فصلنامه شناخت پژوهی مطالعات سیاسی
سال ۱، شماره ۲(۲۱)،۹۲-۷۰
تاریخ دریافت: ۱۴۰۳/۰۸/۰۱
تاریخ پذیرش: ۱۴۰۳/۰۹/۰۳

Journal of cognitive research in political studies
Vol.1, No2 (2), 70-92
Reseived Date: 2024/10/22
Accept Date: 2024/11/23

DOI: 10.20241403/CRPS.2408.1000.1.1.4

فنون مقابله و نبرد هوشمندانه با اخبار جعلی در عصر هوش‌مصنوعی

احمدرضا ترسلی[1]

شماره ۲ (۲)
سال ۱
فصل پاییز ۱۴۰۳
مقاله پژوهشی
تاریخ دریافت:
۱۴۰۳/۰۸/۰۱
تاریخ پذیرش:
۱۴۰۳/۰۹/۰۳
صص: ۹۲-۷۰

چکیده

تشخیص حقیقت از فریب نیازمند ابزارهای نوآورانه در عصری است که در محاصره اطلاعات نادرست است. اطلاعات نادرست مانند اخبار جعلی و شایعات، تهدیدی جدی برای زیست‌بوم‌های اطلاعاتی، زودباوری و اعتماد عمومی است و ظهور هوش‌مصنوعی پتانسیل زیادی برای تغییر شکل چشم‌انداز مبارزه با اطلاعات نادرست دارد. هدف اصلی این مقاله بررسی فنون مقابله و نبرد هوشمندانه با اخبار جعلی در عصر هوش‌مصنوعی است. پژوهش حاضر، از نظر هدف، کاربردی توسعه‌ای؛ از لحاظ رویکرد، کیفی و بر اساس روش فراترکیب انجام شده است. بر این اساس، جامعه آماری پژوهش، شامل مجموعه مقالاتی است که در بازه زمانی سال‌های ۲۰۲۰ تا ۲۰۲۴ میلادی و با موضوع هوش‌مصنوعی، اخبار جعلی، فنون مقابله و نبرد هوشمندانه به چاپ رسیده‌اند که با روش فراترکیب نهایتاً ۳۳ مقاله از میان ۶۳ مقاله انتخاب گردید. در این پژوهش بر اساس نتایج تحلیل مضمون، متون مرتبط مورد بررسی قرارگرفت و نتایج حاصل از روش تحقیق و تحلیل انجام شده در این پژوهش در قالب مدل مفهومی ترکیبی فنون مقابله و نبرد هوشمندانه با اخبار جعلی در عصر هوش‌مصنوعی ارائه گردید.

کلمات کلیدی: هوش‌مصنوعی، یادگیری ماشین، اخبار جعلی، اطلاعات نادرست، مقابله و نبرد هوشمندانه.

[1] دانشکده مدیریت راهبردی دانشگاه عالی دفاع ملی، تهران، ایران

استناد: ترسلی، احمدرضا. (۱۴۰۳). فنون مقابله و نبرد هوشمندانه با اخبار جعلی در عصر هوش‌مصنوعی. شناخت پژوهی مطالعات سیاسی، ۱(۲)، ۹۲-۷۰.

Tarassoli, A. (2024). Techniques to confront and fight smartly with fake news in the age of artificial intelligence. Cognitive research of political studies, 1(2), 70-92

مقدمه

با ظهور اینترنت، الگوهای ارتباطی مردم با افزایش سهم مشارکت در تولید محتوا تغییر کرده است. این امر زمینه انتشار اطلاعات نادرست را نیز تسهیل کرده است. رسانه‌های اجتماعی مانند فیس‌بوک، توییتر و اینستاگرام برای به اشتراک‌گذاری اطلاعات در زمان واقعی و دنبال کردن آخرین اخبار در مورد رویدادهای جاری استفاده می‌شوند (Tarassoli, 2024: 86). با این حال، با توجه به محبوبیت روزافزون کاربران شبکه‌های اجتماعی، این رسانه‌ها برای انتشار اخبار جعلی مناسب شده‌اند، زیرا هیچ سازوکار تأییدی برای اطلاعات به اشتراک گذاشته شده در این رسانه‌ها وجود ندارد. علاوه بر این، اخبار جعلی به دلیل ویژگی انتشار سریع اطلاعات در این رسانه‌ها بسیار سریع منتشر می‌شوند (Sahoo & Gupta, 2021:1). رهبر معظم انقلاب اسلامی نیز در دیدار بسیجیان به‌مناسبت روز بسیج شکست نقشه منطقه‌ای آمریکا به‌وسیله تفکر بسیجی و پرچم‌داری حاج قاسم سلیمانی فرمودند: «امروز مهم‌ترین شیوه‌ی دشمن، جعل و دروغ‌پردازی است. الان مهم‌ترین کاری که دشمن دارد می‌کند [انتشار] دروغ است؛ یعنی همین تلویزیون‌هایی که می‌دانید و می‌بینید مال دشمن است یا همین فضای مجازی، خبر دروغ می‌دهند، تحلیل دروغ می‌دهند[1]». بنابراین انتشار اخبار جعلی تهدیدات بالقوه امنیتی را برای جوامع مختلف از سازمان‌ها گرفته تا دولت در یک کشور ایجاد می‌کند. اخبار جعلی همیشه بخشی از برنامه‌ریزی‌های اطلاعاتی و نظامی بوده است. با این حال، مطمئنا در نتیجه استفاده از رسانه‌های اجتماعی و فناوری هوشمند بدتر نیز می‌شود. این به این دلیل است که فن‌آوری‌های ارتباطی مدرن ابزاری عموماً ارزان و کم مانع برای انتشار اطلاعات به ویژه انتشار اطلاعات نادرست ارائه می‌کنند.

به‌عنوان پیامد اخبار جعلی و نادرست، ممکن است باور عمومی نسبت به حقایق و وقایع در معرض تهدید قرار گیرد. بنابراین، نیاز مبرمی به مقابله و نبرد هوشمندانه با اطلاعات نادرست برای حفاظت از زیست‌بوم‌های اطلاعاتی و حفظ اعتماد عمومی، به ویژه در زمینه‌های پرمخاطره مانند اعتقادات، سلامت عمومی و ارزش‌های دینی و انقلابی وجود دارد.

شناسایی اخبار جعلی که توسط برخی مراجع به‌صورت دستی انجام می‌گردد همواره با مسائل و چالش‌های پیچیده و عدیده‌ای در عصر اطلاعات مواجه است که عبارت‌اند از: (۱)

---

[1] بیانات در دیدار بسیجیان، ۱۴۰۱/۰۹/۰۵

تشخیص تفاوت بین اطلاعات واقعی و نادرست، ردیابی و کنترل آن‌ها بسیار دشوار است، این مسئله، هسته اصلی مشکل تشخیص اخبار جعلی است. روش‌های سنتی و دستی، به دلیل پیچیدگی و حجم بالای اطلاعات، قادر به ارائه راه حل‌های کارآمد و مقیاس‌پذیر نیستند. (۲) با ظهور کلان داده‌ها، ارزیابی دستی کاری زمان‌بر است و افشای اطلاعات نادرست برای تأثیرگذاری بر مخاطبین خیلی دیر اتفاق می‌افتد. این امر باعث می‌شود که اخبار جعلی به سرعت منتشر شده و قبل از اینکه شناسایی شوند، تأثیر خود را بر افکار عمومی بگذارند. (۳) پائین بودن دقت بررسی واقعیت‌ها است که کاربران عموماً به آن اعتماد ندارند و این شرایط در متقاعد کردن افرادی که از قبل به اطلاعات نادرست اعتقاد دارند، بی‌اثر است. این امر باعث می‌شود که اخبار جعلی همچنان در میان مردم نفوذ کند. (۴) اخبار جعلی آینده که به طور مداوم در حال تغییر و تکامل هستند، ممکن است با الگوریتم‌های فعلی قابل تشخیص نباشند. (۵) انتشار سریع اخبار جعلی در شبکه‌های اجتماعی باعث می‌شود که مقابله با آن‌ها دشوارتر شود. (۶) محدودیت‌های رسانه‌های اجتماعی در مجوز دسترسی به اطلاعات که به دلایل مختلفی از جمله حفظ حریم خصوصی کاربران، جلوگیری از انتشار محتوای نامناسب و کنترل جریان اطلاعات اعمال می‌شوند، می‌توانند عواقب ناخواسته‌ای نیز چون شفافیت، افزایش پیچیدگی تشخیص و کاهش اعتماد به منابع داشته باشند (Tarassoli, 2024: 87). از این رو اخبار جعلی از یک سو می‌توانند بر روی افکار عمومی، تصمیم‌گیری‌های سیاسی و اجتماعی تأثیر بگذارند و باعث ایجاد تنش‌های اجتماعی شوند و از سوی دیگر عدم آگاهی کافی کاربران از روش‌های شناسایی اخبار جعلی، موجب افزایش آسیب‌پذیری آن‌ها در برابر اطلاعات نادرست خواهد شد.

مهم‌ترین گام برای جلوگیری از انتشار اخبار جعلی، شناسایی زودهنگام اخبار جعلی است. استفاده از هوش‌مصنوعی برای تشخیص دقیق، صحیح و زود هنگام اطلاعات نادرست بدون شک ارزشمند است و در این مقاله تلاش گردیده تا با بررسی فنون[1] مقابله و نبرد هوشمندانه اخبار جعلی ضمن پیش‌بینی و تشخیص زودهنگام و دقیق اطلاعات نادرست بر میزان دانش و درک مدیران حوزه امنیتی و دفاعی کشور بیافزاید. از این‌رو سؤال اصلی این تحقیق این است که فنون مقابله و نبرد هوشمندانه با اخبار جعلی در عصر هوش‌مصنوعی کدام‌اند؟

[1] Techniques

۱ـ روش‌شناسی پژوهش

از آنجا که این مطالعه به تحلیل داده‌های کیفی می‌پردازد، در زمره پژوهش‌های تفسیری قرار می‌گیرد. با توجه به قلمرو موضوعی تحقیق، جامع آماری آن شامل تمام مجموعه مقالاتی است که در بازهٔ زمانی سال‌های ۲۰۲۰ تا ۲۰۲۴ میلادی و با موضوع هوش‌مصنوعی، اخبار جعلی، فنون مقابله و نبرد هوشمندانه به چاپ رسیده‌اند که با روش فراترکیب نهایتاً ۳۳ مقاله که نزدیک‌ترین محتوا با موضوع تحقیق داشتند از میان ۶۳ مقاله انتخاب گردید. از این روش جهت شناسایی نمای کلی از یافته‌های موجود و بینش‌های جدید پیرامون موضوع تحقیق استفاده گردید. در روش فراترکیب با ترکیب یافته‌های کیفی با هدف دستیابی به بینش‌های وسیعتر، توسعه مفهومی یا سودمندی عملی صورت می‌گیرد، از یافته‌های پژوهش‌های جداگانه فراتر می‌رود و چیزی بیشتر ارائه می‌دهد (Pourkarimi and Azizi, 2024:612). در این پژوهش از راهبرد تحلیل مضمون برای تجزیه‌وتحلیل اطلاعات استفاده گردید. تحلیل مضمون، فرایندی برای تحلیل داده‌های متنی است و داده‌های پراکنده و متنوع را به داده‌هایی غنی و تفصیلی تبدیل می‌کند. تحلیل مضمون فرایندی است که می‌تواند در اکثر روش‌های کیفی به کار رود . به‌طور کلی تحلیل مضمون، روشی است برای:

الف- دیدن متن؛

ب- برداشت و درک مناسب از اطلاعات ظاهراً نامرتبط؛

ج- تحلیل اطلاعات کیفی؛

د-مشاهده نظام‌مند شخص، تعامل، گروه، موقعیت، سازمان و یا فرهنگ؛

ه- تبدیل داده‌های کیفی به داده.

در تحلیل مضمون، واحد تحلیل بیشتر از یک کلمه یا اصطلاح است و به بافت داده‌ها و نکات ظریف آن‌ها بیشتر توجه می‌شود و از شمارش کلمات و عبارات آشکار فراتر می‌رود و بر شناخت و توضیح ایده‌های صریح و ضمنی تمرکز می‌کند (Mousavi Loghman & et.al, 2023: 338).

در تحلیل مضمون، واحد تحلیل بیشتر از یک کلمه یا اصطلاح است و به بافت داده‌ها و نکات ظریف آن‌ها بیشتر توجه می‌شود و از شمارش کلمات و عبارات آشکار فراتر می‌رود و بر شناخت و توضیح ایده‌های صریح و ضمنی تمرکز می‌کند؛ سپس از کدهای مضامین اصلی برای تحلیل عمیق‌تر داده‌ها استفاده می‌شود و می‌توان از فراوانی نسبی مضامین برای مقایسه آن‌ها و تهیه ماتریس مضامین

و ترسیم شبکه مضامین استفاده کرد. مضمون، الگویی است که در داده‌ها یافت می‌شود و حداقل به توصیف و سازماندهی مشاهدات و بیشتر به تفسیر جنبه‌هایی از پدیده می‌پردازد. ابزارهای این روش تحلیل قالب مضامین و تحلیل شبکه مضامین است که به‌طور معمول در تحلیل مضمون به‌کار می‌روند. قالب مضامین فهرستی از مضامین را به‌صورت سلسله مراتبی بازگو می‌کند. شبکه مضامین، براساس روندی مشخص، مضامین پایه (کدها و نکات کلیدی متن)، مضامین سازمان‌دهنده (مضامین به‌دست‌آمده از ترکیب و تلخیص مضامین پایه) و مضامین فراگیر (مضامین عالی) دربرگیرنده اصول حاکم بر متن به‌مثابه کل را نظام‌مند می‌کند؛ سپس این مضامین به‌صورت نقشه‌های شبکه تارنما، رسم و مضامین برجسته هریک از این سه سطح همراه با روابط میان آن‌ها نشان داده می‌شود (Ibid, 339).

۲ـ پیشینه پژوهش

ترسلی (۲۰۲۴)، در تحقیقی با عنوان «کاربرد هوش‌مصنوعی در تشخیص اخبار جعلی» پس از تجزیه‌وتحلیل داده‌ها و محاسبه معیارهای ارزیابی صحت، یادآوری، دقت و امتیاز F1 (میانگین همساز) و مقایسه نتایج نشان داد که مدل یادگیری ماشین جنگل تصادفی در معیارهای صحت و دقت به ترتیب با ۹۸,۳ درصد و ۹۷,۹ درصد و مدل یادگیری تقویت گرادیان در محاسبه معیارهای امتیاز F1 و یادآوری به ترتیب با ۹۷,۷ درصد و ۹۸,۷ درصد بهترین نتیجه در تشخیص اخبار جعلی دارند.

اخگری و ممتازی (۲۰۲۳)، در تحقیقی با عنوان «کاربرد هوش‌مصنوعی در راستی آزمایی اخبار: تشخیص اخبار جعلی با استفاده از متن خبر و اطلاعات منابع منتشرکننده خبر» با استفاده از مجموعه داده «تات» که برای زبان فارسی مناسب بوده و شامل ۱۰۸۱ خبر جعلی و ۱۰۸۱ خبر با برچسب غیرجعلی در حوزه‌های مختلف خبری از ۳۸ کانال تلگرامی است. نشان دادند که استفاده از شبکه عصبی پیچشی، و شناسایی اخبار فارسی منتشرشده در تلگرام و استفاده از متن خبرهای منتشرشده به همراه شناسه کانال ارسال‌کننده خبر به‌عنوان ورودی شبکه توانسته است به صحت ۹۰,۴۶ درصد در تشخیص اخبار جعلی دست یابد.

امامی رودسری (۲۰۲۳)، در تحقیقی با عنوان «روش‌ها و ترفندهای مقابله با اخبار جعلی» نشان داد که ترفندهای مختلفی برای تشخیص اخبار جعلی در رسانه‌های اجتماعی وجود دارد، از جمله بررسی دستی و جستجو در منابع در دسترس، تا شیوه‌های جدیدی که به یادگیری عمیق و فناوری‌های پیشرفته هوش‌مصنوعی وابسته است. برای مقابله با خبرهای جعلی در کشور، ورود شرکت‌های

دانش‌بنیان برای تهیه سامانه‌های مبتنی بر هوش‌مصنوعی با همکاری فعالان رسانه‌ای اهمیت چشم‌گیر دارد.

الاصفا و شاتی[1] (۲۰۲۴)، در تحقیقی با عنوان «نظرسنجی در مورد تشخیص اخبار جعلی در رسانه‌های اجتماعی با استفاده از شبکه‌های عصبی گراف» نشان دادند که شبکه‌های عصبی گراف به‌عنوان راه‌حل برتر به‌ویژه در پردازش داده‌های ساختاریافته گرافی مؤثر هستند و به آن‌ها اجازه می‌دهد تا ارتباطات پیچیده و الگوهای انتشار اخبار در شبکه‌های اجتماعی را با دقت بیشتری مدل کنند.

۳ـ مبانی نظری

چارچوب نظری پژوهش حاضر، مجموعه‌ای نظام‌مند از فرضیه‌های اطمینان بخش و کم و بیش تجربه شده و تعمیم یافته در بررسی ادبیات موضوع (پیشینه پژوهش) است که پایه و پشتوانه تحلیل‌ها و شکل استدلال‌های محقق را در طرح پژوهش تشکیل می‌دهد. در این تحقیق سعی گردید تا در خلال تحلیل داده‌های تحقیق، ضمن بررسی فنون مقابله و نبرد هوشمندانه با اخبار جعلی در عصر هوش‌مصنوعی مصادیق جدیدی برای آن فرضیه‌های کلی ارائه گردد و همانطور که در روش شناسی تحقیق اشاره گردید، با استفاده از رویکرد کیفی در این پژوهش ، مبانی گردآوری و تحلیل داده‌های پژوهش از پارادایم تفسیری گرفته شد که رابطه ای اعم و اخص با چارچوب نظری پژوهش دارد.

۱ـ اخبار جعلی: اصطلاح اخبار جعلی به پدیده‌ای کاملاً فراگیر در سراسر جهان تبدیل شده است و در اصطلاح رایج، چیزی را که فاقد حقیقت و صداقت است القا می کند و ممکن است همه جا در گفتمان سیاسی، عمومی و رسانه‌ای وجود داشته باشد، اما هدف آن تنها فریب دادن، آسیب رساندن، هتک حرمت یا توهین است. با درگیر شدن جهان توسط رسانه‌های رقمی[2]، اخبار جعلی به پدیده‌ای ناشناخته تبدیل شده است که همیشه سعی می کند بر اهداف آسیب‌پذیر انتخاب شده خود تأثیر بگذارد (Soetekouw, et.al, 2024: 460).

۱ـ۱ـ انواع اخبار جعلی: انواع اخبار جعلی منتشره در رسانه‌های اجتماعی بر دو گونه هستند، اطلاعات غلط[3] که بدون قصد فریب، منتشر می‌شوند و اطلاعات نادرست[4] که با قصد

---

[1] Alaa Safaa & Shati
[2] Digital
[3] Misinformation
[4] Disinformation

فریب منتشر می‌شوند. اطلاعات غلط توسط کسانی که اخبار یا مطالب بی‌شماری را با تصور صحت آنها به اشتراک می‌گذارند، منتشر می‌شوند. آن‌ها آن اخبار را به اشتراک می‌گذارند؛ زیرا دوستانشان این کار را انجام داده‌اند. همان‌طور که همه ما می‌دانیم، الگوریتم توصیه هوش‌مصنوعی در رسانه‌های اجتماعی اخبار یا اطلاعاتی را به کاربران خود توصیه می‌کنند. آن‌ها سابقه قبلی کاربران را بررسی می‌کنند و هر زمان که دوستانشان چیزی را مشاهده می‌کنند، به دوستان آن‌ها نیز توصیه و اطلاع می‌دهد که فلان خبر، محتوا یا ویدیو توسط دوستانش شما مشاهده شده است و به‌طور غیرمستقیم آن‌ها را تحت تأثیر قرار می‌دهد تا آن را پسند[1] کنند و با دیگران به اشتراک بگذارند. همه این اشتراک‌گذاری‌ها بدون اطلاع از صحت مطالب به اشتراک‌گذاشته‌شده ادامه می‌یابد و اشتراک‌گذاری در میان کسانی که مشترک سیستم اعتقادی یکسانی هستند یا بخشی از یک نظام اجتماعی- اقتصادی یا سیاسی هستند، سریع‌تر می‌شود. اطلاعات نادرست توسط کسانی که اخبار یا مطالب بی‌شماری را با آگاهی از اصالت (جعلی) آن‌ها و برای به دست آوردن منافع سیاسی یا مالی به اشتراک‌گذاری و انتقال آن به دیگران به اشتراک می‌گذارند، منتشر می‌شوند. در مواردی از حساب‌کاربری‌های مخرب[2] به نام ربات‌ها و اوباش مجازی رسانه‌های اجتماعی[3] استفاده می‌کنند که به شکل انسانی تعامل دارند، با رفتار مخرب در فضای وب به‌دنبال جلب نظر کاربران، ایجاد تشنج و بیان مطالب تحریک‌کننده و توهین‌آمیز هستند. ربات‌ها به راحتی پیام می‌فرستند، پاسخ می‌دهند، هزاران حساب‌کاربری را دنبال می‌کنند و به همان اندازه مطالب[4] اشتراک‌گذاری شده در فیس‌بوک، اینستاگرام، ایکس (توئیتر) و... را فوراً به اشتراک می‌گذارند (Akhtar, et.al, 2024: 9).

۱ـ۲ـ انواع اخبار غلط: اخبار غلط را می‌توان به دسته‌های مختلفی از جمله طعمه کلیک، تقلب، تبلیغات، طنز و هجو و... طبقه‌بندی کرد:

– **طعمه کلیک**[5]: یک متن یا یک پیوند کوچک[6] است که برای جلب توجه و ترغیب کاربران به دنبال کردن (کلیک) آن لینک و خواندن، مشاهده یا گوش دادن به قطعه لینک

---

[1] Like
[2] Malicious account
[3] Social media bots and trolls
[4] Posts
[5] Clickbait
[6] Thumbnail link

شده از محتوای آنلاین طراحی شده است، که معمولاً فریبنده، هیجان انگیز یا گمراه کننده است. طعمه کلیک از عبارات ارجاع دهنده تعریف شده همراه با موارد فوق برای ایجاد شکاف اطلاعاتی[1] استفاده می‌کند، که خواننده را وادار می‌کند تا روی پیوند مرتبط با انتظار برای یافتن اطلاعات مرتبط کلیک کند. کاربران می‌توانند در رسانه‌های اجتماعی با پیام‌هایی که ذاتاً مرتبط با آنها بوده و یا با استفاده از اطلاعات عمومی خودشان تولیدشده، هدف قرار گیرند (Shrestha, et.al, 2024: 1).

- **پروپاگاندا[2] (تبلیغات سیاسی):** پروپاگاندا نوعی ارتباط است که هدف آن تأثیرگذاری بر عقاید، ارزش‌ها یا اَعمال افراد یا گروه‌ها از طریق ارائه اطلاعات به شیوه‌ای مغرضانه یا گمراه کننده است. دولت‌ها، احزاب سیاسی، شرکت‌ها و سایر سازمان‌ها اغلب از آن برای شکل دادن به برداشت‌های عمومی و دستکاری افکار عمومی استفاده می‌کنند و می‌تواند اشکال مختلفی داشته باشد، از جمله تبلیغات، گزارش‌های رسانه‌ای، سخنرانی‌ها و پیام‌های رسانه‌های اجتماعی. پروپاگاندا به‌دنبال آن است که با توسل به احساسات، دستکاری حقایق و استفاده از روش‌های متقاعدکننده، نگرش‌ها و رفتارهای مردم را در جهتی خاص سوق دهد این نوع اخبار را می‌توان با استفاده از مدل‌های تشخیص دستی مبتنی بر واقعیت مانند بررسی‌کننده‌های واقعیت مبتنی بر متخصص و روش‌های جمع‌سپاری شناسایی کرد (Tikkanen, 2024: 34).

-**طنز و هجو:** این نوع اخبار جعلی شکل اغراق آمیزی از داستان ساختگی است که در رسانه‌ها در قالب کمدی گزارش می‌شود. از سبک طنز برای اطلاع‌رسانی به‌روز اخبار به مخاطبان استفاده می‌کند. در اینجا تفاوت اصلی این است که این اخبار جعلی «جعلی بودن» خود را پنهان نمی‌کنند و حتی آن را به رخ می‌کشند ارائه‌دهندگان چنین اخباری به جای روزنامه‌نگار، مجری و کمدین هستند، زیرا مخاطبان می‌دانند که مجریان چنین اخباری را به‌گونه‌ای از رسانه‌ها پخش می‌کنند که مورد علاقه آن‌ها باشد. اگرچه گاهی اوقات روزنامه‌نگاران فاقد صلاحیت سایر رسانه‌ها، نشریات این رسانه را واقعی می‌دانند و اطلاعات تخیلی را واقعی در صفحات خود منتشر می‌کنند. در چنین شرایطی اخبار طنز به خبر جعلی تبدیل می‌شود. این یکی دیگر از شواهدی است که نشان می‌دهد طنز و هجو سریع‌تر از واقعیت‌ها پخش می‌شوند (Kitsa, 2023: 3).

---

[1] Information gap
[2] Propaganda

- **حُقه**[1]: برخلاف اخبار جعلی، که یک دروغ تمام عیار هستند نیمی از اخبار از نوع حُقه و دغل، حقیقت دارند. چنین اخباری برای فریب افکار عمومی و مخاطبان تهیه می‌شوند. آن‌ها عامدانه تولید می‌شوند و به‌گونه‌ای ساخته می‌شوند که گاه در قالب گزارش رسانه‌های اصلی و واقعی تعبیر می‌شوند. این نوع اخبار در سطح وسیع ساخته شده و خسارت هنگفتی را به دنبال دارد. هدف آن بیشتر تمرکز بر یک شخصیت عمومی است (Rastogi & Bansal, 2023:183).

- **دزدی نام، قاب‌بندی، فریب روزنامه‌نگاری**[2]: این نوع اخبار جعلی اساساً هویت یک منبع خبری معتبر را می‌دزدند تا مخاطب باور کند که اخبار آن‌ها درست است. با ایجاد وبگاه شروع می‌شود که از وبگاه خبری معتبر موجود برای فریب مخاطب تقلید می‌کند. آن‌ها حتی یک کپی از یک علامت تجاری مورد استفاده توسط وبگاه واقعی را به نمایش می‌گذارند، آن‌ها همه این ترفندها را برای جلب اعتماد مخاطبان و تثبیت خود در بازار انجام می‌دهند. به‌طور منطقی، مردم مفاهیم خاصی را بر اساس روش ابداع آن درک خواهند کرد، مصرف‌کنندگان معمولاً اگر چیزی را به دو صورت متفاوت قاب‌بندی کنند، به طور متفاوتی درک می‌کنند، اگرچه همه آن‌ها به یک معنا هستند. قاب‌بندی به منظور ارائه تصورات غلط و پنهان کردن واقعیت است، تفاوت بین فریب روزنامه‌نگاری و قاب‌بندی مربوط به نویسنده است، قاب‌بندی را هرکسی می‌تواند انجام دهد، اما فریب‌های روزنامه‌نگاری معمولا توسط روزنامه‌نگاران در یک رسانه معروف انجام می‌شود (Collins, et.al, 2021:253).

۴ـ رویکردهای تشخیص اخبار جعلی

برای مقابله و نبرد هوشمندانه با اخبار جعلی به استناد آیه شریف ۶ از سوره مبارک حجرات مهم‌ترین گام در زمان دریافت هرگونه خبر، بررسی و تبیین است:

يَا أَيُّهَا الَّذِينَ آمَنُوا إِنْ جَاءَكُمْ فَاسِقٌ بِنَبَإٍ فَتَبَيَّنُوا أَنْ تُصِيبُوا قَوْمًا بِجَهَالَةٍ فَتُصْبِحُوا عَلَى مَا فَعَلْتُمْ نَادِمِينَ؛ ای کسانی‌که ایمان آورده‌اید! اگر فاسقی برای شما خبری مهم آورد تحقیق کنید، مبادا (از روی زودباوری و شتاب‌زدگی تصمیم بگیرید و) ناآگاهانه به گروهی آسیب رسانید، سپس از کرده‌ی خود پشیمان شوید.[3]

---

[1] Hoaxes
[2] Name-theft, framing, journalism deception
[3] تفسیر نور، جلد ۹ - صفحه ۱۶۷

از این‌رو با بررسی دیدگاه‌های محققین و پژوهش‌های انجام شده در زمینه تشخیص اخبار جعلی، چهار رویکرد برای بررسی و تشخیص اخبار جعلی شناسائی گردید که فناوری هوش مصنوعی در هر چهار دسته با خودکارسازی فرایند بررسی کاربرد و نقش‌آفرینی برجسته‌ای داشته و محققان در پژوهش‌های خود به انواع الگوریتم‌های مؤثر در آن‌ها اشاره کرده‌اند. آن‌ها عبارت‌اند از: (۱). دانش نادرستی که اخبار را منتقل می‌کند؛ (۲). سبک نوشتن یا نوع محتوای اخبار؛ (۳). الگوهای انتشار یا رفتارهای اجتماعی اخبار؛ (۴). اعتبار منبع اخبار.

ضمن خلاصه‌سازی این چهار رویکرد در شکل ۱ و در جدول ۱ دسته‌بندی و تجزیه‌وتحلیل داده‌های مربوط به رویکردهای تشخیص اخبار جعلی برگرفته از تحلیل مضمون پژوهش‌های انجام گرفته محققان این حوزه نشان داده شده است:

شکل-۱. رویکردهای تشخیص اخبار جعلی (منبع: محقق ساخته)

جدول-۱. دسته‌بندی و تجزیه‌وتحلیل داده‌های مربوط به رویکردهای تشخیص اخبار جعلی

ردیف	نویسنده	مضامین پایه	مضامین سازمان‌دهنده	مضامین فراگیر	منبع
۱	La Barbera, et.al	جمع‌سپاری می تواند ابزار مفیدی برای ارزیابی صحت اظهارات باشد.	جمع‌سپاری[1]	دانش	La Barbera, et.al, 2024:2
۲	Yildirim	مدل‌های یادگیری عمیق ترکیبی، متخصص-جمع‌سپاری، ماشین-جمع‌سپاری، و ادغام روش‌ها از مدل‌های مبتنی	متخصص-جمع‌سپاری، ماشین-جمع‌سپاری		Yildirim, 2023:11

[1] Crowdsource

		بر محتوا و الگوریتم‌های مبتنی بر زمینه اجتماعی نیز برای تشخیص اخبار جعلی پیشنهاد شده‌اند.			
۳	La Barbera, et.al	تحقیقاتی برای ترکیب روش‌های جمع‌سپاری و خودکار برای توسعه رویکردهای انسان در حلقه[1] برای تشخیص اطلاعات غلط انجام شده است.	همکاری جمعی انسان و ماشین		La Barbera, et.al, 2024:3
۴	Dong	بازخورد کاربران حقیقت‌یاب در حلقه یادگیری و استنتاج توان تشخیص اطلاعات نادرست را به روشی غیرمتمرکز افزایش دهند.	یادگیری و استنتاج انسان		Dong, et.al, 2022:47
۵	Godel	مدل‌های یادگیری ماشینی با استفاده از داده‌های جمع‌سپاری بهتر از روش‌های جمع‌بندی ساده، در تشخیص اخبار جعلی عمل می‌کنند	همکاری جمعی انسان و ماشین		Godel, et.al, 2021:3
۶	Jain, et.al	مدل یادگیری عمیق شبکه عصبی پیچشی از اخبار نادرست یا معتبر برای استخراج ویژگی‌های پنهان و ارزیابی دقت، صحت، یادآوری و عملکرد امتیاز f1 استفاده می‌کند.	استخراج ویژگی‌های پنهان[2]	محتوا	Jain, et.al, 2024:6
۷	Alghamdi, et.al	یافتن وابستگی های طولانی مدت و ویژگی‌های نهفته محلی برجسته برای پیشرفت عملکرد تشخیص اخبار ضروری است.	استخراج ویژگی‌های پنهان		Alghamdi, et.al, 2024:4
۸	Abishethvarman, et.al	تحلیل احساسات، جنبه‌ای از پردازش زبان طبیعی،	ویژگی‌های احساسی۳		Abishethvarman, et.al, 2024:1

[1] Human-In-The-Loop (HITL)
[2] Latent features
[3] Syntactical features

		احساسات ظریف درون متن را بررسی می‌کند، روایت‌های پنهان را آشکار می‌کند و بین گزارش‌های معتبر و فریب‌کاری که با دقت ساخته شده تفاوت قائل می‌شود			
۹	Alghamdi, et.al	ترکیبی از ویژگی‌های زبانی و ویژگی‌های روانی برای شناسایی پخش‌کنندگان اخبار جعلی استفاده شده که به دقت ۶۷،۵۰ درصد دست یافتند.	ویژگی‌های زبانی ویژگی‌های روانی		Alghamdi, et.al, 2024:4
۱۰	Ali, et.al	تحلیل متون مقالات خبری برای تشخیص اخبار جعلی شامل استفاده از پردازش زبان طبیعی[1] و فنون یادگیری ماشینی برای استخراج ویژگی‌هایی مانند احساسات، عواطف و الگوهای زبانی از متن است. این قابلیت زمانی که با منابع دانش خارجی ترکیب شود، می‌تواند با توانایی مدل‌ها در تشخیص اشکال پیچیده اطلاعات نادرست استفاده شود.	ویژگی‌های احساسی ویژگی‌های زبانی		Ali, et.al, 2023:6
۱۱	Nasery	پویایی انتشار اطلاعات در رسانه‌های اجتماعی می‌تواند بینش ارزشمندی برای تشخیص اطلاعات درست از نادرست ارائه دهد	پویائی انتشار	انتشار	Nasery, 2024:33
۱۲	Ghosh, et.al	ترکیب شبکه‌های نمودار زمانی[2] و شبکه‌های عصبی مکرر[3] هم ویژگی‌های ساختاری و هم زمانی انتشار	زمان انتشار		Ghosh, et.al, 2024:1

[1] Natural language processing (NLP)
[2] Temporal Graph Network (TGN)
[3] Recurrent Neural Networks (RNNs)

		اطلاعات غلط را به تصویر می‌کشد.			
۱۳	Ghosh, et.al	قابلیت انتشار سریع و گسترده اطلاعات بدون توجه به صحت آن از طریق ساختارهای شبکه پیچیده مشکل شناسایی و توقف انتشار اخبار جعلی در مراحل اولیه آن را افزایش می‌دهد.	انتشار سریع		Ghosh, et.al, 2024:1
۱۴	Ghosh, et.al	یکی از ویژگی‌های مهم برای یک رویداد شامل متن توییت، اطلاعات نمایه کاربر و تعداد ریتوییت‌ها/پسندها/ پاسخ‌ها در زمان رویداد است.	ریتوییت‌ها پسندها/پاسخ‌ها		Ghosh, et.al, 2024:4
۱۵	Liu, et.al	مطالعاتی بر اندازه آبشار، طول عمر آبشار و ویروسی بودن ساختاری برای تجزیه‌وتحلیل الگوهای انتشار اطلاعات نادرست استفاده می‌کنند. اندازه آبشار مربوط به تعداد ارسال‌های تولید شده توسط یک آبشار است. طول عمر آبشار مدت زمانی است که یک آبشار شایعه فعال باقی می‌ماند، یعنی زمان سپری شده بین پخش اصلی و ارسال نهایی و ویروسی بودن ساختاری یک معیار بررسی عمق و وسعت یک آبشار را ارائه می‌دهد.	نمودار آبشاری		Liu, et.al, 2024:5
۱۶	Chan	شخصیت‌های جعلی با توزیع‌های مختلف سن، جنسیت، قومیت، میانگین درآمد، موقعیت جغرافیایی، نمایه‌های سیاسی و غیره توسط الگوریتم‌های خودکار «ربات‌های اجتماعی» اجرا	سن، جنسیت، قومیت، میانگین درآمد، موقعیت جغرافیایی، نمایه‌های سیاسی کاربر	کاربر	Chan, 2024:3

		شده و از کاربران واقعی تقلید می‌کنند. برخی از این ربات‌ها ابتدایی‌تر هستند و ممکن است فقط وظایف استانداردی مانند «پسند کردن» یا «بازتوییت» از کاربران خاص یا ارسال پیام‌های از پیش تعیین‌شده را انجام دهند.			
۱۷	Mittal	مدل‌سازی روش تعامل و تعدد فعالیت کاربر در رسانه‌های اجتماعی از طریق ویرایش و فیلترهای اعمال شده بر روی تصاویر قبل از ارسال آن‌ها در رسانه‌های مختلف اجتماعی برای گرفتن تصاویر است. انجام این کار ممکن است منعکس کننده منابعی باشد که برای تهیه پیام صرف شده است و نشان دهنده نگرش سازنده نسبت به تصویری باشد که به اشتراک می‌گذارد.	مدل‌سازی روش تعامل و تعدد فعالیت کاربر		Mittal, et.al, 2024:4
۱۸	Shahid, et.al	ویژگی‌های مبتنی بر کاربر را میتوان به تجزیه‌وتحلیل نمایه و تجزیه‌وتحلیل اعتبار تقسیم کرد. در تجزیه‌وتحلیل نمایه کاربر، مشخصات کاربر را می توان بر اساس نام کاربری، سن، تصویر نمایه، موقعیت جغرافیایی و وضعیت تأیید حساب مورد تجزیه‌وتحلیل قرار داد. تجزیه‌وتحلیل اعتبار کاربر شامل اطلاعاتی در مورد تعداد دوستان و دنبال‌کنندگان است. به عنوان مثال، حساب‌های ربات معمولاً کاربران بیشتری در فهرست دنبال‌کنندگان خود	نام کاربری، سن، تصویر نمایه، موقعیت جغرافیایی تعداد دوستان و دنبال‌کنندگان		Shahid, et.al, 2022:5

		دارند و خود کاربران بسیار کمتری را دنبال می‌کنند			
۱۹	Steinfeld	در شناسایی اطلاعات نادرست، نشانه‌های مختلف در مورد پیام‌های رسانه‌های اجتماعی و محتوایی که بخشی از متن پیام نیستند، می‌توانند پیش‌بینی کننده‌های قابل اعتماد و مهمی برای اعتبار پیام باشند. این موارد شامل طول پیام/ توییت، استفاده از علامت تعجب یا علامت سوال، تعداد بازتوییت‌ها/ اشتراک‌گذاری‌ها/ پسندها یا منشن‌ها، تعداد دوستان/ دنبال‌کنندگان کاربر، سن حساب، تعداد ایموجی‌ها، تعداد آدرس‌های اینترنتی موجود در پیام/ توئیت، بیوگرافی و وابستگی سیاسی، فراداده‌های تصویر (مکان جغرافیایی، اعتبار تصویر) و پاسخ‌ها به متن توییت/ پیام.	تعداد دوستان/ دنبال‌کنندگان کاربر، سن حساب،		Steinfeld, 2023:3

۵ـ متن‌کاوی

استخراج اطلاعات از متن (متن کاوی) از پردازش زبان طبیعی برای تبدیل داده‌های متنی بدون ساختار به داده‌های عادی و ساختاریافته مناسب برای تحلیل یا هدایت الگوریتم‌های یادگیری ماشین استفاده می‌کند (Lybarger, 2023:3). در یادگیری ماشین، نمایش داده‌ها تا حد زیادی بر صحت نتایج تأثیر می‌گذارد. در رسانه‌های اجتماعی داده‌های متنی که توسط کاربران به اشتراک گذاشته می‌شوند، عموماً به شکل‌های بدون ساختار هستند. به همین دلیل، داده‌های بدون ساختار استخراج شده از رسانه‌های اجتماعی باید با فنون متن‌کاوی و پردازش زبان طبیعی[1] به شکل ساختاریافته تبدیل شوند. مشکل متن‌کاوی را می‌توان به‌عنوان استخراج

[1] Natural Language Processing

اطلاعات معنی‌دار، مفید و ناشناخته قبلی از داده‌های متنی تعریف کرد. روش متن‌کاوی با پیش‌پردازش داده‌ها آغاز می‌شود که شامل سه مرحله است که عبارت‌اند از: بن‌واژه‌سازی متن، حذف کلمات توقف و استخراج ویژگی (Tarassoli, 2024:91). با الگوریتم بردارسازی متن[1]، که در بردار TF-IDF که از روی کلمات می‌تواند ویژگی‌های مختلف را برای یک متن بسازد و در واقع متن را به اعداد قابل فهم برای الگوریتم تبدیل می‌کند. این روش مزایای متعددی را در تجزیه‌وتحلیل متن ارائه می‌دهد. اولاً، با اختصاص وزن‌های بالاتر به اصطلاحاتی که هم در یک سند متداول هستند و هم در مجموعه کلی نادر هستند، اهمیت کلمات را در نظر می‌گیرد. این روش به مدل اجازه می‌دهد تا قدرت تمایز کلمات را به دست آورد و ویژگی‌های کلیدی را برای تجزیه‌وتحلیل شناسایی کند. ثانیاً، تأثیر کلمات رایج (مانند «the» یا «و») را که برای تشخیص اسناد آموزنده نیستند، کاهش می‌دهد (Ibid, 94 and Baguian & Ashley, 2024:4)، به‌طور کلی در این روش با تحلیل متن و تحلیل احساسات نشانه‌های فریب در متن با استفاده از ابزار زبان طبیعی، شناسایی می‌شوند.

۶ـ روش یادگیری ماشینی

ترسلی، ۱۴۰۳، استفاده از الگوریتم‌های طبقه‌بندی یادگیری ماشین را برای شناسایی اخبار جعلی پیشنهاد کرد. نتایج تحقیق بررسی پنج مدل یادگیری ماشین نظارت شده نشان داد که الگوریتم تقویت گرادیان[2] که در معیارهای امتیاز F1 و یادآوری، بالاترین امتیازات را به دست آورد و تعادل بسیار خوبی بین معیارهای دقت و فراخوانی ارائه نمود، از روش تقویت استفاده می‌کند و درخت‌های تصمیم‌گیری در تشخیص اخبار جعلی را یکی یکی می‌سازد، به‌گونه‌ای که هر درخت از درختی که قبلاً آموزش داده شده بهره می‌برد و خطاهای خود را تصحیح و باعث ایجاد یک یادگیرنده قوی می‌شود. الگوریتم جنگل تصادفی[3] که قدرت تشخیص بسیار خوبی در معیارهای صحت و دقت ارائه داد، هر درخت را جداگانه آموزش می‌دهد و نیاز به رأی‌گیری برای تجمیع مدل دارد، به بیان ساده، جنگل تصادفی چندین درخت تصمیم ساخته و آن‌ها را با یک‌دیگر ادغام می‌کند تا پیش‌بینی‌های صحیح‌تر و پایدارتری در تشخیص اخبار جعلی حاصل شوند. ترکیب فنون متن‌کاوی به این مدل اجازه داد تا قابلیت‌های تعمیم و

[1] Text Vectorization
[2] Gradiant boosting
[3] Random forest

عملکرد طبقه‌بندی‌کننده اخبار جعلی را بهبود بخشد. بنابراین، مدل پیشنهادی مبتنی بر طبقه‌بندی کننده‌های ماشینی ظرفیت ایجاد یک محیط غنی از دانش را دارد که می‌تواند به طور معناداری با طبقه‌بندی دقیق اخبار به حداقل رساندن انتشار اخبار جعلی در سکوهای رسانه‌های اجتماعی کمک کند (Tarassoli, 2024:112).

۷ـ رویکرد یادگیری عمیق

یادگیری عمیق زیرمجموعه‌ای از الگوریتم‌های یادگیری ماشینی است که از شبکه‌های عصبی چندلایه (عمیق) برای شبیه‌سازی رفتار مغز انسان و یادگیری روابط پیچیده بین کلان‌داده‌ها استفاده می‌کند. یادگیری عمیق، نیاز به پیش‌پردازش داده کمتری توسط انسان دارد و اغلب می‌تواند نتایج دقیق‌تری نسبت به رویکردهای یادگیری ماشین سنتی ایجاد کند (Ventre, 2020: 57). الگوریتم‌های یادگیری عمیق به‌ویژه شبکه‌های عصبی کانولوشن[1] دقت بالایی را در این زمینه به‌دست آورده‌اند. این مدل‌ها با استفاده از لایه‌های کانولوشن و زیرنمونه‌گیری، ویژگی‌های مهم متن را استخراج کرده و سپس توسط لایه‌های تمام‌متصل برای طبقه‌بندی اخبار به جعلی یا واقعی استفاده می‌شوند (Akhgari & Momtazi, 2023:262). برخی از محققان نیز در سال‌های اخیر با استفاده ترکیبی از انواع الگوریتم‌های یادگیری عمیق شامل شبکه عصبی کانولوشن، شبکه عصبی بازگشتی[2]، شبکه عصبی نمودار[3]، شبکه مبتنی بر توجه[4]، شبکه متخاصم مولد[5]، حافظه کوتاه‌مدت بلندمدت[6] و... به موفقیت‌هایی در تشخیص اخبار جعلی دست یافته‌اند (Ayetiran & Özgöbek, 2024:6).

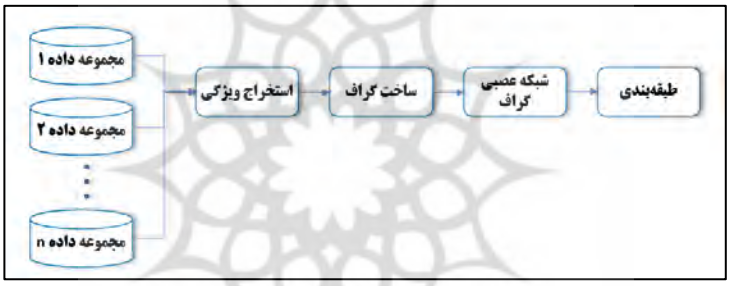
۸ـ روش مبتنی بر نمودار

شبکه‌های عصبی گراف‌پایه[7] به عنوان یک ابزار پیشرفته در تشخیص اخبار جعلی به کار می‌روند. این شبکه‌ها به دلیل توانایی‌شان در مدل‌سازی ارتباطات پیچیده بین داده‌ها، به ویژه در زمینه‌های اجتماعی و متنی، مورد توجه قرار گرفته‌اند. شبکه‌های عصبی گراف[8] می‌توانند

---

[1] Convolutional Neural Network (CNN)
[2] Recurrent Neural Network (RNN)
[3] Graph Neural Network (GNN)
[4] aAttention-based Network
[5] Generative Adversarial Network (GAN)
[6] Long Short-Term Memory (LSTM)
[7] Graph-based Neural Networks (GNN)
[8] Graph Neural Networks(GNN)

به‌طور مؤثری روابط بین کاربران، اخبار و منابع مختلف را مدل‌سازی کنند. این قابلیت به تحلیل عمیق‌تری از چگونگی انتشار اخبار و شناسایی الگوهای مشکوک کمک می‌کند. شبکه‌های عصبی گراف می‌توانند داده‌های غیرساختاری مانند متون و گراف‌های اجتماعی را تحلیل کنند. این ویژگی به آن‌ها این امکان را می‌دهد که به شناسایی نشانه‌های اخبار جعلی بپردازند، به‌ویژه در سکوهای اجتماعی که اخبار به سرعت منتشر می‌شوند. تحقیقات نشان می‌دهند که استفاده از شبکه‌های عصبی گراف در مقایسه با فنون سنتی، دقت بالاتری در تشخیص اخبار جعلی دارد. این شبکه‌ها می‌توانند با یادگیری از ساختارهای گرافی، به شناسایی اخبار واقعی از جعلی کمک کنند (Mahdi & Shati, 2024:1). با مبارزه از طریق شناسائی ربات‌های اجتماعی که اخبار جعلی را منتشر می‌کنند، در شناسایی حساب‌های جعلی در رسانه‌های اجتماعی بسیار مؤثر می‌باشد. فرآیند تشخیص اخبار جعلی با استفاده از مدل شبکه‌های عصبی گراف شامل چهار مرحله کلیدی است که در شکل ۲ نشان داده شده است:



شکل-۲. مدل شبکه‌های عصبی گراف در تشخیص اخبار جعلی (Lakzaei, et.al, 2024:3)

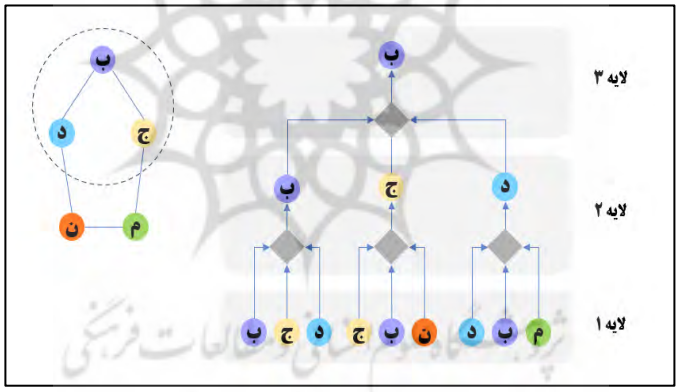
۱- استخراج ویژگی: در این مرحله، مشکلات متعددی در تشخیص اخبار جعلی مانند مجموعه داده‌های ناقص یا مشکلات یافت شده در مجموعه داده‌ها اعم از تصویری، متنی و... وجود دارد. برخی از این مشکلات را می‌توان از طریق کشف ویژگی‌های مهم و مفید در مجموعه داده‌ها برطرف کرد (Madani, et.al, 2024:14).

۲- ساخت گراف: در این مرحله روشی مناسب برای ساخت گراف انتخاب می‌شود که شامل ایجاد گراف تشابه، گراف ناهمگن و یا گراف انتشار می‌باشد. از طریق این فرآیند، مجموعه داده اصلی- که از داده‌های بدون ساختار تشکیل شده بود- به یک فرمت نمودار ساختاریافته تبدیل شد (Lakzaei, et.al, 2024:3).

۳- شبکه عصبی گراف: برای پردازش گراف ایجاد شده در مرحله قبل از شبکه عصبی گراف استفاده می‌شود. هر گره در نمودار دارای یک بردار جاسازی است که توسط این شبکه

تولید شده است. شبکه عصبی گراف‌ها تابعی را می‌آموزند که برای هر گره یک بردار را در یک فضای برداری با ابعاد کم به هم متصل می‌کند. هنگامی که دو گره دارای ویژگی‌ها و مسئولیت‌های ساختاری مشابهی هستند، ساختار شبکه باید به روش حفظ شباهت به گره‌های مجاور در فضای برداری نگاشت شود (Ibid).

شکل ۳ عملکرد یک شبکه عصبی گراف را نشان می‌دهد. ابتدا، برای هر گره، یک همسایه ساخته می‌شود، سپس، هر گره و بردارهای وارده از لایه قبلی آن و از همسایگان اطراف در معرض یک تبدیل خطی (ماتریس وزن قابل آموزش) و یک عملیات تجمیع در هر لایه (مجموع، میانگین، حداکثر، حداقل و...) قرار می‌گیرد. این ساختار به‌طور مکرر به روز می‌شود، با یک جاسازی جدید برای هر گره که در هر تکرار (لایه) محاسبه می‌شود. ممکن است یک سیستم انتقال پیام برای توصیف این فرآیند استفاده شود. سه وظیفه برای هر گره وجود دارد: (۱). تمام بردارهای همسایگان خود را جمع آوری می‌کند (ارسال پیام)، (۲). پیام های جمع آوری شده را با استفاده از یک تابع تجمع پردازش می‌کند و (۳). بردار خود را به روز می‌کند (Alaa Safaa & Shati, 2024:9).



شکل-۳. عملکرد یک شبکه عصبی گراف (Alaa Safaa & Shati, 2024:9)

نتیجه‌گیری

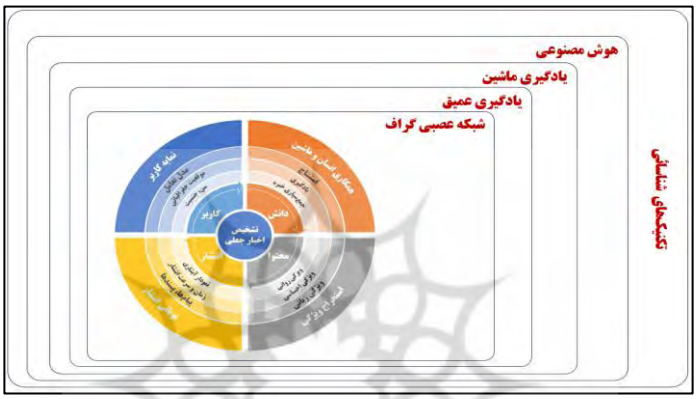
وقتی وارد مباحث فزاینده در مورد مدل‌های تشخیص اخبار جعلی می‌شویم، کاملاً مشخص می‌شود که مدل‌های موجود این بن‌بست را حل نمی‌کنند. زیرا، مطمئناً به مدل‌های پیشرفته‌تر و اثبات شده‌ی علمی بیشتری نیاز خواهیم داشت. قضاوت و تشخیص‌های مبتنی بر افراد ممکن است نتایجی را به همراه داشته باشد اما بررسی دستی مستلزم استفاده از متخصصان

و کارشناسان و مملو از محدودیت‌هایی مانند زمان و درگیری کار است و وقتی با اطلاعات عظیمی مواجه می‌شوند گرفتار می‌شوند. راستی‌آزمایی خودکار و هوشمند مطمئناً به سادگی حجم زیادی از اطلاعات را در یک بازه زمانی کوتاه بررسی می‌کند. این روش با بهره‌گیری از توانایی‌های هوش مصنوعی، به عنوان یک ابزار قدرتمند در تشخیص اخبار جعلی ظهور کرده و با پردازش حجم عظیمی از داده‌ها در زمان بسیار کوتاه، قادر به شناسایی الگوها، تناقضات و نشانه‌های دروغین در اخبار هستند که برای انسان‌ها به سختی قابل تشخیص است. اما این روش نیز در مواجهه با پیچیدگی مداوم و در حال تکامل اخبار جعلی، تغییرات و ظرافت‌های زبانی و اطلاعات ناکافی تشخیص قطعی صحت اخبار و غیره به چالش‌ها و موانعی می‌رسد؛ زیرا اکثر الگوریتم‌های یادگیری ماشین‌های خودکار اخبار جعلی را با استفاده از واژگان و محتواها و سبک‌های متنی تشخیص می‌دهند.

همان‌طور که روش‌ها و سازوکارهای جدید برای بررسی و شناسایی اخبار جعلی در حال تکامل هستند، سازنده‌ها و تولیدکنندگان اخبار جعلی نیز باهوش‌تر می‌شوند و ابزارهایی را برای جلوگیری از هرگونه حمله احتمالی به فریب‌کاری فنی خود ابداع می‌کنند. با این وجود اخیراً، محققان به این نتیجه رسیده‌اند که همکاری انسان و ماشین در قالب استفاده همزمان مدل‌های یادگیری ماشین و یادگیری عمیق از روش‌های مبتنی بر نمودار نتایج بهتری در تشخیص خودکار و بهنگام اخبار جعلی به همراه داشته است. مدل‌های یادگیری ماشین و یادگیری عمیق به دلیل توانایی‌شان در تحلیل حجم عظیمی از داده‌ها و کشف الگوهای پیچیده، ابزارهای قدرتمندی برای تشخیص اخبار جعلی محسوب می‌شوند. این مدل‌ها می‌توانند با بررسی ویژگی‌های مختلف یک خبر مانند سبک نگارش، منابع خبری، تاریخ انتشار و واکنش کاربران، احتمال جعلی بودن آن را ارزیابی کنند. روش‌های مبتنی بر نمودار نیز به‌عنوان ابزاری قدرتمند برای تحلیل شبکه‌های پیچیده و روابط بین موجودیت‌های مختلف مورد استفاده قرار می‌گیرند. در زمینه تشخیص اخبار جعلی، این روش‌ها می‌توانند برای مدل‌سازی روابط بین اخبار، منابع خبری، کاربران و سایر عوامل مرتبط مورد استفاده قرار گیرند. با استفاده از این روش‌ها می‌توان به درک بهتری از نحوه انتشار اخبار جعلی و شناسایی الگوهای انتشار این اخبار دست یافت. اگرچه مدل‌های یادگیری ماشین و یادگیری عمیق پیشرفت‌های چشمگیری داشته‌اند، اما هنوز هم به تنهایی قادر به تشخیص دقیق همه انواع اخبار جعلی نیستند. اخبار جعلی به‌طور مداوم در حال تغییر و تکامل هستند و روش‌های جدیدی

برای تولید و انتشار آن‌ها به کار گرفته می‌شود. بنابراین، ترکیب توانایی‌های انسان در درک مفاهیم پیچیده و قضاوت‌های ارزشی با توانایی‌های ماشین در پردازش حجم عظیمی از داده‌ها و کشف الگوهای پیچیده، می‌تواند به نتایج بهتری در تشخیص اخبار جعلی منجر شود.

از این‌رو در این پژوهش که با هدف بررسی فنون مقابله و نبرد هوشمندانه با اخبار جعلی در عصر هوش‌مصنوعی انجام شد، نتایج حاصل از روش تحقیق و تحلیل انجام شده در این پژوهش در قالب مدل مفهومی ترکیبی فنون مقابله و نبرد هوشمندانه با اخبار جعلی در عصر هوش‌مصنوعی به شرح شکل ۴ ترسیم و ارائه شده است:



شکل-۴. مدل مفهومی ترکیبی فنون مقابله و نبرد هوشمندانه با اخبار جعلی در عصر هوش‌مصنوعی

در این مقاله، ما با بررسی گسترده وجود انواع مختلف اخبار جعلی، مروری بر رویکردهای انتخابی انجام دادیم و در مورد کارایی متنوع و واقعی آن‌ها توضیح دادیم. این مطالعه تا حدودی ماهیت پویا و پیچیده اخبار جعلی را آشکار کرده است که بحث در مورد تشخیص آن را کاملاً چالش برانگیز می‌کند. از این رو، این وظیفه همه ذینفعان زیست‌بوم اطلاعاتی کشور چون خدمات دهندگان و کسب‌وکارهای اینترنتی، خبرگزاری‌ها، دولت، سیاست مداران و دیگران است که گامی جمعی برای توسعه فنون مقابله و نبرد هوشمندانه با اخبار جعلی در عصر هوش‌مصنوعی بردارند.

پیشنهاد

۱. تکثیر حساب‌های ربات‌های اجتماعی مخرب و اوباش اینترنتی چالش واقعی را در ایجاد کرده‌اند. آن‌ها به راحتی اخبار جعلی را در یک بازه زمانی کوتاه در سراسر جهان پخش

می‌کنند، از این‌رو، ضروری است تا برای شناسایی ربات‌های اجتماعی تحقیقاتی در دستور کار محققان این حوزه قرار گیرد.

۲. به‌منظور مبارزه با اطلاعات نادرست، باید برنامه و بودجه‌ای مشخص برای دانشگاه‌ها، مؤسسات تحقیقاتی و سایر محققان و نهادهای دانش‌بنیان برای طراحی، تولید و توسعه سامانه‌های تخصصی این حوزه تأمین و اختصاص داده شود.

۳. به دلیل نارسایی‌های حقیقت‌سنجی مبتنی بر متخصص و مبتنی بر جمع‌سپاری، رویکرد جدید جمع‌سپاری متخصص در تشخیص دستی اخبار جعلی پیشنهاد می‌گردد. از این‌رو نیاز است تا با بررسی جنبه‌های مختلف آن چالش‌ها و فرصت‌های آن شناسایی و تحلیل گردد.

تعارض منافع

بنا بر اظهار نویسندگان، مقاله پیش‌رو فاقد هر گونه تعارض منافع بوده است.

Translated References to English

Abishethvarman, V., Banujan, K., Kumara, B.T.G.S., Ravikumar, N. (2024). Unmasking Fake News with Emotional Cues: A Comparative Study of Sentiment Reasoning Models. *In 2024 4th International Conference on Advanced Research in Computing (ICARC)*. IEEE.

Akhgari, M.R., Momtazi, S. (1402). Application of Artificial Intelligence in News Verification: Detecting Fake News Using News Text and Information from News Publishers. *Media and Communication Research*, 1(1), 243-268. [**In Persian**]

Akhtar, M.M., Masood, R., Ikram, M., Kanhere, S.S. (2024). SoK: False Information, Bots and Malicious Campaigns: Demystifying Elements of Social Media Manipulations. *In Proceedings of the 19th ACM Asia Conference on Computer and Communications Security*.

Alaa Safaa M., Shati, N.M. (2024). A Survey on Fake News Detection in Social Media Using Graph Neural Networks. *Journal of Al-Qadisiyah for Computer Science and Mathematics*, 16(2), 23-41.

Alghamdi, J., Lin, Y., Luo, S. (2024). Enhancing hierarchical attention networks with CNN and stylistic features for fake news detection. *Expert Systems with Applications*, 125024.

Ali, A.M., Ghaleb, F.A., Mohammed, M.S., Alsolami, F.J., Khan, A.I. (2023). Web-informed-augmented fake news detection model using stacked layers of convolutional neural network and deep autoencoder. *Mathematics*, *11*(9), 1992.

Ayetiran, E.F., Özgöbek, Ö. (2024). A review of deep learning techniques for multimodal fake news and harmful languages detection. *IEEE Access*.

Baguian, H., Ashley, H.N. (2024). JOKER Track CLEF 2024: the Jokesters' approaches for retrieving, classifying, and translating wordplay. In Working Notes of the Conference and Labs of the Evaluation Forum (CLEF 2024). *CEUR Workshop Proceedings*, pp. 1811-1817).

Chan, J. (2024). Online astroturfing: A problem beyond disinformation. *Philosophy & Social Criticism*, 50(3), 507-528.

Collins, B., Hoang, D.T., Nguyen, N.T., Hwang, D. (2021). Trends in combating fake news on social media–a survey. *Journal of Information and Telecommunication*, *5*(2).

Dong, X., Sarker, S., Qian, L. (2022, September). Integrating human-in-the-loop into swarm learning for decentralized fake news detection. *In 2022 International Conference on Intelligent Data Science Technologies and Applications (IDSTA)*. IEEE.

Emami Rudsari, Hossein. (2023). Methods and tricks of dealing with fake news. *Media and Communication Research*, 1(1). [**In Persian**]

Godel, W., Sanderson, Z., Aslett, K., Nagler, J., Bonneau, R., Persily, N., & Tucker, J. A. (2021). Moderating with the mob: Evaluating the efficacy of real-time crowdsourced fact-checking. *Journal of Online Trust and Safety*, 1**(**1).

Ghosh, S., Mitra, P., Nakov, P. (2024, May). Clock against Chaos: Dynamic Assessment and Temporal Intervention in Reducing Misinformation Propagation. In *Proceedings of the International AAAI Conference on Web and Social Media*. 1(18) 462-473.

Jain, M. K., Gopalani, D., Meena, Y. K. (2024). ConFake: fake news identification using content-based features. *Multimedia Tools and Applications*, 83(3).

Kitsa, M. (2023). Classification of Fake News in Ukraine and Abroad. *State and Regions. Series: Social Communications*, 1 (53).

Lakzaei, B., Haghir Chehreghani, M., Bagheri, A. (2024). Disinformation detection using graph neural networks: a survey. *Artificial Intelligence Review*, 57(3), 52.

La Barbera, D., Maddalena, E., Soprano, M., Roitero, K., Demartini, G., Ceolin, D., ... & Mizzaro, S. (2024). Crowdsourced Fact-checking: Does It Actually Work? *Information Processing & Management*, 61(5), 103792.

Liu, Z., Zhang, T., Yang, K., Thompson, P., Yu, Z., Ananiadou, S. (2024). Emotion detection for misinformation: A review. *Information Fusion*, 102300.

Lybarger, K., Dobbins, N.J., Long, R., Singh, A., Wedgeworth, P., Uzuner, Ö., Yetisgen, M. (2023). Leveraging natural language processing to augment structured social determinants of health data in the electronic health record. *Journal of the American Medical Informatics Association*, 30(8).

Madani, M., Motameni, H., Roshani, R. (2024). Fake news detection using feature extraction, natural language processing, curriculum learning, and deep learning. *International Journal of Information Technology & Decision Making*, *23*(03), 1063-1098.

Mahdi, A. S., Shati, N. M. (2024). A Survey on Fake News Detection in social media Using Graph Neural Networks. *Journal of Al-Qadisiyah for Computer Science and Mathematics*, 16(2), 23-41.

Mittal, T., Chowdhury, S., Guhan, P., Chelluri, S., Manocha, D. (2024). Towards determining perceived audience intent for multimodal social media posts using the theory of reasoned action. *Scientific Reports*, *14*(1), 10606.

Mousavi Loghman, S.A, Mousavi, S.S., Davoodi Zadeh, L., Pishvaei, M. (2023). Presenting a conceptual model of "family productivity" with emphasis on the economic aspect: a content analysis of social welfare, *Quarterly Scientific Research Journal of Social Welfare*, 23 (91): 323-364. [**In Persian**]

Nasery, M. (2024). *Fake News on social media: From Fake News Lifecycle to Fake News Combat Cycle* (Doctoral dissertation).

Pourkarimi, J., Azizi, M. (2024). Leaders' Competencies in Turbulent Environments (A Meta-Synthesis Study). *Public Administration*, 16(3), [**In Persian**]

Rastogi, S., Bansal, D. (2023). A review on fake news detection 3T's: typology, time of detection, taxonomies. *International Journal of Information Security*, *22*(1).

Sahoo, S.R., Gupta, B.B. (2021). Multiple features-based approach for automatic fake news detection on social networks using deep learning. *Applied Soft Computing*, *100*, 106983.

Shahid, W., Li, Y., Staples, D., Amin, G., Hakak, S., Ghorbani, A. (2022). Are you a cyborg, bot or human? a survey on detecting fake news spreaders. *IEEE Access*, *10*, 27069-27083.

Shrestha, A., Flood, A., Sohrawardi, S., Wright, M., Al-Ameen, M. N. (2024, May). A First Look into Targeted Clickbait and its Countermeasures: The Power of Storytelling. *In Proceedings of the CHI Conference on Human Factors in Computing Systems* (pp. 1-23).

Tarassoli, A.R. (2024). Application of Artificial Intelligence in Detecting Fake News. Quarterly Journal of Defense Preparedness and Technology, 7(2), 82-112. [**In Persian**]

Tikkanen, R. (2024). Intercepted Phone Calls at the Russo-Ukrainian War: Cyberoperation or Propaganda Campaign? *Master's Degree Programme in Information Technology, Cyber Security.*

Ventre, D. (2020). *Artificial Intelligence, Cybersecurity and Cyber Defense*. y ISTE Ltd and John Wiley & Sons, Inc.

Yildirim, G. (2023). A novel hybrid multi-thread metaheuristic approach for fake news detection in social media. *Applied Intelligence*, 53(9), 11182-11202.