

Identifying fake news in light of the right to free access to information

Zahra Askarinejad Amiri*¹
Nastaran Zanjani²
Fatemeh pouramini³
Mohammad Askarinejad Amiri⁴

Abstract

The rise of virtual media has significantly impacted how people communicate in recent years. However, the spread of misleading and suspicious information on social networks poses serious challenges to society. For this reason, detecting false information on these platforms is of great importance. Social networks allow for the rapid production and dissemination of large volumes of diverse content, making it difficult to assess the accuracy and reliability of the information shared. This task requires collaboration between humans and computers, as it cannot be achieved solely through automated systems. At the same time, the right to free access to information is recognized as a fundamental human right in both national and international legal frameworks. Every citizen is entitled to accurate domestic and global news and developments, using the press and other media to express and exchange ideas. This article provides a thorough examination of the methods used to detect online misinformation, commonly referred to as "fake news." Many of these methods focus on analyzing the characteristics of users, the content, and the platforms where the news is published. By exploring these approaches, the research aims to offer recommendations on how to ensure citizens have proper access to accurate information, in line with their right to free access to information.

Keywords: Fake news, right to free access to information, detection of fake news, citizen's right.

1- Assistant Professor, Department of Computer Engineering, Faculty of Refah, Tehran, Iran
askarinejad@refah.ac.ir

2- Assistant Professor, Department of Computer Engineering, Faculty of Refah, Tehran, Iran
zanjani@refah.ac.ir

3- Master's student Computer Engineering at the Faculty of Refah, Tehran, Iran
f.pour67@gmail.com

4- PhD student in law, Islamic Azad University, Ayatollah Amoli Branch, Amol, Iran
m.askarinejad@gmail.com

شناسایی اخبار جعلی

در پرتو حق دسترسی آزاد به اطلاعات

نوع مقاله: پژوهشی

تاریخ دریافت: ۱۴۰۳/۷/۹

تاریخ پذیرش: ۱۴۰۳/۱۰/۸

زهرا عسکری نژاد امیری^{۱*}

نسترن زنجانی^۲

فاطمه پورامینی^۳

محمد عسکری نژاد امیری^۴

چکیده

رشد رسانه‌های مجازی در سال‌های اخیر، تأثیر چشمگیری بر نحوه انتشار و دسترسی به اخبار، از جمله اخبار جعلی داشته است. با این حال، وجود اطلاعات گمراه‌کننده و مشکوک در شبکه‌های اجتماعی، می‌تواند باعث ایجاد مشکلات جدی در جامعه شود و به این دلیل، تشخیص اطلاعات جعلی در شبکه‌های اجتماعی از اهمیت بالایی برخوردار است. در شبکه‌های اجتماعی حجم زیاد و متنوعی از محتوا به سهولت تولید و به سرعت منتشر می‌شود. از این رو تشخیص صحت و سقم اطلاعات در شبکه‌های اجتماعی کاری چالش‌برانگیز است و تنها به کمک رایانه ممکن نبوده، به همکاری بین انسان و رایانه نیاز دارد. از سوی دیگر حق دسترسی آزاد به اطلاعات یکی از حقوق بنیادین بشر در اسناد حقوقی ملی و بین‌المللی است که مورد توجه قرار گرفته است. هر شهروند حق دارد با استفاده از مطبوعات و سایر رسانه‌های نشر عقاید و افکار به درست‌ترین اخبار و اطلاعات داخلی و خارجی و تحولات جهانی دسترسی داشته باشد. این مقاله به صورت جامع روش‌های تشخیص اطلاعات جعلی برخط که اغلب اخبار جعلی نامیده می‌شوند را مورد بحث قرار می‌دهد. بسیاری از این روش‌ها مبتنی بر تشخیص ویژگی‌های کاربران، محتوا و بستر انتشار خبر هستند. سؤال اصلی این مقاله آن است که چه روش‌هایی می‌توانند به بهبود توانایی شهروندان در شناسایی اخبار جعلی کمک کنند؟ نتیجه حاصل از این پژوهش‌ها نشان می‌دهد که استفاده از رویکردهای ترکیبی و ایجاد هم‌افزایی میان انسان و فناوری می‌تواند به ارتقاء کیفیت اطلاعات و کاهش تأثیر اخبار جعلی در شبکه‌های اجتماعی منجر شود. هدف از این پژوهش ارائه راهکارهای مناسب برای شناسایی اخبار جعلی است تا با استمداد از این راهکارها حق دسترسی شهروندان به اطلاعات صحیح تسهیل و هموار گردد.

واژه‌های کلیدی

اخبار جعلی، حق دسترسی آزاد به اطلاعات، تشخیص اخبار جعلی، حق شهروند.

askarnejad@refah.ac.ir

zanjani@refah.ac.ir

f.pour67@gmail.com

m.askarnejad@gmail.com

۱. استادیار گروه مهندسی کامپیوتر دانشکده رفاه، تهران، ایران

۲. استادیار گروه مهندسی کامپیوتر دانشکده رفاه، تهران، ایران

۳. کارشناس مهندسی کامپیوتر دانشکده رفاه، تهران، ایران

۴. دانشجوی دکتری حقوق دانشگاه آزاد اسلامی، واحد آیت‌الله آملی، آمل، ایران

مقدمه

توسعه شبکه جهانی وب و رشد رسانه‌های اجتماعی برخط ارتباط مردم با یکدیگر و به‌اشتراک‌گذاری اطلاعات را به‌طرز چشمگیری متحول کرده است. ویژگی‌هایی از قبیل سهولت استفاده، کم‌هزینه بودن و سرعت بالای رسانه‌های اجتماعی، آن‌ها را به بستر اصلی تعاملات اجتماعی برخط و انتقال اطلاعات تبدیل کرده است (Bigne et al., 2018, pp.1014-1032).

با این حال، به‌دلیل محبوبیت روزافزون رسانه‌های اجتماعی برخط، اینترنت به فضایی ایدئال برای تولید و انتشار اخبار جعلی از قبیل اطلاعات گمراه‌کننده، آگهی‌های جعلی، شایعات و اظهارات سیاسی جعلی تبدیل شده است. حجم انبوهی از اطلاعات باورنکردنی و گمراه‌کننده در این شبکه‌ها تولید و منتشر می‌شوند که به‌عنوان یک تهدید بالقوه برای شبکه‌های اجتماعی برخط به حساب می‌آید و تأثیر منفی عمیقی بر فعالیت‌های اینترنتی مانند خرید برخط می‌گذارد. با انتقال اطلاعات مغرضانه و دروغین، اخبار جعلی می‌توانند اعتماد و باور مردم به مقامات، متخصصان و دولت را از بین ببرند؛ به‌عنوان مثال ۸۸٪ از مشتریان بر اساس نظرهای کاربران خرید می‌کنند و ۷۲٪ از آنها به کسب‌وکاری که دارای بازخورد مثبت از سوی مشتریان است اعتماد دارند (Ahmed, 2017). گسترش اخبار جعلی می‌تواند این روند را خدشه دار کند. اخبار جعلی این روزها همچنان حاکم بر اینترنت است و می‌تواند عواقب سرنوشت‌سازی برای جامعه داشته باشد. بنابراین نیازی جدی برای تولید یک سیستم دقیق جهت شناسایی اخبار جعلی برخط وجود دارد.

حق دسترسی به اطلاعات یکی از اصول بنیادین حکومت‌های مردم‌سالار و از مصادیق مهم نسل اول حقوق بشر و تضمین‌کننده حقوق شهروندی است (اسلامی و دیگران، ۱۴۰۰، ص. ۹۵-۱۲۳). حقوق بشر به‌عنوان حقوق پایه و اولیه، تکلیف و تعهدی بنیادین است که مشروعیت آن مشروط به طرح و تصریح در قانون اساسی نیست؛ اما حقوق شهروندی تنها در یک رابطه حق و تکلیفی وضعی قابل تحقق است و به موازات هر حقی برای شهروند، تکلیفی نیز وجود دارد. شناسایی و تضمین حق مزبور با گسترش سریع شبکه‌های اجتماعی به یکی از مهمترین مصادیق تحقق حقوق شهروندی در نظام حقوقی ایران تبدیل شده است.

اگرچه قانون انتشار و دسترسی آزاد به اطلاعات در سال ۱۳۸۸ به تصویب رسید، در سال‌های اولیه با چالش‌های اجرایی متعددی از جمله ابهام در تعاریف و تأخیر در تشکیل کمیسیون ماده ۱۸ مواجه بود (محسنی و دیگران، ۱۳۹۸، صص. ۳۲۱-۳۵۴). اما بر اساس آخرین گزارش‌های رسمی کمیسیون ماده ۱۸ به طور رسمی از سال ۱۳۹۷ با حضور نمایندگان سه قوه شروع به کار کرده و تا پایان سال ۱۴۰۱ به بیش از ۵۰۰ پرونده رسیدگی نموده است (معاونت حقوقی رئیس جمهور، ۱۴۰۲). ضمناً دیوان عدالت اداری تنها در سال ۱۴۰۱ بیش از ۱۲۰۰ پرونده مرتبط با این قانون را ثبت کرده که نشان‌دهنده استقبال شهروندان از این سازوکار است (سالنامه آماری قوه قضاییه، ۱۴۰۲). علاوه بر این پرونده‌های شاخص مانند پرونده دسترسی به اطلاعات محیط زیستی شرکت‌های پتروشیمی (پرونده شماره ۱۴۰۰۱ در دیوان عدالت اداری) و درخواست انتشار اسناد قراردادهای دولتی (پرونده ۱۴۰۱۱۲ در کمیسیون ماده ۱۸) از موارد عینی اجرای این قانون محسوب می‌شوند (مرکز پژوهش‌های مجلس، ۱۴۰۱). با این وجود، گزارش مرکز مطبوعات و انتشارات قوه قضاییه (۱۴۰۲) نشان می‌دهد که حدود ۴۰٪ از آراء صادره در این حوزه، به دلیل نقص در آیین‌نامه‌های اجرایی با مشکل مواجه شده‌اند که لزوم بازنگری در مقررات اجرایی را نشان می‌دهد.

در بند ۲ اصل سوم قانون اساسی یکی از وظایف دولت جمهوری اسلامی ایران را «بالا بردن سطح آگاهی‌های عمومی در همه زمینه‌ها با استفاده صحیح از مطبوعات و رسانه‌های گروهی و وسایل دیگر» عنوان شده است. متقابل این تکلیف، می‌تواند حق شهروندان بر افزایش آگاهی عمومی و اطلاع رسانی صحیح را نیز عنوان کرد. این اقدام خود شامل دو مرحله پیوسته و درهم‌تنیده می‌باشد: اولین مرحله ارائه اخبار صحیح توسط رسانه‌ها و دوم رد اخبار جعلی و اطلاعات خلاف واقع است (وزیری و دیگران، ۱۳۹۰، صص. ۱۰۷-۱۲۶). در عصر دیجیتال، شبکه‌های اجتماعی به بستری برای انتشار سریع اخبار-اعم از صحیح و جعلی-تبدیل شده‌اند. تشخیص اخبار جعلی به دلیل حجم بالای تولید محتوا و پیچیدگی‌های فنی، نیازمند راهکارهای هوشمند و مشارکت شهروندان است. از طرفی، مصوبه شورای عالی فضای مجازی در خصوص الزامات پیشگیری و مقابله با نشر اطلاعات، اخبار و محتوای خبری خلاف واقع در فضای مجازی تلاش کرده تا با ارائه

چارچوب‌های قانونی، این پدیده را مهار کند. اما پرسش اصلی این است که آیا این مصوبه به‌درستی به تعادل میان مبارزه با اخبار جعلی و حق دسترسی آزاد به اطلاعات پرداخته است؟ حال سوال اینجاست چگونه می‌توان با شناسایی اخبار جعلی میزان دسترسی شهروندان به اطلاعات صحیح مبتنی بر حق دسترسی آزاد به اطلاعات را بهبود بخشید؟ به‌منظور تشخیص اخبار جعلی، بسیاری از سیستم‌های راستی‌آزمایی برخط مانند *Snopes.com* و *PolitiFact.com* مبتنی بر رویکردهای تشخیص انسانی توسط متخصصان هستند. تأخیر زمانی و محدود بودن این سیستم‌ها به صحت‌سنجی اخبار سیاسی باعث می‌شود تا در راستی‌آزمایی سایر موضوعات متنوعی که در حجم زیاد و با سرعت در شبکه‌های اجتماعی منتشر می‌شوند کاربردی نداشته باشند.

در سال‌های اخیر، برای کمک به کاربران برخط در شناسایی اطلاعات صحیح، تحقیقات گسترده‌ای در جهت ایجاد یک چهارچوب مؤثر و خودکار برای کشف اخبار جعلی برخط صورت گرفته است. باوجوداین، شناسایی اطلاعات معتبر از بین میلیون‌ها پیام، به‌دلیل ماهیت پویا و ناهمگن ارتباطات اجتماعی چالش‌برانگیز است. از آنجایی‌که داده‌های اجتماعی برخط، به‌صورت بلادرنگ ایجاد می‌شوند، برای کشف، کاوش و تفسیر اطلاعات جعلی در رسانه‌های اجتماعی برخط، به یک سیستم شناسایی برخط بلادرنگ نیاز است. به‌علاوه تنها به‌کارگیری ویژگی‌های زبانی محتوای اخبار برای آشکار ساختن الگوهای اخبار جعلی کافی نبوده و ویژگی‌هایی دیگری از جمله اعتبار نویسنده خبر و الگوهای انتشار خبر نیز باید در شناسایی اخبار جعلی برخط به‌کار گرفته شوند. به این منظور اخبار جعلی را می‌توان از چهار منظر مختلف شامل: تولیدکننده خبر، محتوای خبر، بستر اجتماعی و گروه خبری جعلی، اهداف آنها، سبک‌های نوشتاری غالب در اخبار جعلی، چگونگی توزیع اخبار جعلی از طریق اینترنت و نحوه تأثیرگذاری آنها بر خوانندگان برخط ایجاد می‌کند.

مقابله قانونی و قضایی با پدیده اخبار جعلی از دیگر راهکارهای ایجاد حق دسترسی آزاد به اطلاعات می‌باشد و در نظام حقوقی ایران جرم نشر اکاذیب مشابه‌ترین مقابله قانونی با اخبار جعلی می‌باشد؛ اما بایستی در تدوین لایحه قضایی مقابله با اخبار جعلی

که یکی از وظایف قانونی قوه قضائیه می‌باشد؛ به تفاوت‌های بین این دو مفهوم توجه داشت؛ بنابراین لازمه این امر شناسایی اخبار جعلی است. با توجه به اهمیت شناسایی اخبار جعلی، این مقاله به بررسی و مقایسه رویکردهای گوناگونی که در این حوزه مورد استفاده قرار گرفته است می‌پردازد. بدین‌منظور با مطالعه مقالات منتشر شده در کتابخانه‌های برخط و مراکز اسناد همچون IEEE, Elsevier, Scopus و ...، و ابتدا رویکردهای مختلف شناسایی اخبار جعلی را بررسی کرده، سپس به مقایسه رویکردهای مختلف شناسایی به کمک نیروهای انسانی و شناسایی خودکار پرداخته‌ایم. این پژوهش با معرفی این رویکردها به دنبال بررسی ابعاد حق دسترسی آزاد به اطلاعات و اجرای آن می‌باشد.

۱. ادبیات پژوهش

امروزه آزادی و حق دسترسی به اطلاعات جزو حقوقی است که به تحقق مفهوم شهروندی که یکی از موضوع‌های اصلی در گفتمان مردم‌سالاری مشارکتی است کمک می‌کند. به‌واقع هر اندازه بر کمیت و کیفیت اطلاعات قابل دسترسی شهروندان افزوده شود به همان میزان مردم‌سالاری فراگیرتر و مشارکت بیشتر می‌شود؛ بنابراین به رسمیت شناختن این حق نه تنها موجب تغییر ساختار حکومت‌ها که سبب تغییر در نقش سنتی دولت نیز می‌شود، به این معنی که دولت‌ها در پرتوی اجرای قوانین آزادی اطلاعات و حق دسترسی، بر خلاف گذشته که مالک اطلاعات بودند به تدریج به واسطه‌ای اطلاعاتی تبدیل خواهند شد (نمک‌دوست تهرانی، ۱۳۸۴، صص. ۶۳-۱۰۸). در فقه اسلام و حقوق موضوعه ایران نیز با تصویب قانون «انتشار و دسترسی آزاد به اطلاعات»، این حق مورد پذیرش قرار گرفت. براساس قانون مذکور، مؤسسات عمومی و خصوصی، موظفند به‌جز در موارد استثناء، اطلاعات مضبوط را در اختیار شهروندان قرار دهند. بعضی اصول به‌دست آمده عبارت‌اند از: آزادی بیان، مشاوره، حق نظارت، حق آگاهی، امانی بودن اطلاعات، نهی از کتمان علم و شفافیت (وزیری و دیگران، ۱۴۰۱، صص. ۱۰۷-۱۲۶). در عصر حاضر آگاهی از رویدادهای زندگی جمعی، از جمله احتیاجات بسیار ضروری انسان، به‌شمار می‌رود. در نظام حقوق بین‌الملل آزادی اطلاعات و اخبار به‌عنوان یکی از حقوق اساسی بشر به‌شمار می‌آید که علاوه بر اهمیت ذاتی، به‌دلیل

کارکرد و پشتیبانی آن از دیگر حقوق، از جایگاه ویژه‌ای برخوردار است؛ بنابراین از یک سو می‌توان گفت که اطلاعات نقش اساسی و غیرقابل انکار در تضمین یک جامعه مردم‌سالارانه دارد و از سوی دیگر در نظام‌های حقوقی بین‌المللی و ملی چنین اتفاق نظری وجود دارد که قلمرو آزادی اطلاعات و اخبار مطلق نبوده و برای آن محدودیت‌هایی ذکر شده است که از جمله مهم‌ترین و مبهم‌ترین این محدودیت‌ها می‌توان به محدودیت اخلاق عمومی اشاره نمود (رهایی و دیگران، ۱۳۹۹، صص ۲۳-۵۶).

مصوبه «الزامات پیشگیری و مقابله با نشر اطلاعات، اخبار و محتوای خبری خلاف واقع در فضای مجازی» که توسط شورای عالی فضای مجازی تصویب شده است، چهارچوبی برای شناسایی، برخورد و محدودسازی انتشار اخبار نادرست در سکوه‌های دیجیتال تعیین می‌کند. این مصوبه در پاسخ به نگرانی‌های مربوط به تأثیرات مخرب اخبار جعلی بر امنیت ملی، نظم عمومی و اعتماد اجتماعی طراحی شده است. اهداف اصلی مصوبه عبارتند از: کاهش انتشار اخبار جعلی در فضای مجازی، مسئولیت‌پذیری سکوها (شبکه‌های اجتماعی، رسانه‌های خبری برخط ...) در قبال محتوای منتشرشده، حفظ امنیت روانی و اجتماعی در برابر اطلاعات گمراه‌کننده، هماهنگی بین نهادهای نظارتی (مانند پلیس فتا، مرکز ملی فضای مجازی و سازمان‌های رسانه‌ای) (شورای عالی فضای مجازی، ۱۴۰۲).

به هر حال بررسی روش‌های مختلف تشخیص اخبار جعلی می‌تواند راهگشای دولت‌ها در ایجاد حق دسترسی آزاد به اطلاعات باشد. علاوه بر این امر می‌تواند هم‌راستا با اهداف مصوبه شورای عالی فضای مجازی باشد. به جهت اهمیت این موضوع در ادامه به رویکردهای شناسایی اخبار جعلی پرداخته شده است.

۲. شناسایی اخبار جعلی

در شناسایی اخبار جعلی می‌توان به دو رویکرد کلی اشاره کرد: یکی شناسایی به کمک نیروهای انسانی که در آن متخصص به بررسی محتویات اینترنتی در حوزه‌های مختلف پرداخته و جعلی یا واقعی بودن آنها را تشخیص می‌دهند. تارنمای PolitiFact.com به‌عنوان یک مرجع برای ارزیابی صحت ادعاها و اظهارات مقامات منتخب، اندیشمندان، نویسندگان، وبلاگ‌نویسان، تحلیل‌گران سیاسی و دیگر اعضای رسانه شناخته می‌شود. ویراستاران

این تارنما، یک کلمه یا عبارت کلیدی و متن کامل یک ادعا را به دقت بررسی کرده و صحت آن را ارزیابی می‌کنند. این سکو به کاربران امکان می‌دهد تا به متن کامل بیانیه‌ها، داستان‌ها، وعده‌ها و به‌روزرسانی‌های بررسی شده دسترسی داشته باشند. تارنمای دیگر این بخش، طیف گسترده‌ای از حوزه‌ها از جمله خودرو، تجارت، رایانه، جرم، فریب و کلاهبرداری، تاریخ و... را دربرمی‌گیرد. به کمک دانش افراد و سازمان‌های حرفه‌ای می‌تواند ارزیابی‌های جامعی را برای انواع مختلف منابع چاپی ارائه دهد و آنها را از نظر راستی‌آزمایی رتبه‌بندی کند.

روش دیگری در شناسایی اخبار جعلی توسط انسان وجود دارد که صحت‌سنجی اطلاعات از طریق جمعیت است تارنمای Fiskkit² یک سکو برای تحلیل و بحث پیرامون اخبار و مقالات است که به کاربران این امکان را می‌دهد تا ادعاهای موجود در مقالات خبری را به‌طور دقیق بررسی و تحلیل کنند. کاربران می‌توانند بخش‌های مختلف یک مقاله را علامت‌گذاری کنند و به آنها امتیاز صحت، منطقی و اهمیت بدهند. این تارنما به‌نوعی بستری برای بحث و نقد گروهی است که به کاربران اجازه می‌دهد تا اطلاعات و ادعاها را بررسی کرده و دیدگاه‌های مختلف را با شواهد و دلایل مستدل به چالش بکشند. هدف این تارنما ارتقای تفکر نقادانه و افزایش دقت در مصرف اخبار است. با این حال این روش نوپاست و نتایج حاصل از آن در مقایسه با صحت‌سنجی مبتنی بر خبرگان، به دلیل سوگیری سیاسی افراد و حاشیه‌نویسی متناقض آنها کمتر قابل اعتماد است، از این رو، در این روش اغلب نیاز به فیلتر کردن کاربران غیرمعتبر و تحلیل نتایج متناقض صحت‌سنجی وجود دارد (Di et al., 2018, pp.124-130). باید توجه داشت که روش تشخیص انسانی فرایندی زمان‌بر و پرهزینه بوده، نیازمند دخالت گسترده انسان می‌باشد؛ بنابراین ایجاد بسترهای شناسایی اخبار جعلی خودکار ضروری به‌نظر می‌رسد. در این‌راستا تعدادی مراجع برخط صحت‌سنجی اخبار وجود دارند که به تعدادی از آنها اشاره می‌کنیم.

Classify.news فضایی برای شناسایی برخط مقالات جعلی است. هدف آن ساختن مدلی برای تشخیص اعتبار یک مقاله صرفاً بر اساس محتوای متنی آن با استفاده از الگوریتم‌های

1. Snopes.com
2. Fiskkit.com

یادگیری ماشین است. این یک سکوی برخی است که به تحلیل و طبقه‌بندی اخبار می‌پردازد. این تارنما از تکنیک‌های یادگیری ماشینی و هوش مصنوعی برای طبقه‌بندی محتوای خبری استفاده می‌کند. هدف این تارنما این است که کاربران بتوانند به راحتی اخبار را بر اساس دسته‌بندی‌های مختلف مانند موضوع، منبع و میزان سوگیری سیاسی فیلتر کنند. همچنین به تحلیل صحت و اعتبار منابع خبری نیز می‌پردازد تا به کاربران در انتخاب منابع قابل اعتماد کمک کند. روش کار به این صورت است که مقالات خبری برچسب زده شده را جمع‌آوری می‌کنند و بر اساس نمونه‌های معتبر و غیرمعتبر سیستم را آموزش می‌دهند. از دو نوع مدل پیش‌بینی «صرفاً مبتنی بر متن» به روش دسته‌بندی‌کننده بیز چندجمله‌ای و «صرفاً مبتنی بر مفهوم» به روش دسته‌بندی‌کننده بوستینگ تطبیقی در این سیستم استفاده می‌شود.

Factmata.com: یک پروژه سرمایه‌گذاری شده توسط گوگل می‌باشد که به صحت‌سنجی آماري ادعاها به کمک الگوریتم‌های هوش مصنوعی و یادگیری ماشین می‌پردازد. این یک سکوی مبتنی بر هوش مصنوعی است که برای شناسایی و مقابله با اطلاعات نادرست و سوگیری در محتواهای برخی طراحی شده است. این تارنما از فناوری‌های پیشرفته پردازش زبان طبیعی (NLP) و یادگیری ماشینی استفاده می‌کند تا محتوای رسانه‌ای، مقالات خبری و پست‌های شبکه‌های اجتماعی را تجزیه و تحلیل کرده و میزان صحت، سوگیری، تبلیغات سیاسی (پروپاگاندا)، و محتوای نفرت‌پراکن یا گمراه‌کننده را شناسایی کند. Factmata به ناشران، برندها و کسب‌وکارها کمک می‌کند تا اعتبار و کیفیت محتوای خود را ارتقا دهند و از انتشار اطلاعات نادرست جلوگیری کنند.

پس می‌توان گفت یکی از اهداف کلیدی مصوبه شورای عالی فضای مجازی در زمینه مقابله با اخبار خلاف واقع، کاهش انتشار اخبار جعلی از طریق مکانیسم‌های شناسایی کارآمد است. در این راستا، رویکردهای مبتنی بر مشارکت نیروهای انسانی به عنوان راهکاری مؤثر مورد توجه قرار گرفته‌اند (شورای عالی فضای مجازی، ۱۴۰۲). این روش‌ها ترکیبی از تخصص انسانی و صحت‌سنجی جمعی را می‌تواند به کار می‌گیرند تا هم از دقت بالاتری برخوردار باشند و هم از محدودیت‌های روش‌های کاملاً خودکار بکاهند. در

این بخش به برخی از این موارد اشاره شده است. اما باید در نظر داشت که راهکار عملی جهت اجرای آن باید در فضای جامعه شکل گیرد.

این تارنماها الگوریتم‌های هوش مصنوعی و یادگیری ماشین را در صحت‌سنجی اخبار جعلی به کار می‌گیرند. با این حال Factmata.com فقط می‌تواند ادعاها یا اظهاراتی را که حاوی اطلاعات آماری است بررسی کند و Classify.news عمدتاً از روش‌های با نظارت یادگیری ماشینی که به مجموعه داده‌ها و برچسب‌های آموزشی با کیفیت بالا نیاز دارند، استفاده می‌کند؛ لذا مسائل و مشکلات کشف اخبار جعلی، محققان و دست‌اندرکاران را برای ایجاد مدل‌های نظام‌مند و رویکردهای پیش‌بینی سازگار، خودکار و جامع ترغیب می‌کند. حال سؤال اینجاست که امروزه با تشخیص اخبار جعلی مبتنی بر روش‌های مذکور وظایف دولت‌ها جهت تبیین و استفاده از این روش‌ها چیست؟ شناسایی اخبار جعلی مبتنی بر روش‌های موجود و صحت‌سنجی آن می‌تواند مسیری روشن برای دولت‌ها جهت ایجاد حق دسترسی آزاد ارائه کند. بررسی این امر می‌تواند به دولت‌ها کمک کند تا راهکارهایی برای فراهم کردن دسترسی آزاد به اطلاعات را پیشنهاد دهند. با این حال، چالش‌هایی مانند سوگیری‌های شناختی، کمبود نیروهای متخصص و هزینه‌بر بودن فرآیندهای انسانی، لزوم تلفیق این روش‌ها با فناوری‌های پیشرفته را آشکار می‌سازد (Lewandowsky et al., 2020). در همین راستا، مصوبه خلاف واقع شورای عالی فضای مجازی بر توسعه سیستم‌های ترکیبی انسان-ماشین تأکید دارد که در آن الگوریتم‌های هوش مصنوعی اخبار مشکوک را غربال می‌کنند و نیروهای انسانی به عنوان مرحله نهایی راستی‌آزمایی عمل می‌نمایند. از این رو، با انجام اعتبارسنجی خودکار اخبار جعلی، می‌توان به راهبردهای مؤثری برای این منظور دست یافت.

۳. صحت‌سنجی خودکار اخبار جعلی

با توجه به اینکه سازندگان اخبار جعلی اغلب از ابزارها و فناوری‌های پیشرفته برای گمراه کردن افکار عمومی استفاده می‌کنند، قوانین مربوط به حفاظت از اطلاعات باید به گونه‌ای طراحی شوند که علاوه بر جلوگیری از نقض حریم خصوصی، امکان ردیابی و شناسایی

منابع اخبار جعلی را نیز فراهم کنند؛ به عنوان مثال، ایجاد شفافیت در عملکرد رسانه‌ها و ملزم کردن آن‌ها به انتشار منابع خبری می‌تواند به کاهش انتشار اخبار جعلی کمک کند. نحوه اجرای سیستم تشخیص جعلی اخبار را می‌توان به صورت زمان واقعی و غیربرخط طبقه‌بندی کرد. در سیستم تشخیص غیربرخط، معمولاً از مدل‌های یادگیری ماشین به صورت دسته‌ای برای شناسایی اخبار جعلی استفاده می‌شود. سیستم تشخیص غیربرخط برای طبقه‌بندی اخبار جعلی برخط مهم است؛ زیرا می‌تواند اطلاعات غیرعادی را به صورت توصیفی مورد تجزیه و تحلیل قرار دهد، از جمله انتخاب تأثیرگذارترین ویژگی‌ها برای تمایز اطلاعات دروغین در بین مقادیر زیادی پیام‌های اجتماعی. دسته‌بندی غیربرخط می‌تواند بر اساس انواع اطلاعات برخط، به تشخیص نظرهای جعلی، هجویات، فریبکاری‌ها، اخبار سیاسی و برخی از زمینه‌های مرتبط دیگر، مانند تشخیص شایعه، تشخیص تله‌کلک، شناسایی اسپمر و تشخیص ربات پردازد (Shu et al., 2019, pp.43-65).

با این حال، سیستم غیربرخط محدود است. مجموعه داده‌های مورد استفاده ممکن است ویژگی‌های اساسی اخبار جعلی برخط را نشان ندهند و مدل‌های یادگیری آموزش دیده در یک سیستم غیربرخط، در شرایط دیگر پاسخگو نباشد. در سیستم تشخیص بلادرنگ، روش‌های مختلف تجزیه و تحلیل بلادرنگ برای تعیین جعلی بودن یا نبودن اطلاعات اجتماعی استفاده می‌شود. با استفاده از روش‌های تحلیلی پیش‌بینی‌کننده بر روی اطلاعات در زمان واقعی، تجزیه و تحلیل بلادرنگ می‌تواند کاربرد روش‌های غیربرخط را بهبود بخشد. مطالعات کمی در زمینه کشف اخبار جعلی بلادرنگ وجود دارد؛ به عنوان مثال، شائو و همکارانش یک سیستم مصورسازی بلادرنگ برای تجزیه و تحلیل اطلاعات غیرعادی در توئیتر ایجاد کرد (Shao et al., 2016, pp.745-750). ایجاد یک سیستم برخط مؤثر، چالش‌برانگیز است؛ زیرا داده‌های ارتباطات برخط حساس به زمان، مداوم و ناهمگن هستند. با وجود این، سیستم تشخیص برخط ابزاری قدرتمند برای گرفتن ماهیت پویای اطلاعات برخط و مبارزه با اخبار جعلی برخط است و باید در آینده مورد توجه بیشتری قرار بگیرد.

در صحت‌سنجی خودکار اخبار به صورت غیربرخط دو نوع رویکرد مبتنی بر اجزاء خبر و مبتنی بر روش‌های داده‌کاوی وجود دارد (Oshikawa et al., 2018). در رویکردهای

مبتنی بر اجزاء، چهار مؤلفه اصلی برای خبر در نظر گرفته می‌شود که عبارت‌اند از: سازنده اخبار جعلی، مصرف‌کننده اخبار جعلی، محتوای اخبار جعلی و بستر اجتماعی خبر جعلی. ربات‌های اجتماعی الگوریتم‌های رایانه‌ای هستند که برای نمایش رفتارهای شبیه به انسان طراحی شده‌اند و به‌طور خودکار محتوا تولید می‌کنند و در رسانه‌های اجتماعی با انسان تعامل دارند (Zhao, 2013, p.21). بسیاری از ربات‌ها به‌طور خاص برای توزیع شایعات، اسپم و بدافزارها، طراحی شده‌اند. سایبورگ‌ها به ربات‌های کمکی انسان اشاره دارد. پس از ثبت‌نام توسط یک شخص، سایبورگ می‌تواند توییت ارسال کرده و در شبکه‌های اجتماعی شرکت کند. مشابه ربات‌های اجتماعی، حساب‌های مخرب سایبورگ با انتشار اطلاعات و پیام‌های جعلی، کاربران اجتماعی برخط را گمراه کرده، از آنها سوءاستفاده می‌کنند که ممکن است منجر به آسیب رساندن به اعتماد اجتماعی شود.

تحلیل محتوای اخبار مؤلفه دیگر از صحت‌سنجی اخبار است. یک خبر حاوی محتوای فیزیکی (مانند عنوان، متن، شکل یا ویدیو) و محتوای غیرفیزیکی (مانند هدف، عقیده) است. با بهره‌گیری از تجزیه و تحلیل عمیق محتوای اخبار، می‌توانیم الگوهای زبانی و سبک‌های نوشتن را برای اخبار واقعی و جعلی مورد تجزیه و تحلیل قرار دهیم، و سپس از مهم‌ترین ویژگی‌ها برای تشخیص اخبار جعلی برخط استفاده کنیم. مطالعات فعلی در مورد تجزیه و تحلیل محتوای اخبار، یا زبانی و معنایی هستند یا مبتنی بر سبک.

هر دو تحلیل زبانی و معنایی مبتنی بر مطالعات کلاسیک و علمی زبان طبیعی هستند. با استخراج اطلاعات مفید از محتوای اخبار، تجزیه و تحلیل زبانی و معنایی می‌تواند الگوهای زبان انگلیسی، ساختارها و معانی اخبار را تحلیل کند. بیشتر سازندگان اخبار جعلی برای جلوگیری از شناسایی، از راهبردهای مخصوص نوشتن استفاده می‌کنند. هدف اولیه تحلیل زبانی تطبیق صلاحیت زبانی نویسنده خبر با مشاهده قالب‌های زبانی و کشف الگوهای نوشتن است. "Bag-of-words" و "n-grams" متداول‌ترین روش‌ها برای نمایش متون خبری خام است. در "Bag-of-words" با در نظر گرفتن هر کلمه به‌عنوان واحد منفرد و یکسان، متن خبر خام می‌تواند به‌عنوان مجموعه کلمات آن، صرف‌نظر از دستور زبان و دستور کلمه ارائه شود. در "n-gram"، متن خبری با دنباله‌ای از n آیتم نمایش داده می‌شود،

این آیتم‌ها می‌توانند واج‌ها، هجاها، حروف یا کلمات باشند. با این حال سادگی این دو رویکرد منجر به برخی کاستی‌های واضح در پردازش متن خام می‌شود؛ برای نمونه، مدل "n-gram" دارای پراکندگی زیاد است و نمی‌تواند نمونه‌های خبری را که حاوی نشانه‌های ناشناخته است تفسیر کند. "bag-of-words" ممکن است با نادیده گرفتن بستر و معنای کلمات، اطلاعات قابل‌توجهی را از دست بدهد. از این رو از رویکردهای تحلیل معنایی استفاده می‌شود که به فرایند توصیف ساختارهای نحوی اخبار از سطح عبارات گرفته تا سطح معنایی اشاره دارد. با ترکیب کردن مدل "n-gram" با مدل "deep syntax"، تجزیه و تحلیل معنایی می‌تواند میزان سازگاری بین تجربه شخصی سازنده اخبار و محتوای آن را کشف کند؛ به عنوان مثال سازندگان اخبار جعلی غالباً برای جلب توجه خوانندگان از عناوین مبالغه‌آمیز استفاده می‌کنند؛ بنابراین عنوان اخبار جعلی معمولاً نامربوط یا مغایر با محتوای اخبار است درحالی‌که عنوان یک خبر واقعی غالباً با محتوای بدنه خبر مطابقت دارد. مثال دیگر نظرهای کاربران برخط است که اگر جعلی باشد ممکن است حاوی تضاد یا اشتباه باشد؛ زیرا نظردهندگان فریبکار تجربه‌ای در مورد ویژگی‌ها و خدمات مربوط به محصولات ندارند. محققان می‌توانند با استفاده ترکیبی از تجزیه و تحلیل سازنده اخبار و تجزیه و تحلیل معنایی، سازگاری بین پیشینه کاربر و محتوای اخبار را بررسی کنند که این امر پیشرفت خوبی را برای طبقه‌بندی و کشف اطلاعات نادرست نشان داده است.

وجود نشانه‌های مشکوک مانند تعداد URLها، هشتک‌ها، اشاره کردن به دیگران و کلمات با حروف بزرگ در داده‌های ارتباطات اجتماعی از ویژگی‌های خوبی برای شناسایی تحلیل سبک نوشتن است (Zhao et al., 2013, pp.677-686). این در حالی است که تجزیه و تحلیل سبک غیرفیزیکی به بررسی جنبه‌های غیرفیزیکی اخبار، مانند پیچیدگی و خوانایی متن خبر می‌پردازد. سازندگان خبر جعلی معمولاً زمان بیشتری برای نوشتن جملات نیاز دارند و در هنگام نوشتن اشتباه می‌کنند؛ بنابراین برخی از الگوهای خاص استفاده از کلیدها برای تحلیل سبک نوشتن قابل ردیابی هستند؛ به عنوان مثال کلید delete و backspace بیشتر وقتی استفاده می‌شود که یک سازنده اخبار جعلی می‌خواهد پیغام غیرواقعی بنویسند. تحلیل بستر اجتماعی، مطالعه چگونگی توزیع سریع و گسترده اطلاعات اجتماعی و نحوه تعامل

کاربران برخط با یکدیگر است. با این حال، بسیاری از رویکردهای اخیر برای کشف اخبار جعلی برخط مربوط به تجزیه و تحلیل محتوای مستقیم اخبار است. محدود مطالعاتی از تجزیه و تحلیل بستر اجتماعی برای پیش‌بینی اطلاعات برخط غیرعادی استفاده می‌کنند. با الگوبرداری کمی از الگوی توزیع داده‌های ارتباطات اجتماعی، تجزیه و تحلیل بسترهای اجتماعی قادر به پیش‌بینی‌های اولیه از سرعت انتشار، مقیاس و تأثیر انتشار اطلاعات و ارائه دیدگاه‌های متنوع و عمیق برای نمایش اخبار جعلی برخط است. در این راستا دو نوع تجزیه و تحلیل بسترهای اجتماعی پیشنهاد می‌شود که می‌توانند نامزدهای بالقوه‌ای برای افزایش عملکرد روش‌های موجود برای کشف اخبار جعلی باشند که عبارت‌اند از تجزیه و تحلیل شبکه کاربر و تجزیه و تحلیل الگوی توزیع هستند.

از سوی دیگر تجزیه و تحلیل الگوی توزیع، مطالعه‌ای برای تجزیه و تحلیل ویژگی‌های انتشار اطلاعات است (Zhao et al., 2013, pp.677-686). الگوهای انتشار اخبار برخط می‌تواند اطلاعات ارزشمند و آموزنده‌ای را برای پیشنهاد وقایع و پیام‌های غیرعادی فراهم کند. در سال‌های اخیر، بسیاری از محققان در حال کار بر روی تشخیص الگوی غیرعادی برای اطلاعات اجتماعی هستند که می‌تواند به کاربران برخط کمک کند داده‌های متفاوت، غیرمعمول یا غیرمنتظره را کشف کنند؛ اما، کشف الگوی توزیع مشکوک برای اخبار جعلی، به دلیل ماهیت ناهمگن و پویای رفتارهای اجتماعی برخط پیچیده و چالش‌برانگیز است (Yang et al., 2012, pp.1-7). در نتیجه، سیستم‌های مصورسازی^۱ پیشرفته همیشه به‌منظور پاسخگویی به چالش‌های فوق، با الگوریتم‌های یادگیری ماشین کلاسیک همراه هستند. به‌طور خلاصه، اطلاعات فراوانی هر روز از طریق رسانه‌های اجتماعی پخش می‌شود و بیشتر اوقات، تجزیه و تحلیل محتوای اخبار برای ساختن یک سیستم مؤثر برای کشف اطلاعات کاذب کافی نیست. با الگوبرداری کمی از الگوی توزیع داده‌های ارتباطات اجتماعی، تجزیه و تحلیل بسترهای اجتماعی قادر به پیش‌بینی‌های اولیه از سرعت انتشار، مقیاس و تأثیر انتشار اطلاعات و ارائه دیدگاه‌های متنوع و عمیق برای نمایش اخبار جعلی برخط است.

۴. شناسایی اخبار جعلی به صورت خودکار مبتنی بر داده‌کاوی

روش دیگر شناسایی اخبار جعلی به صورت خودکار رویکردهای مبتنی بر داده‌کاوی است که در ادامه به بررسی آن می‌پردازیم. بر اساس دسته‌های مختلف تجزیه و تحلیل مؤلفه‌ها، انواع مختلفی از ویژگی‌ها را می‌توان از داده‌های ارتباطات برخط استخراج کرد و برای یادگیری ساخت مدل استفاده کرد. مدل‌های یادگیری ماشین موجود را می‌توان به سه رویکرد مدل‌های نظارت‌شده، نیمه‌نظارتی و بدون نظارت تقسیم‌بندی کرد (Afroz et al., 2012, pp.461-475). در این مقاله، مدل‌های یادگیری نظارت‌شده و مدل‌های یادگیری بدون نظارت عمده‌تاً برای کشف اخبار جعلی برخط مورد بحث قرار می‌گیرند.

الگوریتم‌های یادگیری ماشین نظارت‌شده مانند درخت تصمیم‌گیری^۱، جنگل تصادفی^۲، ماشین بردار پشتیبانی^۳، رگرسیون لجستیک^۴ و K نزدیکترین همسایه^۵ به‌طور گسترده برای شناسایی کلاهبرداری‌های برخط و طبقه‌بندی اطلاعات فریبده استفاده شده است. عملکرد یک مدل یادگیری نظارت‌شده به شدت به کیفیت مجموعه داده دارای برچسب بستگی دارد. ولی تولید یک مجموعه داده گسترده و با کیفیت خوب برای کشف اخبار جعلی دشوار است، چرا که اولاً مجموعه داده برخط در دنیای واقعی معمولاً بزرگ، ناقص، بدون ساختار، بدون برچسب و دارای عوامل مزاحم است. ثانیاً هر روز مقدار زیادی از اطلاعات نادرست با اهداف متنوع و ویژگی‌های زبانی مختلف از طریق رسانه‌های اجتماعی ایجاد می‌شود که در نتیجه به دست آوردن برچسب درستی مستدل برای داده‌ها را دشوار می‌کند. با توجه به محدودیت‌های مدل یادگیری نظارت‌شده که در بخش قبل به آن اشاره شد، مدل یادگیری بدون نظارت برای حل مسئله در دنیای واقعی کاربردی‌تر و عملی‌تر است. با این حال، تنها مطالعات کمی وجود دارد که به‌طور مستقیم بر روی کشف اخبار جعلی برخط از طریق رویکردهای بدون نظارت کار کند. بیشتر آنها بر تحلیل شباهت معنایی یا عقیده‌کاوی متمرکز شده‌اند و اندازه‌گیری تشابه بدون نظارت را برای تشخیص نظرهای

-
1. Decision Tree
 2. Random Forest
 3. Support Vector Machine (SVM)
 4. Logistic Regression
 5. K-nearest Neighbor

کاربران جعلی به کار برده‌اند (Ahmed, 2017). با ترکیبی از شباهت کلمه‌ای و شباهت ترتیب کلمات، رویکرد پیشنهادی در پژوهش دیگری می‌تواند نظرهای تقریباً تکراری کاربران را با دقت بالا شناسایی کند. ونگ و همکارانش یک چهارچوب عقیده‌کاوی بدون نظارت برای تصاویر در رسانه‌های اجتماعی ارائه داده‌اند (Wang & Li, 2015, pp.1584-1591). با بهره‌گیری از روابط بین اطلاعات بصری و اطلاعات متنی مرتبط، روش آنها می‌تواند عقیده مستتر در تصاویر اجتماعی را از دو مجموعه داده بزرگ پیش‌بینی کند. این روش بر اساس اندازه‌گیری شباهت کسینوسی، و با استفاده از امتیازات و ویژگی‌های زمانی، به‌طور خودکار ویژگی‌های تفکیک‌کننده راویان راستگو و متقلب را تشخیص می‌دهند. ژانگ و همکارانش معتقدند که الگوریتم‌های یادگیری بدون نظارت مسیرهای عملی و ضروری برای کشف اخبار جعلی برخط هستند و باید در تحقیقات آینده اولویت اصلی قرار گیرند (Zhang & Ghorbani, 2020, pp.102025). آنها انواع مدل یادگیری بدون نظارت را برای تشخیص اخبار جعلی پیشنهاد می‌کنند که عبارت‌اند از: خوشه‌بندی^۱، تحلیل داده دورافتاده^۲، تحلیل تشابه معنایی^۳، و نهفتگی اخبار بدون نظارت^۴.

۵. اخبار جعلی از منظر حقوقی

اخبار جعلی (Fake News) به‌عنوان یکی از چالش‌های اصلی در عصر ارتباطات نوین، مسائل پیچیده‌ای را در حوزه حقوقی ایجاد کرده است. از منظر حقوقی، اخبار جعلی می‌تواند مصداق نقض حقوق اساسی شهروندان از جمله حق دسترسی به اطلاعات صحیح و دقیق باشد (Smith, 2020). در برخی کشورها، انتشار اخبار جعلی می‌تواند تحت قوانین مربوط به افترا، تشویش اذهان عمومی، یا حتی جرائم امنیت ملی مورد پیگرد قرار گیرد (Brown & Jones, 2019). از منظر حقوقی، انتشار اخبار جعلی به‌طور مستقیم با مفاهیم مربوط به جرم فریب عمومی، سوءاستفاده از اطلاعات و نقض اعتماد عمومی مرتبط است. سازندگان اخبار جعلی، چه انسان باشند و چه ماشین (ربات‌های اجتماعی یا سایبورگ‌ها)، در صورت

-
1. Cluster analysis
 2. Outlier analysis
 3. Semantic similarity analysis
 4. Unsupervised news embedding

انتشار اطلاعات کذب و گمراه‌کننده، می‌توانند تحت پیگرد قانونی قرار گیرند. سازندگان اخبار جعلی برخط می‌توانند انسان یا ماشین باشند. سازندگان اخبار جعلی هم شامل نویسندگان خبر می‌شود و هم کاربرانی که اخبار جعلی را ناخواسته منتشر می‌کنند. مصرف‌کنندگان خبر هم می‌توانند دانش‌آموزان، رأی‌دهندگان، والدین، افراد سالخورده و... باشند. ماشین‌ها، ربات‌های اجتماعی و سایبورگ‌ها رایج‌ترین سازندگان اخبار جعلی ماشینی هستند.

سازندگان اخبار جعلی با گمراه کردن مصرف‌کنندگان، اعتماد اجتماعی را تحت تأثیر قرار می‌دهند که این خود نوعی نقض اعتماد عمومی محسوب می‌شود. این موضوع به‌ویژه در زمینه‌های سیاسی و اقتصادی می‌تواند اثرهای بسیار منفی داشته باشد و ممکن است به‌عنوان جرمی علیه نظم عمومی و مردم‌سالاری در نظر گرفته شود. با این حال، تعیین حدود قانونی برای مقابله با اخبار جعلی، همواره با چالش‌هایی همراه بوده است؛ زیرا هرگونه اقدام قانونی ممکن است با آزادی بیان به‌عنوان یکی از حقوق بنیادین بشر تضاد پیدا کند (Johnson, 2018). سازمان ملل متحد نیز تأکید کرده است که تلاش برای محدود کردن اخبار جعلی باید در چهارچوب استانداردهای بین‌المللی حقوق بشر، از جمله اصل تناسب و ضرورت، انجام شود (United Nations, 2017).

علاوه بر این، بسیاری از کشورها به‌منظور مقابله با انتشار اخبار جعلی، قوانینی نظیر «قانون مسئولیت سکوها» را تصویب کرده‌اند که بر اساس آن، سکوهای دیجیتال موظف به نظارت بر محتوای منتشرشده توسط کاربران خود هستند (Taylor, 2021). این قوانین، گرچه می‌توانند تأثیر مثبتی بر کاهش انتشار اطلاعات نادرست داشته باشند؛ اما در برخی موارد ممکن است به سانسور غیرضروری منجر شوند (Williams, 2022)؛ بنابراین، تلاش‌های گسترده‌ای برای تجزیه و تحلیل حساب‌های مخرب در رسانه‌های اجتماعی صورت گرفته است. حساب‌های مخرب، رفتاری متفاوت از کاربران قانونی دارند؛ بنابراین در تحلیل سازنده و مصرف‌کننده خبر با توجه به رخ‌نمای کاربران، رفتار پستی و زمانی آنها و نیز اعتبار حساب می‌توان به اطلاعات ارزشمندی از آنان دست یافت. طبق قوانین بسیاری از کشورها، هم سازندگان و هم توزیع‌کنندگان اخبار جعلی ممکن

است مسئولیت کیفری داشته باشند. حتی اگر فردی به صورت ناخواسته و بدون اطلاع از جعلی بودن اطلاعات، آن را منتشر کند، می‌تواند تحت برخی قوانین مربوط به انتشار غیرقانونی اطلاعات مسئول شناخته شود. اطلاعات اولیه رخنمای کاربر شامل زبان حساب کاربری، مکان‌های جغرافیایی حساب کاربری، زمان ایجاد حساب کاربری، تعداد پست / توئیت، رفتار پستی و زمانی یک حساب، میانگین زمان بین دو پست متوالی، فراوانی پاسخ، اشتراک‌گذاری، و اشاره کردن^۱ مؤلفه‌هایی هستند که به تشخیص صحت خبر منتشر شده توسط کاربر کمک می‌کنند. به کمک زمان‌سنج‌ها و برنامه‌های خودکار، دیده شده که حساب‌های مشکوکی مانند ربات‌های اجتماعی و سایبورگ‌ها در یک بازه زمانی خاص فعال هستند. در مقابل کاربران انسانی قانونی از نظر زمانی رفتارهای پیچیده‌ای دارند. شدت بالای پاسخ‌ها و رفتارهای ذکر شده ممکن است نشانگر سطح بالای مشکوک بودن یک حساب اجتماعی باشد. همچنین می‌توان اعتبار حساب را با توجه به تعداد دوستان و دنبال‌کنندگان^۲ بررسی کرد. تعداد دنبال‌کنندگان یک کاربر قانونی شبکه‌های اجتماعی اغلب به دوستان خود نزدیک است. با این حال، ربات‌های اجتماعی معمولاً دوستان بیشتری نسبت به دنبال‌کنندگان دارند. چو و همکارانش معادله‌ای برای محاسبه اعتبار حساب با تعداد دنبال‌کنندگان و دوستان پیشنهاد می‌کنند. فرمول (۱) را ببینید:

$$Account_{Reputation} = \frac{follower}{follower+friend} \quad (1)$$

آن‌ها مشاهده کردند که یک فرد مشهور همیشه دارای امتیاز شهرت نزدیک به ۱ است، درحالی‌که این امتیاز برای یک ربات اجتماعی مشکوک نزدیک به صفر است (Horne & Adali, 2017, pp.759-766).

از منظر حقوقی، اگر کاربرانی که با سازنده خبر در تعامل هستند (چه با «لایک» یا اظهارنظر) به انتشار اخبار نادرست کمک کنند، می‌توانند مسئولیت مدنی یا حتی کیفری داشته باشند. در برخی کشورها، بازنشر یا تأیید اخبار نادرست، به‌ویژه در زمینه‌های حساس مانند مسائل سیاسی یا امنیتی، می‌تواند جرم تلقی شود. این به‌معنای افزایش مسئولیت فردی کاربران برای بررسی صحت اخبار قبل از تعامل با آن‌هاست. صحت‌بخشی

1. Mentioning
2. Followers

از اخبار برخط را می‌توان در شبکه سازنده اخبار شناسایی کرد. کاربران مختلف برخط، سوابق تحصیلی، تجربیات کاری و علایق مختلفی دارند؛ بنابراین کاربران در رسانه‌های اجتماعی تمایل دارند گروه‌هایی را تشکیل دهند که شامل کاربران همفکر باشد؛ بنابراین این فرضیه منطقی است که از کاربران برخطی که با سازنده خبر، در تعامل بسیار هستند، می‌توان برای پیش‌بینی سطح صحت اخبار استفاده کرد. در واقع تجزیه و تحلیل شبکه کاربر، مطالعه‌ای برای آشکار کردن تعامل بین کاربران برخط است؛ به‌عنوان مثال، اگر بسیاری از حساب‌های غیرعادی یا غیرقابل اعتماد، یک قطعه از خبر را "like" کرده باشند یا در مورد آن اظهارنظر کرده باشند، این خبر به احتمال زیاد حاوی اطلاعات غلط و گمراه‌کننده است.

از منظر حقوقی، مقابله با اخبار جعلی نیازمند توازن دقیق میان حفاظت از حق دسترسی آزاد به اطلاعات و حمایت از آزادی بیان است. این موضوع همچنان یکی از دغدغه‌های اصلی محافل حقوقی و سیاست‌گذاری است و نیاز به تدوین قوانین جامع‌تری دارد که به‌طور همزمان حقوق اساسی شهروندان را حفظ کنند و از انتشار اطلاعات نادرست جلوگیری کنند (Anderson, 2019)؛ بنابراین بررسی ویژگی‌های اخبار جعلی می‌تواند گامی مؤثر در این راستا باشد.

۶. ویژگی‌های اخبار جعلی

همان‌طور که اشاره شد در رویکردهای مبتنی بر یادگیری نظارت شده مجموعه‌ای از ویژگی‌های اخبار مورد تجزیه و تحلیل قرار می‌گیرند. در ادامه ویژگی‌های مهم و رایج مورد استفاده برای کشف اخبار جعلی برخط را مورد بحث قرار می‌دهیم. بر اساس مؤلفه‌های اخبار جعلی که در جدول ۱ مورد بحث قرار گرفته است، سه نوع اصلی از مجموعه ویژگی‌ها وجود دارد: ویژگی‌های مبتنی بر سازنده/کاربر، ویژگی‌های مبتنی بر محتوای خبر، ویژگی‌های مبتنی بر بستر اجتماعی.

جدول (۱) دسته‌بندی ویژگی‌های اخبار جعلی

ویژگی‌های اخبار جعلی								
ویژگی‌های مبتنی بر سازنده/کاربر			ویژگی‌های مبتنی بر محتوای اخبار			ویژگی‌های مبتنی بر بستر اجتماعی		
ویژگی‌های مبتنی بر زمان	ویژگی‌های مبتنی بر شبکه	ویژگی‌های مبتنی بر توزیع	ویژگی‌های بصری	ویژگی‌های مبتنی بر سبک	ویژگی‌های زبانی و نحوی	ویژگی‌های رفتار کاربر	ویژگی‌های اعتبار کاربر	ویژگی‌های پروفایل کاربر

۶-۱. ویژگی‌های مبتنی بر سازنده/کاربر

ویژگی‌های مبتنی بر سازنده/کاربر به‌طور گسترده برای شناسایی حساب‌های برخط مشکوک مورد استفاده قرار می‌گیرند. هدف از این ویژگی‌ها به‌دست آوردن ویژگی‌های منحصر به فرد حساب‌های مشکوک یا حساب‌های ماشینی است. از این دسته ویژگی‌ها می‌توان به موارد زیر اشاره کرد. مشخصات تارنمای کاربر شامل اطلاعات اصلی کاربر مانند نام حساب، اطلاعات جغرافیایی، داده‌های ثبت‌نام کاربر، تأیید شده یا خیر، دارای توضیحی است یا خیر. مؤلفه اعتبار کاربر شامل امتیاز اعتبار کاربر، تعداد دوستان و دنبال‌کنندگان کاربر، نسبت بین دوستان کاربر و دنبال‌کنندگان، تعداد کل توییت‌ها/پست‌های کاربر و همچنین ویژگی‌های رفتار کاربر که به منظور دستیابی به الگوی رفتار کاربر هم برای کاربران فریبکار و هم کاربران مشروع مورد استفاده قرار می‌گیرند. یک ویژگی معمولی رفتار کاربر، امتیاز ناهنجاری کاربر است که با استفاده از تعداد تعامل کاربر در یک پنجره زمانی، تقسیم بر میانگین تعامل ماهانه کاربر، برای تشخیص اطلاعات غیرعادی به‌کار برده می‌شود (Zhao et al., 2013, pp.1391-1400).

۶-۲. ویژگی‌های مبتنی بر محتوای اخبار

ویژگی‌های مبتنی بر محتوای اخبار از سر نخ‌های آشکار برای کشف اخبار جعلی است. این مشخصه‌ها را می‌توان به ویژگی‌های زبانی و نحوی، ویژگی‌های مبتنی بر سبک نوشتن و ویژگی‌های بصری طبقه‌بندی کرد.

ویژگی‌های زبانی و نحوی به مؤلفه اساسی، ساختاری و معناشناسی زبان طبیعی اشاره دارد. اگرچه محتوای خبری جعلی همیشه به عمد برای گمراه کردن کاربران برخط تولید می‌شود؛ اما ویژگی‌های مبتنی بر زبان و نحو هنوز منابع ارزشمندی برای تجزیه و تحلیل اخبار مشکوک هستند و آنها را می‌توان از این نظر به سه سطح کلمه، جمله و محتوا طبقه‌بندی کرد:

۶-۲-۱. سطح کلمه

بسته کلمات^۱، n-gram، فراوانی کلمه، فروانی وزنی IDF^۲ رایج‌ترین ویژگی‌های زبانی برای پردازش زبان طبیعی هستند. به علاوه وجود نشانه‌های خاص و مشکوک در محتوای خبر می‌تواند برای شناسایی اطلاعات جعلی استفاده شود (Pennebaker et al., 2003, pp.547-577). علائم خاص و مشکوک شامل علامت تعجب، علامت سؤال، چندین علامت تعجب، اشاره به کاربر، هشتگ، علامت لبخند، علامت اخم، حروف بزرگ، کلمات برجسته، می‌باشد. مشابه علائم مشکوک، وجود کلمات مربوط به سبک همچنین می‌تواند برای کشف اخبار جعلی برخط استفاده شود. کلمات مربوط به سبک عبارت‌اند از: کلمات توقف، علائم نگارشی، نقل‌قول، کلمات منفی (مانند نه، هرگز، نه، علی‌رغم، تردید، جعل، تخریب، انکار، تقلب و نادرست و...)، کلمات محاوره‌ای/سوگند، عبارات پرسشی چرا، چه، چه زمان، چه کسی، اسامی، ضمایر شخصی، ضمایر ملکی، قیود، حروف ندا، افعال، کلمات کمی‌ساز، کلمات مقایسه‌ای، علامت‌های تعجب، عبارات عامیانه.

LIWC^۳ یک برنامه تحلیل متن است که کلمات را در مقوله‌های معنادار از لحاظ

-
1. Bag-of-words
 2. Term frequency-inverted document frequency (TF-IDF)
 3. Linguistic Inquiry and Word Count

روان‌شناختی شمارش می‌کند. LIWC شامل پنج دسته اصلی و زیرشاخه‌های مختلف مانند اجتماعی، عاطفی، شناختی، ادراکی است. بر اساس LIWC، محققان می‌توانند تعداد کلمات احساسی در محتوای اخبار را حساب کنند که ممکن است به تعیین سطح کلی احساسات موجود در خبر کمک کند. کلمات احساسی شامل کلمات تحلیلی، کلمات حسی، کلمات بیان‌کننده علت، کلمات عدم توافق، کلمات موقت، کلمات یقین، واژه‌های متمایز، کلمات وابستگی، کلمات قدرت، کلمات بارز، کلمات خطرناک، کلمات دغدغه‌های شخصی (شغل، اوقات فراغت، دین، سرمایه)، کلمات با لحن احساسی، کلمات هیجانی (عصبانیت، غم) و... هستند. سایر ویژگی‌های زبانی کلمات، مانند خوانایی کلمات (امتیاز خواندن بر اساس تعداد هجا در کلمات) و نسبت نوع-نشانه (تعداد کلمات منحصربه‌فرد تقسیم بر تعداد کل کلمات در اسناد/پست‌ها/توییت‌ها) می‌تواند سادگی و تنوع واژگان در اخبار را نشان دهد (Castillo et al., 2011, pp.675-684) و همچنین می‌تواند برای تجزیه و تحلیل محتوای جعلی اخبار استفاده شود.

۶-۲-۲. سطح جمله

ویژگی‌های جمله به تمام ویژگی‌های مهم در مقیاس جمله اشاره دارد و شامل برچسب‌گذاری پاره‌های گفتار^۱، طول متوسط جمله، میانگین طول یک توییت/پست، فراوانی علائم نگارشی، کلمات دستوری، عبارت در یک جمله، تقارن متوسط جمله (مثبت، خنثی یا منفی)، پیچیدگی جمله و... . به‌طور خاص محققان پیچیدگی جمله، عمق درخت نحوی هر جمله، عمق درخت نحوی عبارات اسمی و عمق درخت نحوی عبارات فعلی را با استفاده از Stanford parser محاسبه می‌کنند که برای کشف اخبار جعلی برخط استفاده شده است (Horne & Adali, 2017, pp.759-766).

۶-۲-۳. سطح محتوا

ویژگی‌های محتوا به اطلاعات خام محتوای متا اخبار اشاره دارد. امتیاز کلی احساسات محتوای خبری به‌عنوان ابزاری قدرتمند برای اخبار جعلی و تحلیل نویسنده مشکوک تأیید شده است. در مطالعات قبلی، از SentiStrength برای شناسایی شدت احساسات مثبت و منفی

در سند خام استفاده شده است. به‌طور مشابه، ویژگی‌های موضع‌محور^۱ رفتار حمایتی یا انکاری یک تویییت/پست را نشان می‌دهند و می‌توانند برای ارزیابی وضعیت احساسی اخبار استفاده شود. سایر ویژگی‌ها در سطح محتوای خبر، از قبیل عنوان خبری پررنگ، موضوعات خبری (زندگی اجتماعی، سیاست، فناوری و امنیت سایبری، تجاری و مالی و...)، قطعیت خبر، تعداد نشانه‌های خاص یا نمادها در کل اخبار، وجود پیوندهای (لینک‌ها) خارجی یا URLها نیز سرخ‌های مهمی برای شناسایی برخط اخبار جعلی هستند.

همان‌طور که اشاره شد علاوه بر ویژگی‌های زبانی و نحوی خبر، مؤلفه‌های مبتنی بر سبک نیز می‌توانند برای بررسی محتوای خبر مورد استفاده قرار گیرند. این مؤلفه‌ها جهت نشان دادن ویژگی‌های مختلف سبک نوشتن نویسندگان اخبار جعلی به‌کار می‌روند. اگرچه بیشتر اوقات، نویسندگان اخبار جعلی سعی در تقلید از شیوه نوشتن یک نویسنده اخبار عادی دارند تا خوانندگان برخط را فریب دهند؛ اما هنوز هم تفاوت‌هایی وجود دارد که می‌تواند باعث ایجاد تمایز بین سازندگان اخبار جعلی و اخبار واقعی شود. به‌منظور شناسایی نظرهای جعلی، در پژوهشی ویژگی‌های استفاده از کلیدها^۲ برای هر دوی محتواسازان جعلی و نظردهندگان واقعی برخط مطالعه شده است. با تمرکز بر ویژگی‌های الگوی ویرایش (مانند تعداد پاک‌کردن‌ها، کلیک موس و استفاده از arrow keyها) و ویژگی‌های بازه زمانی (مانند مدت زمان کل گزارش، فاصله متوسط بین کلمات) محققان دریافتند که سازندگان محتوای جعلی برای اتمام نوشتن به زمان بیشتری نیاز دارند و اشتباهات بیشتری انجام می‌دهند (Ahmed, 2017).

علاوه بر متون موجود در خبر، تصاویر یا فیلم‌های موجود در یک محتوای خبری نیز نشانه‌های اساسی برای شناسایی اطلاعات مشکوک یا فریبنده هستند. مطالعات اخیر ویژگی‌های بصری را برای شناسایی اطلاعات غلط برخط بررسی می‌کنند. این ویژگی‌های بصری شامل تعداد تصاویر یا فیلم‌ها، امتیاز وضوح، امتیاز انسجام، نمودار توزیع شباهت، امتیاز تنوع، امتیاز خوشه‌بندی، نسبت تصویر، نسبت چند تصویر، نسبت تصویر داغ، نسبت تصویر طولانی، و... هستند.

1. Stance-based
2. Keystroke

۳-۶. ویژگی‌های مبتنی بر بستر اجتماعی

ویژگی‌های مبتنی بر بستر اجتماعی برای انعکاس الگوی توزیع اخبار برخط و تعامل بین کاربران برخط طراحی شده‌اند و می‌توان آنها را در سه نوع زیر خلاصه کرد: ویژگی‌های مبتنی بر شبکه، ویژگی‌های مبتنی بر توزیع و ویژگی‌های مبتنی بر زمان. تجزیه و تحلیل مبتنی بر شبکه بر روی گروهی از کاربران برخط که از جنبه‌های مختلف، از قبیل مکان، پیشینه تحصیلات و عادات، مشابه هستند تمرکز دارد. ویژگی‌های مبتنی بر شبکه بر اساس شبکه‌های خاص، انتخاب و استخراج می‌شوند و می‌توانند برای مطالعه ویژگی‌های منحصر به فرد شبکه‌های خاص و شباهت و عدم تشابه حساب‌های برخط مختلف استفاده شوند. شو و همکارانش یک جمع‌بندی مختصر از تجزیه و تحلیل مبتنی بر شبکه‌های مختلف، مانند شبکه وضعیت، شبکه همزمان، شبکه دوستی و شبکه انتشار ارائه می‌دهد (Shu et al., 2019, pp.43-65).

ویژگی‌های مبتنی بر توزیع می‌توانند به کشف الگوی متمایز انتشار اخبار برخط کمک کنند. معمولاً می‌توان یک درخت تکثیر را برای سهولت در توصیف ماهیت توزیع یک خبر ساخت. ویژگی‌های مربوط به درخت تکثیر شامل درجه ریشه در یک درخت تکثیر، حداکثر تعداد زیر درخت، حداکثر/متوسط درجه و عمق درخت و غیره می‌باشد. علاوه بر این، برخی از ویژگی‌های دیگر مانند تعداد بازتوییت‌ها/بازنشرها برای توییت/پست اصلی و بخشی از توییت‌ها/پست‌هایی که برای یک حساب برخط مجدد توییت می‌شوند، برای بررسی تأثیر، محبوبیت و سطح مشکوک بودن اخبار جعلی برخط به کار برده می‌شوند.

از ویژگی‌های مبتنی بر زمان نیز می‌توان برای توصیف رفتار ارسال اخبار برخط به صورت سری زمانی استفاده کرد. این ویژگی‌ها برای شناسایی فعالیت‌های ارسال مشکوک مناسب هستند و می‌توانند برای نشان دادن سطح نادرستی اخبار برخط استفاده شوند. از ویژگی‌های متداول زمانی که معمولاً استفاده می‌شود، عبارت‌اند از فاصله بین دو پست، دفعات ارسال، پاسخ دادن و نظر دادن برای یک حساب خاص، زمان و روزی که اطلاعات اصلی ارسال می‌شود/به اشتراک گذاشته می‌شود یا نظر داده می‌شود.

در حال حاضر اکثر مطالعات سعی در ارزیابی درست یا نادرست بودن اطلاعات برخط دارند؛ اما شناسایی هر گونه اخبار دروغ پرتطرفدار یا بالقوه در اسرع وقت بسیار مهم است. به همین دلیل در کنار سیستم‌های تشخیص اخبار جعلی، پیش‌بینی زود هنگام خبر جعلی و مداخله در انتشار اخبار جعلی نیز مورد توجه محققان قرار گرفته است. پیش‌بینی زود هنگام اخبار جعلی با یادگیری از داده‌های گذشته، تلاش می‌کند که اخبار جعلی یا شایعات نوظهور را حتی قبل از وقوع آنها کشف کند. برخی پژوهشگران معتقدند که به کمک اطلاعات مربوط به قطبش^۱ و تأیید کاربر، قادر به شناسایی موضوعاتی هستند که مستعد سوءاطلاعات هستند (Vicario et al., 2019, pp.1-22). برخی پژوهشگران نیز مسئله تشخیص شایعات زود هنگام در رسانه‌های اجتماعی را بررسی کرده‌اند (Wu et al., 2017, pp.99-107).

پیش‌بینی زود هنگام اخبار جعلی می‌تواند قبل از وجود اخبار جعلی، اطلاعات نادرست بالقوه را به کاربران برخط یادآوری کند. مداخله در انتشار اخبار جعلی می‌تواند به کاربران برخط کمک کند تا تأثیرات منفی اخبار جعلی را پس از انتشار اخبار جعلی پاک کنند. فرج‌تبار و همکارانش معتقدند که ترکیب یادگیری تقویتی^۲ با مدل شبکه فرایند نقطه^۳ (Farajtabar et al., 2017, pp.1097-1106)، تأثیر اخبار جعلی در رسانه‌های اجتماعی را کاهش می‌دهند. حال وظیفه دولت‌هاست تا با به‌کارگیری روش‌های پیشنهاد شده فضای تشخیص خودکار اخبار جعلی را برای کاربران ارائه دهند تا بتوانند بستر حق دسترسی آزاد به اطلاعات و حق دسترسی اطلاعات صحیح را در جامعه ایجاد کنند.

بحث و نتیجه‌گیری

حق دسترسی آزاد به اطلاعات یکی از اصول اساسی حقوق بشر و حقی است که وظیفه دولت‌ها در بسترسازی این حق است. اخبار جعلی یکی از تهدیدکننده‌ترین آسیب‌ها در شبکه‌های اجتماعی است. اخبار جعلی می‌توانند توسط افراد سو دجو به‌منظور تأثیرگذاری بر تصمیمات کلیدی روزمره افراد، از جمله در زمینه‌های بازار سهام، انتخاب‌های بهداشتی،

-
1. Polarization
 2. Reinforcement learning
 3. Point process network

خریدهای برخط، تحصیلات و حتی انتخابات ریاست جمهوری، مورد سوءاستفاده قرار گیرند؛ از این رو، اهمیت شناسایی و پالایش این اخبار در راستای حق دسترسی آزاد به اطلاعات بسیار حیاتی است. کشف خودکار اخبار جعلی برخط یک کار بسیار مهم اما چالش برانگیز است. این پژوهش روشن می‌سازد که تشخیص صحت اطلاعات در بستر دیجیتال، نیازمند همکاری و هم‌افزایی میان انسان و فناوری است. از آنجاکه حق دسترسی آزاد به اطلاعات یکی از حقوق اساسی بشر به‌شمار می‌رود، لازم است شیوه‌های مؤثری برای شناسایی و مقابله با اخبار جعلی اتخاذ گردد. در نهایت، با بهره‌گیری از روش‌های شناسایی اخبار جعلی و ارتقای سطح آگاهی شهروندان درباره اهمیت صحت اطلاعات، می‌توان به تحقق حق دسترسی آنها به اطلاعات صحیح و معتبر کمک شایانی کرد. این امر نه تنها به بهبود فضای ارتباطی در جامعه کمک می‌کند، بلکه می‌تواند به تقویت مردم‌سالاری و افزایش مشارکت اجتماعی نیز بینجامد.

در این پژوهش در مورد درک عمیق جنبه‌های مهم اخبار جعلی برخط، مانند نشردهنده/ سازنده خبر جعلی، گروه‌های هدف، محتوای خبر و بستر اجتماعی بحث کردیم. توصیف واضح اخبار جعلی برخط می‌تواند نقش به‌سزایی در تجزیه و تحلیل داده‌های ارتباطات اجتماعی و تشخیص اطلاعات غیرعادی داشته باشد. همچنین با مقایسه رویکردهای تشخیص موجود، فهرست کامل مجموعه‌ای از ویژگی‌های اخبار جعلی برای صحت‌سنجی اخبار ارائه دادیم.

حق دسترسی آزاد به اطلاعات نه تنها یک اصل اساسی در ایجاد جوامع مردم‌سالارانه و شفاف است بلکه ابزاری حیاتی برای توانمندسازی شهروندان و ارتقاء آگاهی عمومی می‌باشد.

این پژوهش با مراجعه به مصوبه شورای عالی فضای مجازی در زمینه مقابله با اخبار خلاف واقع، به بررسی و نقد رویکردها و روش‌های پیش‌بینی‌شده در این سند می‌پردازد. همچنین، راهکارهای ترکیبی انسان‌محور و فناوری‌محور را برای تشخیص اخبار جعلی پیشنهاد می‌دهد. یافته‌ها نشان می‌دهد که ایجاد تعادل میان مقابله با اطلاعات نادرست و

حفظ حق دسترسی آزاد به اطلاعات، نیازمند اصلاح برخی سیاست‌های فعلی و بهره‌گیری از راهکارهای هوشمند است.

برای تحقق این حق، اقدامات مشترک از سوی دولت‌ها، رسانه‌ها و نهادهای مدنی ضروری است. دولت‌ها باید با ایجاد زیرساخت‌های فناوری مناسب و حذف موانع قانونی، زمینه را برای دسترسی آسان به اطلاعات فراهم کنند. رسانه‌ها نیز باید متعهد به ارائه اطلاعات دقیق و قابل اعتماد باشند. در پایان، آگاهی و آموزش درباره شناسایی اخبار و اطلاعات معتبر، ضرورتی حیاتی برای مقابله با چالش‌های اطلاعاتی فعلی است. با این رویکرد، می‌توان به ایجاد جامعه‌ای هوشمندتر، آگاه‌تر و مشارکت‌جو دست یافت.

منابع

- اسلامی، جهانگیر و پیشقدم، کاظم، م. (۱۴۰۰). بررسی تطبیقی مفهوم آزادی در منشور حقوق شهروندی و اعلامیه جهانی حقوق بشر. *پژوهش‌های حقوقی*، ۲۰(۴۵)، ۹۵-۱۲۳.
- سالنامه آماری قوه قضاییه. (۱۴۰۲). آمار پرونده‌های حقوق شهروندی. تهران: مرکز آمار و فناوری اطلاعات قوه قضاییه.
- رهایی، ملکی، مسلم. (۱۳۹۸). آزادی اطلاعات و اخبار و محدودیت اخلاق عمومی در حقوق بین‌الملل و اسلام. *پژوهش تطبیقی حقوق اسلام و غرب*، ۶(۱)، ۲۳-۵۶.
- شورای عالی فضای مجازی. (۱۴۰۲). مصوبه الزامات پیشگیری و مقابله با نشر اطلاعات، اخبار و محتوای خبری خلاف واقع در فضای مجازی. تهران، ایران: مرکز ملی فضای مجازی.
- محسنی، هاشمی، جاوید، عباسی. (۱۳۹۸). تحلیل حقوقی نسبت‌سنجی حق دسترسی عموم به اطلاعات با تحقق حقوق شهروندی با تأکید بر نظام حقوقی ایران. *پژوهش حقوق عمومی*، ۲۰(۶۲)، ۳۲۱-۳۵۴.
- مرکز پژوهش‌های مجلس شورای اسلامی. (۱۴۰۱). گزارش عملکرد قانون انتشار و دسترسی آزاد به اطلاعات. تهران: انتشارات مجلس.
- معاونت حقوقی رئیس‌جمهور. (۱۴۰۲). آمار عملکرد کمیسیون ماده ۱۸ قانون انتشار و دسترسی آزاد به اطلاعات.

- مرکز مطبوعات و انتشارات قوه قضائیه. (۱۴۰۲). چالش‌های اجرای قانون انتشارات و مطبوعات. تهران: معاونت اجتماعی قوه قضائیه.
- نمک‌دوست تهرانی. (۱۳۸۴). آزادی اطلاعات و حق دسترسی: بنیان دموکراسی. مجلس و راهبرد، ۴۲(۱)، ۶۳-۱۰۸.
- وزیری، آباد، م. ب.، جبار. (۱۳۹۰). اصول حاکم بر گردش آزاد اطلاعات در اسلام و حقوق موضوعه. فصلنامه علمی حقوق و فقه تربیت و کودک، ۲(۷)، ۱۲۶-۱۰۷.
- Afroz, S., Brennan, M., & Greenstadt, R. (2012). Detecting hoaxes, frauds, and deception in writing style online. IEEE symposium on security and privacy.
- Ahmed, H. (2017). *Detecting opinion spam and fake news using n-gram analysis and semantic similarity*.
- Anderson, M. (2019). Legal Challenges of Fake News: Balancing Freedom of Expression and Public Order. *International Journal of Law and Communication*, 12(3), 45-63.
- Bigne, E., Andreu, L., Hernandez, B., & Ruiz, C. (2018). The impact of social media and offline influences on consumer behaviour. An analysis of the low-cost airline industry. *Current Issues in Tourism*, 21(9), 1014-1032 .
- Brown, P., & Jones, R. (2019). *Fake News and Its Legal Implications: A Comparative Analysis*. Cambridge University Press.
- Castillo, C., Mendoza, M., & Poblete, B. (2011). *Information credibility on twitter. Proceedings of the 20th international conference on World wide web*.
- Di, R., Wang, H., Fang, Y., & Zhou, Y. (2018). *Fake comment detection based on time series and density peaks clustering Algorithms and Architectures for Parallel Processing: ICA3PP 2018 International Workshops*, Guangzhou, China, November 15-17, 2018, Proceedings 18
- Farajtabar, M., Yang, J., Ye, X., Xu, H., Trivedi, R., Khalil, E., Li, S., Song, L., & Zha, H. (2017) *Fake news mitigation via point process based intervention. International conference on machine learning*.
- Horne, B., & Adali, S. (2017). *This just in: Fake news packs a lot in title, uses simpler, repetitive content in text body, more similar to satire than real news*. Proceedings of the international AAAI conference on web and social media,
- Johnson, L. (2018). Freedom of Speech and the Threat of Fake News. *Oxford Legal Studies*, 18(2), 78-91.
- Lewandowsky, S., et al. (2020). *The Debunking Handbook 2020*. George Mason University.
- Oshikawa, R., Qian, J., & Wang, W. Y. (2018). *A survey on natural language processing for fake news detection*. arXiv preprint arXiv:1811.00770 .
- Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual review of psychology*, 54(1), 547-577 .
- Shao, C., Ciampaglia, G. L., Flammini, A., & Menczer, F. (2016). *Hoaxy: A platform for tracking online misinformation*. Proceedings of the 25th international conference

companion on world wide web.

- Shu, K., Bernard, H. R., & Liu, H. (2019). *Studying fake news via network analysis: detection and mitigation*. Emerging research challenges and opportunities in computational social network analysis and mining, 43-65 .
- Smith, J. (2020). The Rise of Fake News in the Digital Era: Legal Perspectives. *Harvard Journal of Law & Technology*, 33(1), 15-28.
- Taylor, K. (2021). Platform Responsibility and Misinformation: Legal Frameworks in Practice. *Journal of Internet Regulation*, 9(4), 92-104.
- United Nations. (2017). Human Rights-Based Approach to Misinformation and Disinformation. UN Human Rights Council.
- Vicario, M. D., Quattrociocchi, W., Scala, A., & Zollo, F. (2019). Polarization and fake news: Early warning of potential misinformation targets. *ACM Transactions on the Web (TWEB)* 22-1, (2) 13.
- Wang, Y., & Li, B. (2015). Sentiment analysis for social media images. 2015 IEEE international conference on data mining workshop (ICDMW).
- Williams, T. (2022). Censorship or Protection? The Double-Edged Sword of Fake News Laws. *Journal of Media Law*, 14(1), 33-50.
- Wu, L., Li, J., Hu, X., & Liu, H. (2017). Gleaning wisdom from the past: Early detection of emerging rumors in social media. Proceedings of the 2017 SIAM international conference on data mining.
- Yang, F., Liu, Y., Yu, X., & Yang, M. (2012). Automatic detection of rumor on sina weibo. Proceedings of the ACM SIGKDD workshop on mining data semantics.
- Zhang, X., & Ghorbani, A. A. (2020). An overview of online fake news: Characterization, detection, and discussion. *Information Processing & Management*, 57(2), 102025.
- Zhao, Z., Lee, W. C., Shin, Y., & Song, K.-B. (2013). An optimal power scheduling method for demand response in home energy management system. *IEEE transactions on smart grid*, 4(3), 1391-1400.