentals of erbium-doped fibre amplifiers arrays (Periodical style—Submitted for publication),” IEEE J. Quantum Electron., submitted for publication

[25] Kulkarni, Sumit Milind, and Vidya Dhamdhere. "Automatic detection of fake profiles in online social networks." Open access international journal of science and engineering 3.1 (2018): 70-73. M. Young, The Techincal Writers Handbook. Mill Valley, CA: University Science, 1989.

[26] M. Smruthi, N. Harini,” A Hybrid Scheme for Detecting Fake Accounts in Facebook”, International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-7, Issue-5S3, February 2019.

[27] Nambouri Sravya, Chavana Sai praneetha, S. Saraswathi,” Identify the Human or Bots Twitter Data using Machine Learning Algorithms”, International Research Journal of Engineering and Technology (IRJET), Volume: 06 Issue: 03 | Mar 2019 www.irjet.net, e-ISSN: 2395-0056, p- ISSN: 2395-0072.

[28] B. Erçahin, Ö. Aktaş, D. Kilinç, and C. Akyol, ‘‘Twitter fake account detection,’’ in Proc. Int. Conf. Comput. Sci. Eng. (UBMK), Oct. 2017, pp. 388–392, 2017.

[29] F. Ahmed and M. Abulaish, “An MCL-based approach for spam profile detection in online social networks,” in Proc. IEEE 11th Int. Conf. Trust, Secur. Privacy Comput. Commun., 2012, pp. 602–608.

[30] K. Zarei, R. Farahbakhsh, and N. Crespi, “Deep dive on politician impersonating accounts in social media,” in *Proc. IEEE Symp. Comput.Commun.*, Jun. 2019, pp. 1–6.

[31] K. Zarei, R. Farahbakhsh, and N. Crespi, “Typification of impersonated accounts on instagram,” in *Proc. IEEE 38th Int. Perform. Comput. Commun. Conf.*, 2019, pp. 1–6.

[32] S. Revathi and M. Suriakala, “Profile similarity communication matching approaches for detection of duplicate profiles in online social network,” in *Proc. IEEE 3rd Int. Conf. Comput. Syst. Inf. Technol. Sustain. Solutions*, 2018, pp. 174–182.

[33] M. Suriakala and S. Revathi, “Privacy protected system for vulnerable users and cloning profile detection using data mining approaches,” in *Proc. IEEE 10th Int. Conf. Adv. Comput.*, 2018, pp. 124–132.

[34] P. Pourghomi, M. Dordevic, and F. Safieddine, “Facebook fake profile identification: Technical and ethical considerations,” *Int. J. Pervasive Comput. Commun.*, vol. 16, pp. 101–112, 2020.

[35] P.Wanda and H. J. Jie, “Deepprofile: Finding fake profile in online social network using dynamic CNN,” *J. Inf. Secur. Appl.*, vol. 52, pp. 1–13, 2020.

[36] K. S. Adewole, N. B. Anuar, A. Kamsin, and A. K. Sangaiah, “SMSAD: Aframework for spammessage and spam account detection,” Multimedia Tools Appl., vol. 78, no. 4, pp. 3925–3960, 2019.

[37] M. M. Swe and N. N. Myo, “Fake accounts detection on twitter using blacklist,” in Proc. IEEE/ACIS 17th Int. Conf. Comput. Inf. Sci., 2018, pp. 562–566.

[38] M. Egele, G. Stringhini, C. Kruegel, and G. Vigna, “Towards detecting compromised accounts on social networks,” IEEE Trans. Dependable Secure Comput., vol. 14, no. 4, pp. 447–460, Jul./Aug. 2017.

[39] Awan, M.J., Khan, M.A., Ansari, Z.K., Yasin, A. and Shehzad, H.M.F, “Fake profile recognition using big data analytics in social media platforms,”International Journal of Computer Applications in Technology, 2022 , pp.215-222.

[40] H. Shen, B. Wang, X. Liu, and X. Zhang, "Social Spammer Detection via Convex Nonnegative Matrix Factorization," IEEE Access, vol. 10, pp. 91192-91202, 2022.

[41] A. Tandon, S.K. Guha, J. Rashid, J. Kim, M. Gahlan, M. Shabaz, and N. Anjum, "Graph-Based CNN Algorithm to Detect Spammer Activity over Social Media," IETE Journal of Research, pp. 1-11, 2022.

[42] M. Srinivas, A.D. Sai, V. Nikhil, and V. Ramana, "Spammer Detection in Social Networks using ML and NLP," in 2022 International

Conference on Sustainable Computing and Data Communication Systems (ICSCDS), pp. 114-118, IEEE, Apr. 2022.

[43] M.P. Ramesh, K.N.V.P.S.B. Ramesh, C. Revanth, C.V. Sandhyarani, D.N.S. Manikanta, and L. Ramesh, "Spammer Detection and Fake User Identification on Social Networks," Mathematical Statistician and Engineering Applications, vol. 71, no. 4, pp. 5197-5212, 2022.

[44] C. Hema, "The Machine Learning Model for Identifying Bogus Profiles in Social Networking Sites," in Sustainability: Cases and Studies in Using Operations Research and Management Science Methods, Cham: Springer International Publishing, 2023, pp. 67-80.

زیرنویس ها

۱ Online
۲ Sock Puppets
۳ Srinivas Rao
٤ Support Vector Machine
٥ Baysian
٦ Rohit Raturi
۷ Bag of Word
۸ Sumit Milind Kulkarni
۹ Vidya Dhamdhere
۱۰ M. Smruthi
۱۱ N. Harini
۱۲ K-Nearest Neighbor
۱۳ Random Forest
۱٤ Supervised Machine Learning
۱٥ Cross Validation
۱٦ K-fold
۱۷ Nambouri Sravya
۱۸ Erşahin
۱۹ Ahmad
۲۰ Abulaish
۲۱ Zarei
۲۲ TermFrequency-Inverse Document Frequency
۲۳ Convolutional Neural Network
۲٤ Revathi
۲٥ Neural Sstochastic Contraction Metrix
۲٦ Suriakala
۲۷ Revathi
۲۸ Logis Rule
۲۹ Pourghomi
۳۰ Wanda
۳۱ Jie
۳۲ Recall
۳۳ Adewole
۳٤ Swe
۳٥ Myo
۳٦ Awan
۳۷ Shen
۳۸ convex nonnegative matrix factorization
۳۹ Tandon
٤۰ Srinivas
٤۱ Ramesh
٤۲ Hema