

شبکه‌های اجتماعی در خدمت نوآوری و کارآفرینی اجتماعی (مطالعه موردی: توییتر)

علی اصغر سعدآبادی^۱، فراز مهر آیین^۲، زهره رحیمی راد^۳

تاریخ دریافت: ۱۴۰۰/۵/۱۵ تاریخ پذیرش: ۱۴۰۰/۶/۲۱

چکیده

امروزه، کاربران بسیاری در شبکه‌های اجتماعی مختلف، عقاید و نظرات خود را با سایرین به اشتراک می‌گذارند. حجم انبوه داده‌های خام تولیدشده در شبکه‌های اجتماعی، آنها را به مخزنی سرشار از افکار، ایده‌ها و دغدغه‌های تعداد بی‌شماری از کاربران تبدیل کرده که توجه به آنها می‌تواند در تصمیم‌گیری‌های خرد و کلان، نقشی تعیین‌کننده داشته باشد. این پژوهش درصدد است با استفاده از روش‌های متن‌کاوی در توییتر، به شناسایی دغدغه‌های اجتماعی افراد جامعه بپردازد تا در طرح‌های آتی نوآوری و کارآفرینی اجتماعی، از آنها استفاده شود. برای دستیابی به این هدف، توییت‌های مرتبط با نوآوری اجتماعی و کارآفرینی اجتماعی، با به‌کارگیری روشی نوین، از توییتر، استخراج و با استفاده از تکنیک‌ها و الگوریتم‌های متن‌کاوی، از جمله: مدل‌سازی موضوعی LDA و تحلیل احساسات، بررسی شدند. یافته‌های متن‌کاوی نشان می‌دهد کاربران، در حوزه کارآفرینی اجتماعی، بیشتر به موضوعاتی مانند: لزوم توجه به مسئله اشتغال در چابهار، برگزاری رویدادهای رایگان به‌منظور شناسایی ایده‌ها، فرصت‌ها و مشکلات جذب سرمایه در پروژه‌های کارآفرینی پرداخته‌اند. در حوزه نوآوری اجتماعی نیز، کاربران به موضوعاتی همچون: لزوم توجه به معلولین، ساخت مراکز دندان‌پزشکی، تأسیس گلخانه در مناطق جنوبی کشور و توانمندسازی بانوان، توجه بیشتری نشان داده‌اند. علاوه بر این، تحلیل احساسات توییت‌های فارسی درباره کارآفرینی و نوآوری اجتماعی، حاکی از مثبت بودن نتایج است. با توجه به نتایج به‌دست‌آمده، می‌توان در برنامه‌ریزی‌ها و تصمیم‌گیری‌های آینده، طرح‌هایی را برای رفع این دغدغه‌ها در جامعه اجرا کرد که در نهایت، این امر می‌تواند به ارتقای میزان رضایتمندی مردم و سطح رفاه آنها بینجامد.

واژه‌های کلیدی

کارآفرینی اجتماعی، داده‌کاوی، توییتر، نوآوری اجتماعی، عقیده‌کاوی.

* این مقاله بر اساس نظر گروه دبیران و سردبیر فصلنامه، پژوهشی است.

۱. استادیار گروه سیاست‌گذاری علم و فناوری پژوهشکده مطالعات بنیادین علم و فناوری دانشگاه شهید بهشتی، تهران، ایران
a_sadabadi@sbu.ac.ir
۲. کارشناسی ارشد مدیریت تکنولوژی، گروه مدیریت، دانشکده مدیریت و حسابداری، دانشگاه شهید بهشتی، تهران، ایران
azarinmehr.pub@gmail.com
۳. دکتری سیاست‌گذاری علم و فناوری، گروه مدیریت صنعتی، دانشکده علوم اقتصادی و اداری، دانشگاه مازندران، مازندران، بابلسر
zrahimi.rad@gmail.com

مقدمه

در سال‌های اخیر، مفهوم نوآوری اجتماعی در بین دانشگاهیان فراگیر شده است، هر چند تعریف آن مبهم است (ماکیس^۱ و همکاران، ۲۰۱۸). مبنای ما در این مقاله، تعریفی است که مک کالوم و همکاران، از نوآوری اجتماعی ارائه داده‌اند (مک کالوم و همکاران، ۲۰۰۹)؛ یعنی نوآوری اجتماعی به‌عنوان یک رابطه اجتماعی، کارگزار جمعی و توانمندساز. در همین راستا، موضوع پرطرفدار نوآوری، صرفاً به‌عنوان تلاشی اجتماعی برای افزایش رفاه جامعه تعریف می‌شود که برای مشخص شدن بهتر مفهوم آن، باید معانی و ارتباط بین فعالیت‌های اجتماعی، نوآوری اجتماعی و رفاه اجتماعی به‌خوبی ارزیابی شود (هوالدت و شوارتز^۲، ۲۰۱۷).

البته تلاش‌هایی برای ارائه تعریفی جامع در ادبیات نیز توسط برخی از پژوهشگران انجام شده است؛ واکاوی تعریف‌های موجود و ارائه تعریفی واضح و یکپارچه از نوآوری اجتماعی، امری ضروری به‌حساب می‌آید. تعریف جامع و واضح می‌تواند به برداشت و درک مشترک و منطقی این موضوع کمک کند (خدایاری و نیل‌فروشان، ۱۳۹۸).

با وجود تلاش‌ها و پژوهش‌هایی که در ادبیات، در ارتباط با نوآوری اجتماعی انجام شده، هنوز تعریفی جامع برای نوآوری اجتماعی ارائه نشده است. از طرفی، امروزه شبکه‌های اجتماعی، کاربران زیادی دارند که به‌صورت روزانه، در آنها به فعالیت می‌پردازند. این شبکه‌ها، برای تأثیرگذاری بر فرایند یادگیری، فرصت‌های حل مسئله و پیاده‌سازی ایده‌ها، پتانسیل زیادی دارند. علاوه بر این، می‌توان با کمک گرفتن از این شبکه‌ها- به‌عنوان عاملی مهم در شکل‌گیری ایده‌ها- به‌خوبی، نوآوری را تشخیص داد. از این رو، شبکه‌های اجتماعی می‌توانند باعث ارتقای فرایندهای نوآوری شده و فرصت‌های بیشتری را برای یادگیری فراهم کنند (کولک^۳، ۲۰۱۳).

به همین دلیل است که شبکه‌های اجتماعی برای مدیران کسب‌وکارهای

1. Marques
2. Howaldt & Schwarz
3. Kolleck

اجتماعی، به ابزاری مهم در راستای مدیریت روابط آنها با عموم مردم تبدیل شده‌اند. همچنین این شبکه‌ها به کانالی محبوب برای مردمی تبدیل شده‌اند که دوست دارند ایده‌های خود را به اشتراک گذاشته و روابطی جدید را شکل دهند. در این میان، محتواهای تولیدشده توسط کاربر، منبعی بسیار مناسب برای بررسی پدیده‌های اجتماعی است. با توجه به این مسئله، داده‌کاوی شبکه‌های اجتماعی، راهی مؤثر برای بررسی الگوها و نظرات عموم مردم نسبت به موضوعی خاص است (شن و هو، ۲۰۱۸). نکته دیگر آن که امروزه شبکه‌های اجتماعی به دلیل نقش مهمی که در ارتباطات میان‌فردی دارند، بیشتر مورد توجه پژوهشگران قرار گرفته‌اند. افزایش شمار کاربران و پست‌های آنان، باعث شده استفاده از سیستم‌های پیشنهادگر در این شبکه‌ها معمول شود؛ سیستم‌های پیشنهادگر، موضوعات مفید و موردعلاقه کاربران را به‌آسانی در دسترس آنها قرار می‌دهند (میرزاآقازاده و همکاران، ۱۳۹۴).

یکی از شبکه‌های اجتماعی پرطرفدار، توئیتر است که افراد در آن، به‌طور روزانه، دیگر کاربران را از اتفاقات روزمره خود آگاه می‌کنند و اطلاعاتی همچون تجربیات خود را با آنها به اشتراک می‌گذارند (نظری و همکاران، ۱۳۹۸). این شبکه، یکی از بزرگ‌ترین و محبوب‌ترین وبسایت‌های میکرو بلاگینگ^۱ است و کاربران، از آن برای بیان نظرات خود استفاده می‌کنند، از این رو، اگر آنالیز خوبی بر روی داده‌های آن انجام گیرد، اطلاعاتی مفید برای تصمیم‌گیری به دست می‌آید (پرهانتو و همکاران^۳، ۲۰۲۰). همچنین توئیتر، مهم‌ترین رسانه اجتماعی است که پژوهشگران از آن برای به اشتراک‌گذاری برون‌دادهای پژوهشی خود در سطح بین‌المللی استفاده می‌کنند و قابلیت‌های فراوانی را برای مطالعه اثرگذاری اجتماعی پژوهش دارد. برخلاف آنچه تصور می‌شود، توئیتر برای محتواهای علمی نیز مناسب است و توئیتهای زیادی در مورد مقالات علمی در آن، بین کاربران به اشتراک گذاشته می‌شود (عرفان‌منش، ۱۳۹۷). اولوا و همکاران^۴ (۲۰۱۶) بیان می‌کنند که امروزه

1 - Shen & Ho

2 - Microblogging : به اشتراک‌گذاری مطالب و به‌روزرسانی‌های کوتاه

3 - Prehanto

4 - Ulloa

تویبیتر، به ابزاری برای اندازه‌گیری نبض بشریت تبدیل شده است؛ بلافاصله صدها جنبش اجتماعی جهانی، سازمان‌دهی و از طریق این شبکه اجتماعی به اشتراک گذاشته می‌شود و بر درک موضوع موردبحث تأثیر می‌گذارد. با توجه به این موارد، می‌توان گفت تویبیتر، برای بررسی پدیده‌های اجتماعی مانند: نوآوری اجتماعی و کارآفرینی اجتماعی، فضای مناسبی است.

از آن جا که نوآوری اجتماعی و کارآفرینی اجتماعی، هر دو، بیشتر پدیده‌های اجتماعی به حساب می‌آیند، باید در تصمیم‌گیری‌ها، به نظر و عقیده مردم توجه کرد تا بهتر بتوان معضلات اجتماعی را شناخت. در واقع، اگر تصمیم‌گیرندگان، در تصمیم‌گیری‌های خود، به نظرات مردم توجه نداشته باشند، ممکن است به درستی، معضلات اجتماعی را شناسایی نکنند. با این مقدمه، می‌توان گفت در دنیای امروز، لازم است کسب‌وکارها اطلاعات خود را از شبکه‌های اجتماعی دریافت کنند. علاوه بر این، با توجه به اینکه شبکه‌های اجتماعی، محل نشر نظرات افراد مختلف جامعه هستند، می‌توان با استفاده از آنها، اطلاعات مناسبی را برای تصمیم‌گیری‌ها استخراج کرد. با در نظر گرفتن موارد گفته‌شده، در این پژوهش به کاوش نظرات مردم بیرامون دو پدیده نوآوری اجتماعی و کارآفرینی اجتماعی، با استفاده از داده‌های تویبیتر پرداخته‌ایم تا بتوانیم احساسات مردمی و موارد نهفته در تویبیت‌هایی با موضوعات مطرح‌شده را استخراج کنیم.

مرور ادبیات

نوآوری اجتماعی را می‌توان راه‌حلی نوین و مردمی برای بسیاری از معضلات اجتماعی دانست، امروزه، نمونه‌های بارزی از این پدیده را در جامعه شاهد هستیم، مانند: نقش نوآوری اجتماعی در طرح‌های عدالت غذایی و... به باور بسیاری از محققان، در شرایطی که جوامع با چالش‌های جدی اجتماعی روبه‌رو هستند، نوآوری اجتماعی می‌تواند با درگیر کردن همه افراد، راه‌حلی نوآورانه را برای حل معضلات ارائه دهد (کشتکار، ۱۳۹۷). همچنین، این پدیده می‌تواند علاوه بر به‌کارگیری منابع اجتماعی، بازیگران دیگر را برای پذیرش راهکار مجاب کند، به طوری که راهکار، به بخشی از جامعه تبدیل شود. هدف نوآوری اجتماعی این است که ارزشی را که می‌تواند برای جامعه خلق کند، به حداکثر برساند (سعدآبادی و همکاران، ۱۳۹۵).

تعاریف نوآوری اجتماعی از دیدگاه محققان مختلف، در جدول شماره ۱، قابل مشاهده است.

جدول شماره ۱. تعاریف مختلف نوآوری اجتماعی (چیتسازان و صادقی‌کیا، ۱۳۹۷)

منبع	تعریف نوآوری اجتماعی
Murray et al., 2010	ایده‌های جدید محصولات، خدمات، یا مدل‌ها هستند که به صورت هم‌زمان، نیازهای اجتماعی را برطرف می‌کنند و روابط یا همکاری‌های اجتماعی جدیدی را خلق می‌کنند.
Bacon et al, 2008	به ایده‌ها، نهادها یا شیوه‌های جدیدی اشاره دارد که هدف‌شان، رسیدگی به مسائل اجتماعی است، بنابراین، ممکن است شامل شیوه‌های جدید فعالیت برای کاهش فقر و تبعیض، یا خدمات و سازمان‌های جدید برای مراقبت از بیماران باشد.
Social Innovation Europe, 2012	ایده‌ها، نهادها یا شیوه‌های کاری جدید هستند که در مقایسه با رویکردهای موجود، به شکلی مؤثرتر، به نیازهای اجتماعی رسیدگی می‌کنند.
OECD, 2010	با به کارگیری مشارکت افراد، برای بهبود رفاه اجتماعی اقدام می‌کند و هدف آن، ارائه راه‌حل برای مسائل فردی و اجتماعی است.
Joshi, 2010	نوآوری پایدار یا فراگیر یا نوآوری اجتماعی، نوآوری‌هایی هستند که برای کسب‌وکار، مشتری، محل یا اجتماع، ارزش افزوده ایجاد کنند.
Pol & Ville, 2009	یک نوآوری، هنگامی نوآوری اجتماعی نامیده می‌شود که پتانسیل بهبود کمیت یا کیفیت زندگی مردم را داشته باشد.
Bisgaard, 2009	محصولات جدید، خدمات جدید، مدل‌های جدید کسب‌وکار، فرایندهای جدید کانال‌های توزیع جدید و غیره، آن‌گاه که به صورت هم‌زمان، به حل چالش‌های جهانی مربوط به موضوعات محیطی و اجتماعی می‌پردازند.
Phills , et al 2008	راه‌حلی جدید برای مشکلات اجتماعی که نسبت به راه‌حل‌های موجود، مؤثرتر، کارآمدتر، پایدارتر یا عادلانه‌تر است و ارزش ایجادشده ناشی از آن، در درجه اول، به جای افراد، به جامعه- به عنوان یک کل- تعلق می‌گیرد.
Mulgan, 2006	به فعالیت‌ها و خدمات نوآورانه‌ای اشاره دارد که هدف‌شان، رسیدگی به نیازهای اجتماعی است و عمدتاً نیز توسط سازمان‌هایی که اهداف اجتماعی برای‌شان اولویت دارد، انجام می‌شوند.
TEPSIE, 2014	به عنوان رویکردهای جدید در پرداختن به نیازهای اجتماعی تعریف می‌شود. نوآوری‌های اجتماعی، در اهداف و ابزارهای‌شان اجتماعی هستند؛ آنها با به کارگیری و تسهیل دسترسی ذی‌نفعان به منابع و قدرت، به تغییر روابط اجتماعی کمک می‌کنند.

نوآوری اجتماعی در ادبیات، موضوعی بسیار محبوب و پرکاربرد است و محققان زیادی از جنبه‌های مختلف، آن را بررسی کرده‌اند، با وجود این، در همه قسمت‌های ادبیات نیز، تعریفی جامع درمورد آن ارائه نشده است.

نوآوری اجتماعی می‌تواند به شیوه‌های گوناگون، تأثیرگذار باشد، به‌عنوان مثال، نوآوری اجتماعی در ترکیب با فناوری دیجیتال می‌تواند در بحث بررسی‌های برنامه‌ریزی فضایی، موردبررسی قرار گیرد. اهمیت و محبوبیت نوآوری اجتماعی در سال‌های اخیر، به‌دلیل چالش‌های مالی، اجتماعی و همچنین پیشرفت تکنولوژی بیشتر شده است. نوآوری اجتماعی دیجیتال، امروزه بیش از پیش در برنامه‌های شهرهای هوشمند وارد شده است. کارکرد و مزایای نوآوری اجتماعی دیجیتال، در برنامه‌ریزی فضا سازی مشهود است. به‌عنوان مثال، برنامه‌ریزی مبتنی بر همکاری و راه‌حل‌های مختلف، از اثرات نوآوری اجتماعی دیجیتال است (آنچلیدو و پسالوتوگوا، 2018).

نوآوری اجتماعی در سیستم‌های غذایی نیز اثراتی قابل‌توجه داشته و اثر آن در گذار به سیستم‌های غذایی پایدارتر، کاملاً مشهود است؛ به‌طوری‌که نوآوری اجتماعی در میان مردم برای رسیدن به ظرفیت بیشتر به کار گرفته شده است، به‌عنوان مثال، در زنجیره ارزش نان و گندم در توسکانی ایتالیا، درمورد ارتباط بین ذات نوآوری اجتماعی در سطح ابتکارات مردمی که در جست‌وجوی تغییرات انقلابی، با هدف تعریف مجدد فرایند تولید و مصرف (از طریق مداخلات اجتماعی برای برآورده کردن نیازهای اجتماعی) هستند، بینش‌هایی اضافه شده است (روسی و بوچی، ۲۰۱۸).

پیاده‌سازی طرح‌های نوآوری اجتماعی همیشه آسان نیست و در بعضی از مواقع، با موانع متعددی روبه‌رو می‌شود که موجب به‌تعویق افتادن پیاده‌سازی این طرح‌ها یا عدم‌اجرای آنها می‌شود و یا باعث می‌شود در صورت اجرا، نتیجه مطلوب پیش‌بینی‌شده را به همراه نداشته باشند. در راستای بررسی موانع نوآوری اجتماعی به‌منظور ارائه نوآوری‌های جدید برای حل مشکلات اجتماعی، همواره چالش‌هایی وجود دارد. همچنین ایجاد محیطی مناسب برای نوآوری، می‌تواند فرایند پیاده‌سازی نوآوری را آسان‌تر کند. استفاده از پنج مارک ۳ یا مطالعه تاریخی کشورهای خارجی می‌تواند روشی مناسب برای توسعه نوآوری اجتماعی باشد.

به‌عنوان مثال، مطالعات نشان می‌دهد روسیه نیازمند توسعه مشارکت‌های عمومی- خصوصی در اجرای نوآوری اجتماعی است؛ اما این اتفاق نیفتاده است. اجرای نوآوری اجتماعی به افزایش آگاهی مردم نیز نیاز دارد، زیرا تمرکز نوآوری اجتماعی بر مشارکت مردم در حل مشکلات است. همچنین استفاده از فناوری‌های ICT و میانجی‌گری نوآوری، باید جزء مکانیسم نوآوری اجتماعی باشد.

کارآفرینی اجتماعی

از آن جا که نوآوری اجتماعی و کارآفرینی اجتماعی، اهدافی مشترک دارند، اشتراکات زیادی با یکدیگر نیز دارند. کارآفرینی اجتماعی، به‌عنوان یک مفهوم و اصطلاح، در اواخر دهه ۱۹۹۰ مطرح شد. کارآفرینان اجتماعی و بنگاه‌های اجتماعی، بازیگرانی هستند که از طریق نوآوری، تغییر اجتماعی ایجاد می‌کنند. ورود اصطلاح «نوآوری اجتماعی» به صحنه، آن هم با محبوبیتی زیاد، استفاده از کارآفرینی اجتماعی را تا حدی به عقب رانده و بیانگر این درک است که کارآفرینان اجتماعی، تنها بازیگرانی نیستند که نوآوری اجتماعی ایجاد می‌کنند. امروزه به توسعه کارآفرینی اجتماعی، به‌عنوان مأموریتی اجتماعی و انسان‌دوستانه و یکی از راهکارهای حل برخی از مسائل اجتماعی توجه می‌شود. کارآفرینی، بر بسیاری از فعالیت‌های کشورها، تأثیری مثبت و مستقیم دارد؛ از یک طرف، باعث توسعه اقتصادی و از طرف دیگر، موجب ایجاد فرصت شغلی می‌شود (مرجانیم و صدری، ۱۳۹۲). کارآفرینی اجتماعی به‌منزله ایجاد راه‌حل‌های نوآورانه به‌منظور رفع مشکلات اجتماعی و بسیج ایده‌ها، ظرفیت‌ها و منابع موردنیاز برای تغییرات اجتماعی است (نداف^۱ و همکاران، ۲۰۱۸).

نوآوری اجتماعی، عاملی مکمل و اصلاح‌گر در آغاز طرح‌های دولتی است. علاوه بر این، ممکن است از ابتکارات محلی شکل بگیرد، به‌گونه‌ای که افراد سعی می‌کنند مشکلات محله خود را حل کنند. اتحادیه اروپا- که اخیراً از فعالیت‌های مختلف نوآوری اجتماعی حمایت می‌کند- نوآوری اجتماعی را مجموعه‌ای از ایده‌ها، خدمات و محصولات می‌داند که هدف آنها، برآورده کردن نیازهای اجتماعی است. نوآوری اجتماعی ممکن است در همه بخش‌های اقتصاد رخ دهد و فقط محدود

1 - Nadaf

شبکه‌های اجتماعی در خدمت نوآوری ...

به بخش‌های غیرانتفاعی نیست.

از آن جا که نوآوری اجتماعی بر پایهٔ ابتکارات بخش خصوصی است، توجهات زیادی را به خود جلب کرده که می‌تواند مقدمات و ترتیبات دولتی را تغییر دهد. کارآفرینی اجتماعی نیز در بخش‌هایی با نوآوری اجتماعی همکاری می‌کند، همان‌طور که ممکن است کارآفرین‌های عادی، به‌طریقی به نوآوری اجتماعی کمک کنند. درواقع، فعالیت‌های کارآفرینی اجتماعی ممکن است آگاهانه یا ناآگاهانه منجر به نوآوری اجتماعی شوند. کارآفرینان اجتماعی، تولیدکننده‌های ارزش اجتماعی به حساب می‌آیند. درضمن، کارآفرینان اجتماعی باید نسبت به کارآفرین‌های معمولی اثرات اجتماعی بیشتری داشته باشند.

نوآوری اجتماعی ممکن است محصول کارآفرینی عادی نیز باشد. درواقع، نوآوری اجتماعی به کارآفرینی اجتماعی نیاز ندارد، چون منابع دیگری نیز برای نوآوری اجتماعی وجود دارد. بنابراین در بخش خصوصی، نوآوری اجتماعی نیازی به کارآفرینی اجتماعی ندارد، اما مانند هر نوع نوآوری دیگر، به کارآفرینی نیاز دارد (گروت و دانکبار،^۱ ۲۰۱۴).

رسانه‌های اجتماعی و داده‌کاوی

رسانه‌های سنتی مانند: روزنامه، تلویزیون و رادیو، ابزارهایی مؤثر برای انتقال اطلاعات بوده‌اند، هرچند در این روش‌ها، ارتباط، یک‌طرفه بوده و درنتیجه، امکان ثبت بازخورد و نظر در آنها وجود نداشته و اینک، با روی کارآمدن شبکه‌های اجتماعی، این مشکل و خلأ حل شده است (سیندانی^۲ و همکاران، ۲۰۱۹). حدود ۵/۲ میلیارد نفر در جهان از رسانه‌های اجتماعی استفاده می‌کنند. از پیشگامان این عرصه می‌توان فیسبوک^۳، اینستاگرام^۴، تویتر^۵ و... را نام برد؛ در تویتر به‌تنهایی، در هر روز، تقریباً ۵۰۰ میلیون پست به اشتراک گذاشته می‌شود (مهناز

1 - Groot & Dankbaar

2 - Sindhani

3 - Facebook

4 - Instagram

5 - Twitter

عباسی و همکاران، ۱۳۹۷). محتوای تولیدشده توسط کاربران^۱ یا همان UGCها می‌توانند شامل نظرات، عقاید و توصیفات کاربران باشند که به‌صورت محتوا یا یکدیگر به اشتراک می‌گذارند. همچنین مطالعه این‌گونه محتواها می‌تواند برای بوم‌کسب‌وکارها و مدیران بسیار مهم باشد، زیرا تجزیه و تحلیل این اطلاعات می‌تواند بینش خوبی به مدیران بدهد و حتی باعث تعریف مجدد بعضی از استراتژی‌ها شود. علاوه بر این، رشد UGCها و توسعه بعضی از تکنولوژی‌ها، مانند: داده‌های کلان یا داده‌کاوی و یادگیری ماشین، باعث ایجاد روش‌های متعدد بهینه‌سازی اطلاعات و استخراج اطلاعات معنادار از UGCها شده است، به همین دلیل، داده‌کاوی می‌تواند روشی مناسب برای تجزیه و تحلیل محتواهای تولیدشده توسط کاربران در شبکه‌های اجتماعی باشد (ساورا و بنت^۲، ۲۰۱۹).

در بسیاری از مواقع، بیشتر افراد، تفاوت چندانی بین داده‌کاوی و متن‌کاوی^۳ قائل نمی‌شوند و این دو را مشابه و زیرمجموعه یکدیگر می‌دانند. درحالی‌که این دو، تفاوت‌های کم، ولی مهمی با یکدیگر دارند. در تفاوت داده‌کاوی و متن‌کاوی می‌توان گفت داده‌کاوی شامل روش‌هایی برای اکتشاف از انواع داده‌ها مانند: داده‌های عددی، چندرسانه‌ای و... است، الگوریتم‌های عمومی داده‌کاوی می‌توانند تمام انواع داده را پردازش کنند، اما متن‌کاوی به‌عنوان زیرمجموعه داده‌کاوی عمومی در نظر گرفته می‌شود. تفاوت اصلی میان داده‌کاوی و متن‌کاوی، همان نوع داده و الگوریتم‌های مربوط به آنهاست. همچنین داده‌کاوی را می‌توان به‌عنوان زمینه‌ای از تحلیل داده‌های کمی (عددی) و متن‌کاوی را زمینه‌ای از داده‌های کیفی (متن) در نظر گرفت (عباسی: ۱۳۹۵). تفاوت اصلی داده‌کاوی و متن‌کاوی، در ویژگی استخراج یا الگوهای استفاده‌شده در کار با داده‌هاست؛ در داده‌کاوی، استخراج ویژگی، از داده‌های ساختاریافته می‌آید، درحالی‌که در متن‌کاوی، از نیمه ساختاریافته یا ساختاریافته می‌آید، اگرچه بیشتر داده‌ها به‌صورت غیرساختاریافته است. علاوه بر این، باید به این نکته نیز توجه کرد که مرحله پیش‌پردازش در متن‌کاوی،

1 - UGC : User Generated Content

3 - Text Mining

به دلیل آماده کردن داده‌های غیرساخت‌یافته به ساخت‌یافته است (مایواتی^۱ و همکاران، ۲۰۱۸). برای جمع‌بندی مطالب گفته‌شده، تفاوت‌های داده‌کاوی و متن‌کاوی در جدول شماره ۲ درج شده است:

جدول شماره ۲. مقایسه داده‌کاوی و متن‌کاوی (عباسی : ۱۳۹۵)

عنوان	داده‌کاوی	متن‌کاوی
موضوع موردبررسی	داده‌های عددی (ساختاریافته)	متون (غیرساخت‌یافته و نیمه‌ساخت‌یافته)
ساختار موضوع	بانک‌های اطلاعاتی رابطه‌ای	متون فاقد شکل
روش‌ها	یادگیری ماشین، درخت تصمیم‌گیری، شبکه‌های عصبی، رگرسیون و...	الگوریتم‌های خوشه‌بندی طبقه‌بندی، زبان‌شناسی، هسته‌شناسی و...
دوره توسعه	پیاده‌سازی وسیع از سال ۱۹۹۴	پیاده‌سازی وسیع از سال ۲۰۰۰

امروزه از داده‌کاوی و متن‌کاوی برای بهره‌مندی از داده‌های موجود در اینترنت استفاده‌های فراوانی می‌شود و به کمک این دو می‌توان از داده‌های حجیم و بسیار زیاد موجود در اینترنت- که در نگاه اول، بی‌استفاده به نظر می‌رسند- اطلاعاتی معنادار را استخراج کرد که می‌تواند به تصمیم‌گیرندگان، سیاست‌گذاران و یا حتی مدیران کسب‌وکارها بینش‌های خوبی را ارائه دهد.

یکی از کاربردهای اصلی داده‌کاوی و متن‌کاوی، تحلیل نظرات کاربران در سطح اینترنت است. اگر این‌گونه داده‌ها به‌خوبی تحلیل شده و تصمیمات درستی بر پایه آن گرفته شود، می‌توان مشارکت مردم در این‌گونه سایت‌ها را افزایش داد. یکی از سایت‌های معروف جهانی، سایت تریپ ادوایزر است که در زمینه گردشگری فعالیت می‌کند و اطلاعات مفیدی را در اختیار گردشگران قرار می‌دهد. با توجه به داده‌های هتل‌ها، رستوران‌ها و اماکن دینی شهر لیسبون، می‌توان گفت مشارکت کاربران در رسانه‌های اجتماعی، از دو منظر حائز اهمیت است؛ نخست، در جهت منفعت رسانه اجتماعی و دوم، چگونگی افزایش مشارکت کاربران در رسانه‌ها. به زبانی ساده‌تر، افزایش مشارکت کاربران در رسانه‌های اجتماعی، علاوه بر استخراج نتایج دقیق‌تر، افزایش اعتبار همان رسانه را به دنبال خواهد داشت و هرچه اعتبار رسانه بیشتر باشد، مشارکت کاربران در آن رسانه بیشتر خواهد بود (سهرابی و ریسی، ۲۰۱۶).

۱. مایواتی و داده‌کاوی برای بررسی روندها و استفاده از آنها در مطالعات،

می‌تواند در شکل‌گیری ایده‌های نوین، نقشی بسزا داشته باشد. به‌عنوان مثال، برای جهت‌دهی به صنعت فینتک کره، داده‌های ورودی Naver (معروف‌ترین سایت کره‌ای) مورد مطالعه قرار گرفته و ۲۰ کلمهٔ پرتکرار در سال ۲۰۱۵ و اوایل ۲۰۱۶، مرتبط با حوزه‌های پرداخت، بانک، پلتفرم و دیگر حوزه‌های مرتبط با فینتک، به دست آمده است. برای بررسی داده‌ها، از تحلیل احساسات، بررسی مدل‌های موضوعی و دیگر تکنیک‌های متن‌کاوی استفاده شده است (لی و همکاران، ۲۰۱۷).

امروزه شبکه‌های اجتماعی به دلیل نقش زیادی که در ارتباطات میان‌فردی دارند، بیشتر مورد توجه پژوهشگران قرار می‌گیرند. این شبکه‌ها، علاوه بر محلی برای ایجاد روابط، به رسانه‌ای محبوب در میان مردم برای به‌اشتراک‌گذاری ایده‌هایشان تبدیل شده‌اند. محتواها و نظرات مردم در شبکه‌های اجتماعی، می‌تواند منبعی مفید برای بررسی پدیده‌های اجتماعی باشد. البته این نکته نباید فراموش شود که با توجه به قوانین زبانی، کاراکترها و عبارات غیرمرتبط و زائد، باید از متون موجود در شبکه‌های اجتماعی حذف شوند و سپس با ترکیب کلمات هم‌معنی، ارتباط کلمات و ضریبی از قوی بودن ارتباط آنها به دست آید؛ اطلاعات و روابط به دست آمده می‌تواند بینشی مفید را در اختیار بازیابان یا مدیران کسب‌وکارهای اجتماعی قرار دهد (شن و هو، ۲۰۱۸).

در این میان، توییتر یکی از شبکه‌های اجتماعی محبوب شناخته می‌شود، حتی در کشور ما نیز با وجود برخی محدودیت‌ها، کاربران ایرانی زیادی از آن استفاده می‌کنند. وجود کاربران متعدد باعث می‌شود روزانه پست‌ها و اطلاعات زیادی در این شبکهٔ اجتماعی منتشر شود. از این داده‌های خام می‌توان با کمک گرفتن از داده‌کاوی و متن‌کاوی اطلاعات مفیدی را استخراج کرد.

همان‌طور که اشاره شد، برخلاف آنچه تصور می‌شود، توییتر برای محتواهای علمی نیز مناسب است، به‌گونه‌ای که طبق بررسی‌ها، بیش از ۷۸ درصد اشتراک‌گذاری برون‌داده‌های علمی در شبکه‌های اجتماعی، متعلق به توییتر بوده است. همچنین، مطالعات نشان می‌دهند روند به‌اشتراک‌گذاری مقالات علمی در

تویبیتز، به میزان چشمگیری روبه‌افزایش است. بیشتر تویبیت‌ها مربوط به کاربرانی از کشورهای آمریکا، انگلستان و استرالیا بوده است. علاوه بر اینها، نتایج پژوهشی منتشرشده در مجله‌های علوم پزشکی و علوم میان‌رشته‌ای، به میزان گسترده‌تری تویبیت شده است (عرفان‌منش، ۱۳۹۷).

مدل‌سازی موضوعی یکی از تکنیک‌های پرکاربرد متن‌کاوی و داده‌کاوی است. این تکنیک، الگوریتم‌های متفاوتی، مانند: تخصیص پنهان دیریکله (LDA) و آنالیز معنایی پنهان (LSA) دارد. به‌وسیلهٔ مدل‌سازی موضوعی، می‌توان موضوعات موجود در داده‌های خام ورودی را شناسایی کرد. شیخا (۲۰۲۰) در مقاله‌ای، به داده‌کاوی داده‌های تویبیتز در مورد کووید-۱۹ پرداخته است؛ در این پژوهش، از اطلاعات تویبیتز برای داده‌کاوی اطلاعات مربوط به کووید-۱۹ از ۳ تا ۳۱ مارچ استفاده شده و ۱۰۰ هزار تویبیت استخراج شده است. این داده‌ها با استفاده از تکنیک‌های پردازش زبان طبیعی، همچون: معکوس تکرار در تکرار سند (TF-IDF)، ریشه‌یابی، توکن‌سازی، آنالیز معنایی پنهان (LSA)، تخصیص پنهان دیریکله (LDA) انجام شده است. همچنین برای کاهش ابعاد خوشه‌بندی حاصل از مدل‌های LDA و LSA، از الگوریتم‌های خوشه‌بندی مانند: HDBSCAN و K-Means استفاده و به مقایسهٔ این دو نیز پرداخته شده است. نتایج این پژوهش نشان می‌دهد مدل‌سازی موضوعی براساس الگوریتم LDA، عملکرد به‌مراتب بهتری نسبت به LSA دارد.

اثر چندین تکنیک مدل‌سازی موضوعی در جنبه‌های مختلف پردازش طبیعی زبان فارسی، کاملاً مشهود و مفید است. بررسی‌ها نشان می‌دهد که این تکنیک به اندازه‌گیری مؤثر تکرار تشبیهات کمک می‌کند. مدل‌های تشخیص تشبیه بر پایهٔ LDA، می‌تواند برای پردازش طبیعی زبان فارسی مفید باشد (قوبدل، ۲۰۱۴).

در زمان پاندمی کرونا، محتواهای زیادی در تویبیتز در این‌باره منتشر شد. ایران، چین، کرهٔ جنوبی و ایتالیا جزء اولین کشورهایی بودند که با اولین موج انتشار ویروس کرونا مواجه شدند. تویبیتز یکی از پلتفرم‌های وسیعی است که ایرانیان داخل و خارج از کشور، از آن برای به‌اشتراک‌گذاری عقاید، نظرات و احساسات خود استفاده می‌کنند؛ با توجه به حجم زیاد اطلاعات منتشرشده در مورد کرونا در ایران، می‌توان با استفاده از داده، موضوعات بحث‌شده در میان کاربران را

موردبررسی قرار داد تا واکنش آنها به پاندمی در گذر زمان مشخص شود. برای این کار، علاوه بر ابزارهای مدل‌سازی موضوعی، از تفسیرهای غیرماشینی، انسانی و دستی نیز می‌توان بهره برد. این چهارچوب و متدولوژی- که برای ردیابی پاسخ جامعه به کووید ۱۹ استفاده شده- را می‌توان به‌عنوان یک چهارچوب جامع در نظر گرفت که بتوان از آن در موضوعات دیگر و بررسی واکنش‌های مردم به آنها استفاده کرد و درواقع، این مدل تنها محدود به موضوع کرونا و ایرانیان نخواهد بود (حسینی و همکاران، ۲۰۲۰)

استفاده از میکروبلگ‌ها مانند: توییتر و پیام‌رسان‌های پیامکی در دوران‌های دشوار و بروز فجایع، اطلاعاتی ارزشمند را ارائه می‌دهد که با توجه به همین مسئله، می‌تواند منبعی مناسب برای سازمان‌های در ارتباط با فجایع باشد، زیرا باعث می‌شود اطلاعات حیاتی آنها همیشه به‌روز باشد. این اطلاعات با پردازش‌ها و مصورسازی‌های بیشتر و مفیدتر، می‌تواند کمک کنند کارشناسان زودتر به نتایج کاربردی برسند. برای بررسی داده‌های توییتر مرتبط با فجایع، می‌توان از مدل‌سازی موضوعی بهره برد. در کنار آن، مصورسازی و وزن‌دهی به کلمات نیز قابل‌استفاده هستند، البته باید به این نکته توجه کرد که مدل‌سازی موضوعی LDA، یک الگوریتم غیرنظارت‌شده است که اگر داده‌های اولیه، به‌خوبی پیش‌پردازش و طبقه‌بندی‌شده باشد، مشکلی به وجود نمی‌آید (کیریف و همکاران، ۲۰۰۹).

تحلیل احساسات

با رشد روزافزون نقدها و نظرات موجود در اینترنت، تحلیل احساسات، روزبه‌روز توجهات بیشتری را در زمینه پردازش طبیعی زبان و بازیابی اطلاعات، به خود جلب می‌کند. متن‌کاوی زبان فارسی همیشه چالش‌های خود را دارد؛ یکی از این چالش‌ها، کاربرد بیش از حدِ پسوندها و وجود فاصله در بعضی از کلمات چندقسمتی فارسی و نیم‌فاصله‌ها است (باقری و همکاران، ۲۰۱۳).

توییتر به‌عنوان یک شبکه اجتماعی متن‌محور، حاوی متون زیادی است که روزانه، توسط کاربران زیادی منتشر می‌شوند. چنین حجم زیادی از متون به‌اشتراک گذاشته‌شده توسط کاربران مختلف، می‌تواند منبعی بسیار مناسب برای

مدیران باشد تا احساسات مردم را درک کنند. البته تحلیل احساسات در توییت‌ها به‌علت وجود محدودیت کاراکتر و گاه، زبان طنز بعضی از توییت‌ها، با چالش‌هایی روبه‌روست، ولی از آن‌جا که روزانه، توییت‌های بسیاری در این شبکه منتشر می‌شود، توییت‌ها می‌تواند منبعی بسیار غنی برای تحلیل احساسات به‌شمار رود.

عباسی و همکاران (۲۰۱۸) در مقاله‌ای، به تحلیل احساسات داده‌های توییت‌ها با استفاده از الگوریتم‌های داده‌کاوی پرداخته‌اند. محققان در این مقاله بیان داشته‌اند که کاربران زیادی در جهان، نظرات خود را راجع به موارد مختلف در شبکه‌های اجتماعی به اشتراک می‌گذارند. تجزیه و تحلیل احساسات، شامل: فرایند شناسایی نظرات در داده‌های ساختاریافته و بدون ساختار یا نیمه‌ساختاریافته است. بررسی نظرات راجع به یک سازمان، راهی مؤثر و مفید برای نظارت بر احساسات افراد، نسبت به نام تجاری، کسب‌وکار، مدیران و.. است. امروزه سازمان‌ها و برندهای تجاری مختلف برای شناسایی احساسات و نظرات مردم از خوشه‌بندی و تحلیل احساسات داده‌های بدون برچسب شبکه‌های اجتماعی استفاده می‌کنند.

مدل‌سازی موضوعی و تحلیل احساسات، از آن دسته تکنیک‌های داده‌کاوی هستند که می‌توان به‌خوبی از آنها در کنار هم استفاده کرد. از این طریق، می‌توان علاوه بر شناسایی موضوعات نهفته در متون، به احساسات کاربر نسبت به آنها نیز پی برد.

مقاله کایلا و پراساد (۲۰۲۰)، به بررسی جریان اطلاعات در توییت‌ها در زمان پاندمی کرونا پرداخته است. در این پژوهش، توییت‌های مرتبط با هشتگ کرونا و ویروس شناسایی شدند و به‌وسیله تحلیل احساسات و مدل‌سازی موضوعی LDA، موردپردازش قرار گرفتند؛ طبق این تحقیق، توییت‌ها، محبوب‌ترین واسطه‌ها برای کسب اطلاعات در طی پاندمی کرونا بوده که این اطلاعات، با پردازش، می‌توانند برای دولت‌ها و سازمان‌ها- به‌منظور کنترل ترس عمومی و پخش اطلاعات مفید- مورد استفاده قرار بگیرند. نتایج این مطالعه نشان می‌دهد که جریان اطلاعات در توییت‌ها، دقیق و قابل‌اطمینان و همچنین مرتبط با ویروس کرونا است. بررسی LDA توانست موضوعاتی کاملاً مرتبط با کرونا را از میان توییت‌ها استخراج کند. علاوه بر این، تحلیل احساسات نیز نشان می‌دهد بعضی از توییت‌ها با واکنش‌های

منفی، نظیر ترس، روبه‌رو بودند و بعضی دیگر، حاوی احساسات مثبت، نظیر امید و اعتماد بودند.

روش تحقیق

در این پژوهش، سعی داریم با رویکرد استقرایی، با استفاده از توییت‌ها، نتیجه‌گیری کلی داشته باشیم. از طرفی، پژوهش حاضر از لحاظ هدف، یک پژوهش کاربردی به حساب می‌آید. چون هدف آن، مشخص کردن چگونگی استفاده از الگوریتم‌های متن‌کاوی شبکه اجتماعی توییت‌ها در زمینه نوآوری اجتماعی است.

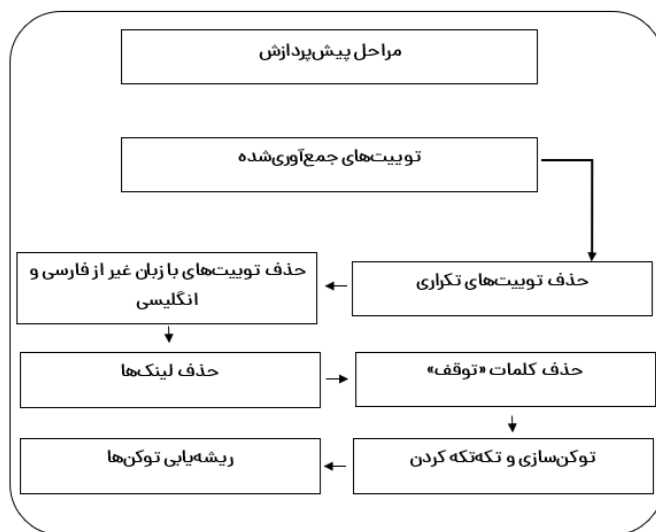
داده‌های موردنیاز این پژوهش، از تاریخ ۱۶ ژانویه ۲۰۱۶ تا ۲۵ آگوست ۲۰۲۰ با استفاده از توییت‌های حاوی کلمه کلیدی و هشتگ‌های: کارآفرینی اجتماعی و نوآوری اجتماعی و #نوآوری-اجتماعی و #کارآفرینی-اجتماعی، به زبان فارسی، جمع‌آوری شده است. همچنین از آن جا که پژوهش، درمورد نظرات مردمی است و با متون توییت‌های کاربران و آنالیز متنی آنان سروکار دارد، کیفی به حساب می‌آید از طرفی، تعداد نظرات و اعمال وزن‌های مختلف به نظرات نیز، هدف پژوهشگر بوده است، در نتیجه، بخشی از این پژوهش، کمی خواهد بود؛ در واقع، این پژوهش، ترکیبی از کمی و کیفی است.

برای جمع‌آوری اطلاعات از توییت‌ها، راه‌ها و روش‌های مختلفی وجود دارد که می‌بایست با توجه به ماهیت پژوهش، بهترین راه را انتخاب کنیم. استفاده از کتابخانه و پکیج‌های موجود و مرتبط با زبان برنامه‌نویسی پایتون، از روش‌های استخراج اطلاعات از توییت‌ها است.

با توجه به ماهیت پژوهش و همچنین محدودیت‌های API‌های رسمی ارائه‌شده توییت‌ها، برای استخراج اطلاعات، روش چهارم، یعنی استفاده از کتابخانه‌ها و پکیج‌های برنامه‌نویسی پایتون و کتابخانه getOldTweets نسخه ۳ که با نسخه ۳ زبان برنامه‌نویسی پایتون نیز به خوبی سازگار است، انتخاب شد. در مجموع، ۱۰۲۵ توییت به زبان فارسی جمع‌آوری شد. تاریخ توییت‌های مذکور از ۱۶ ژانویه ۲۰۱۶ تا ۲۵ آگوست ۲۰۲۰ است.

در راستای پیاده‌سازی مدل این پژوهش، بعد از جمع‌آوری اطلاعات، آنها را مورد پیش‌پردازش قرار دادیم تا بتوان بهترین نتیجه را از آن گرفت. در این پژوهش، از ابزار متن‌کاوی فارسی‌یار استفاده شده است. این ابزار، مجموعه‌ای از سرویس‌ها، ابزارها و خدمات سفارش متن‌کاوی را به پژوهشگران و کسب‌وکارها ارائه می‌دهد. پروژه‌های فارسی‌یار، کدهای اپلیکیشن اندرویدی فارسی‌یار، پیکره‌ها و نمونه کدهای استفاده از ابزارهای مختلف فارسی‌یار، به‌صورت متن‌باز در دسترس عموم قرار دارند. در ادامه، به بررسی مراحل پیش‌پردازش و پیاده‌سازی مدل اشاره می‌کنیم.

در ابتدا، توییت‌های تکراری را حذف کردیم، لینک‌ها را از متن توییت‌ها شناسایی و آنها را از بین متون پاک کردیم. سپس توییت‌های با زبان غیرفارسی را از مجموعه داده‌ها حذف کردیم، زیرا ممکن است فردی با هشتگ به زبان فارسی، توییتی به زبان دیگر ثبت کند. سپس، جملات را توکن‌سازی یا به‌عبارتی، قطعه‌قطعه کردیم و در مرحله بعد، «کلمات ایست» که در واقع، کلماتی هستند که به‌تنهایی، معنای خاصی ندارند را از کلمات شناسایی‌شده حذف کردیم. در آخر نیز از آن‌جا که در زبان فارسی، کلمات زیادی دارای پسوند و پیشوند هستند، ریشه کلمات را به‌منظور پردازش‌های مراحل بعد به دست آوردیم. در شکل ۱، خلاصه فرایند پیش‌پردازش داده‌ها قابل‌مشاهده است.



شکل ۱. مراحل پیش‌پردازش داده اولیه

پس از پیش‌پردازش اطلاعات، سراغ استخراج موضوعات و تحلیل احساسات رفتیم. برای تحلیل احساسات، از ابزار تحلیل احساسات فارسی‌یار کمک گرفتیم؛ برای این کار، داده‌های پیش‌پردازش‌شده را به‌عنوان ورودی، به این ابزار دادیم و این ابزار، داده‌ها را براساس نمرهٔ احساسی آنالیزشده، طبقه‌بندی کرد، طبق داکيومنت این ابزار:

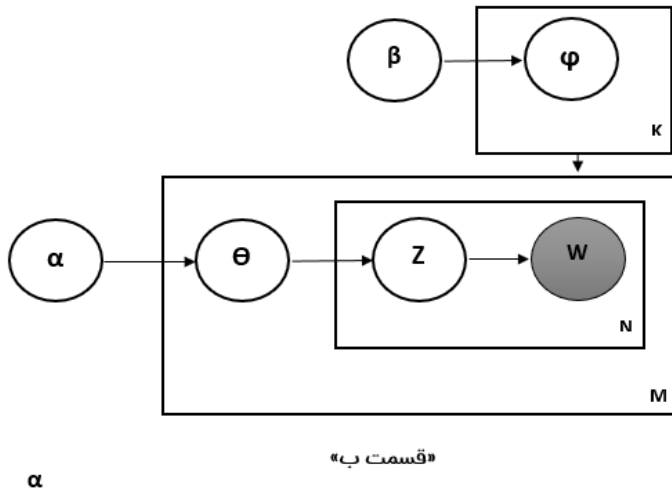
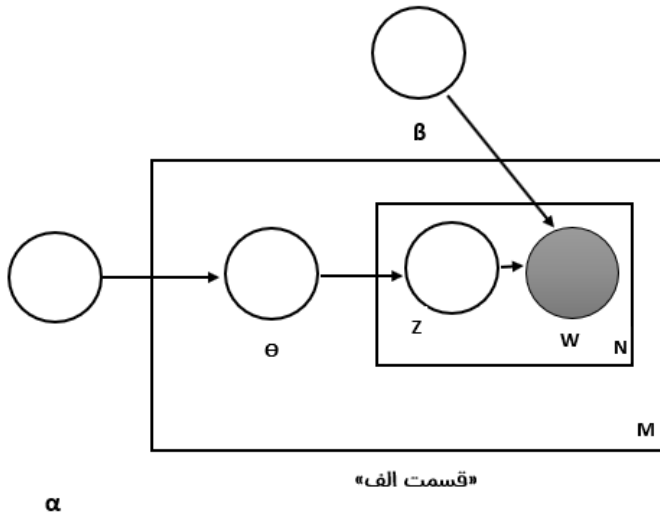
- اگر احساس به‌دست‌آمده، با عدد یک برابر باشد، جمله، خنثی و احساسی در آن شناسایی نشده است؛
- اگر احساس به‌دست‌آمده، با عدد دو برابر باشد، جمله، حاوی احساس مثبت است؛
- اگر احساس به‌دست‌آمده، با عدد صفر برابر باشد، جمله، حاوی احساس منفی است.

البته این ابزار، فعلاً به‌صورت آزمایشی عرضه شده و در حال توسعه است، اما قطعاً هر چقدر بیشتر از آن استفاده شود، در آینده، عملکردی بهتر خواهد داشت. در مرحلهٔ استخراج موضوع، برای آنالیز موضوعات توییت‌های اصلی، از مدل‌سازی موضوعات براساس تخصیص پنهان دیریکله (LDA) استفاده کردیم؛ هدف اصلی از استفاده از LDA در اینجا، پیدا کردن موضوعاتی است که در توییت‌ها در حال بحث هستند (حسینی و همکاران، ۲۰۲۰). LDA مدلی برای آزمایش مجموعه‌داده‌های بزرگ برای تشخیص شباهت در آنهاست. LDA مثالی از مدل‌سازی موضوعی است که در نمونهٔ استخراج‌شده، به موضوعات مختلف تقسیم‌بندی می‌شود. LDA در ابتدا توسط Pritchard و Stephans به‌عنوان یک تکنیک یادگیری ماشین، توسعه داده شد و بعدها کاربرد آن توسط Blei و همکاران، توسعه داده شد (ساورا و بنت ، ۲۰۱۹). مدل‌سازی موضوعات LDA، یک الگوریتم غیرنظارتی یادگیری ماشین است که تاپیک‌ها یا موضوعات را از مجموعهٔ اسناد شناسایی می‌کند؛ این تاپیک‌ها مجموعه‌ای از کلمات هستند و بر این فرض شناسایی می‌شوند که هر یک از اسناد را می‌توان با توزیع تاپیک‌ها توصیف و هر تاپیک را می‌توان با تقسیم‌بندی کلمات توصیف کرد. در الگوریتم‌های نظارتی، رایانه،

پردازش‌های خود را بر روی داده‌هایی از پیش برچسب‌گذاری شده انجام می‌دهد، ولی در الگوریتم‌های غیرنظارتی، داده‌ها از قبل، برچسب‌گذاری نشده‌اند و رایانه سعی می‌کند ساختارهای پنهان را از این داده‌ها استخراج کند (لیوید ، ۲۰۱۳). در شکل ۲، مدل ابتدایی LDA نمایش داده شده که قسمت «الف» تصویر، مدل ابتدایی و قسمت «ب»، بر پایه توزیع موضوع است.

پارامترهای مدل عبارت‌اند از:

- M تعداد متن‌ها.
- N تعداد کلمات در هر متن.
- K تعداد موضوعات.
- α : پارامتر توزیع پیشین دیریکله برای عناوین به‌ازای هر متن است .
- β : پارامتر توزیع پیشین دیریکله برای توزیع کلمات به‌ازای هر موضوع (تاپیک) است .
- θ_i : توزیع عناوین برای متن i -ام است .
- ϕ_k : توزیع کلمات برای موضوع (تاپیک) k -ام .
- Z_{mn} : موضوع m -امین کلمه در n -امین متن .
- W_{mn} : m -امین کلمه در n -امین متن .
- ϕ : ماتریس $K \times V$ توزیع-دیریکله پارامترهای موضوع- کلمه و V تعداد واژه‌هاست.



شکل ۲. مدل ابتدایی LDA

خوشبختانه، با وجود کتابخانه‌های پایتون، دیگر نیازی به این محاسبات پیچیده وجود ندارد و تمام کارها به وسیله زبان برنامه‌نویسی پایتون و کتابخانه‌های موجود و مرتبط در آن انجام می‌شود. در این پژوهش، از کتابخانه genism پایتون استفاده شده است؛ تمام مدل‌های این کتابخانه، مورد آزمایش و آموزش قرار

گرفته‌اند. کتابخانهٔ `genism` هیچ‌گونه محدودیتی ندارد، به‌صورت کاملاً متن‌باز در اختیار کاربران قرار می‌گیرد و کاربران می‌توانند هر مجموعه‌داده با هر حجمی را به‌وسیلهٔ آن، مورد پردازش قرار دهند.

در این مقاله نیز از کتابخانهٔ `genism` و ابزار `LdaModel` استفاده شده است؛ این مدل، ورودی‌های مختلفی می‌گیرد که بعضی از آنها اجباری و بعضی اختیاری‌اند. این ابزار، برای شروع به‌کار، به دیکشنری به‌دست‌آمده از مجموعه‌داده و مجموعه‌ای از متون، به‌عنوان ورودی نیاز دارد و تعداد موضوعات استخراج‌شدهٔ نهایی و متغیرهای مدل `lda`، دیگر ورودی‌های اختیاری هستند. درصورتی‌که ورودی‌های اختیاری را نیز به این مدل بدهیم، ورودی داده‌شدهٔ ما اولویت بیشتری یافته و مقدار ورودی داده‌شدهٔ ما را برای محاسبهٔ مدل در نظر می‌گیرد.

برای استخراج موضوع از متن توییت‌ها، به توسعهٔ مدل و بررسی و آزمون این مدل، نیاز است تا بتوانیم بهترین عملکرد آن را متناسب با ورودی داده‌شده به آن تشخیص دهیم؛ بنابراین، به تعریف برخی از معیارها می‌پردازیم. سرگشتگی یکی از معیارهای ارزیابی است و به‌طور گسترده، برای ارزیابی مدل زبان از آن استفاده می‌شود. این مفهوم، نشان می‌دهد که چگونه یک مدل، با داده‌های جدیدی که قبلاً مشاهده نکرده، رفتار می‌کند و واکنش غیرطبیعی نشان می‌دهد و یا به‌نوعی شگفت‌زده می‌شود. معیارهای وابستگی موضوع، با اندازه‌گیری میزان شباهت معنایی بین کلمات (با امتیاز بالا) موجود در موضوع به دست می‌آید. این اندازه‌گیری به تمایز بین موضوعات قابل‌تفسیر معنایی و موضوعات مصنوعی استنباط آماری کمک می‌کند. هر چه در مدل ما، میزان سرگشتگی، کمتر و میزان وابستگی موضوعی، بیشتر باشد، مدلی بهینه‌تر خواهیم داشت.

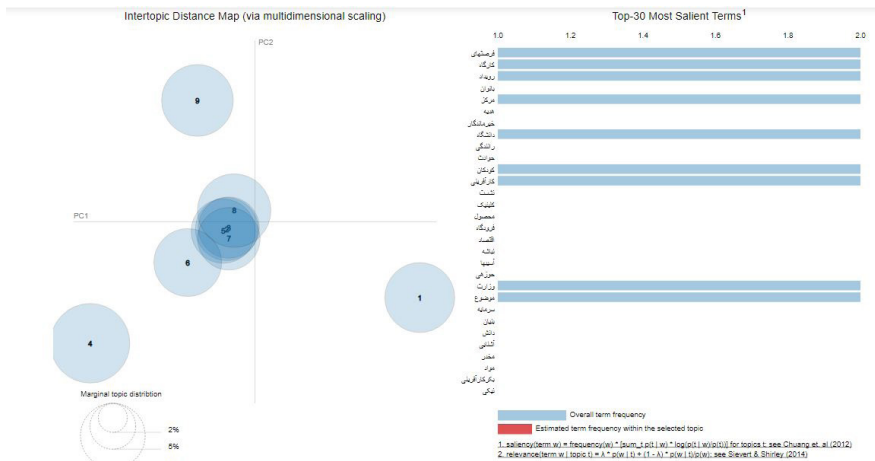
برای مصورسازی نتایج مدل `LDA` نیز، از کتابخانهٔ `pyLDAvis` استفاده کردیم. این کتابخانه به کاربران کمک می‌کند بیشتر و راحت‌تر با موضوعات استخراج‌شده از مدل `LDA` ارتباط برقرار کنند. این کتابخانه، از اطلاعات استخراج‌شدهٔ مدل `LDA` استفاده می‌کند و یک خروجی مصور تحت‌توب به کاربر نشان می‌دهد.

ابر کلمات، ابزاری بسیار مناسب برای مصورسازی کلمات و تکرار آنها در حجم وسیعی از متون هستند. در این پژوهش نیز، از ابزار `word_cloud_fa` برای رسم ابر کلمات استفاده شده است.

یافته‌ها

نتایج تحلیل احساسات نشان می‌دهد که در متون نوآوری اجتماعی تعداد ۲۱ توییت با حس منفی، ۵۷ توییت با حس مثبت و ۱ توییت با حس خنثی تشخیص داده شده‌اند و در متون کارآفرینی اجتماعی، تعداد ۱۱۷ توییت با حس منفی، ۲۱۳ توییت با حس مثبت و ۱۷ توییت با حس خنثی تشخیص داده شده‌اند که البته نباید وجود عبارات طنزآمیز و کنایه‌آمیز را در زبان توییت‌ر فارسی نادیده گرفت، زیرا وجود طنز و کنایه در متون، قدرت تشخیص احساسات نهفته را کاهش می‌دهد.

در استخراج موضوع، بعد از انجام تست‌ها و بررسی هایپرپارامترها بر روی ۷۵ درصد داده‌ها، مشخص شد که مدل ما با هایپرپارامترهای آلفا برابر با ۰/۹۱ و اتا برابر با ۰/۹۱ و ۹ تعداد موضوع، بهترین میزان وابستگی با عدد ۰/۷۰ و سرگشتگی ۷/۱۹- را به دست آورده که میزان بسیار مناسبی است. با توجه به نتایج به‌دست‌آمده از مدل حاصل از متون توییت نوآوری اجتماعی، مدل را با بهترین عملکرد خود اجرا کردیم و نتایج زیر به دست آمد. در مرحله بعد، می‌بایست با توجه به کلمات موضوع به‌دست‌آمده را در قالب یک جمله توصیف کرد (ساورا و همکاران، ۲۰۱۹).



شکل ۳. تصویری از خروجی pyLDA به‌دست‌آمده برای موضوعات نوآوری اجتماعی

در تصویر ۳، توزیع موضوعات مختلف در نواحی مختلف معنایی دیده می‌شود. همچنین در گوشه سمت چپ تصویر نیز راهنمای میزان هم‌پوشانی معنایی موضوعات استخراج‌شده، به نمایش گذاشته شده است. با توجه به توزیع موضوعات و هم‌پوشانی‌ها، می‌توان گفت موضوعات ۱، ۴، ۶ و ۹ با دیگر موضوعات، تداخل ندارند و به موضوعاتی جداگانه اشاره می‌کنند، درحالی‌که سایر موضوعات، از نظر معنایی به موضوعاتی واحد اشاره می‌کنند. در ادامه، موضوعات استخراج‌شده از توییت‌های نوآوری اجتماعی را مشاهده می‌کنیم:

جدول ۳. موضوعات نوآوری اجتماعی

موضوع	توضیح موضوع
موضوع ۱	احداث گلخانه و دندان‌پزشکی در جنوب کشور
موضوع ۴	برگزاری رویدادهایی برای آگاهی از ایده‌های بانوان
موضوع ۶	فعالیت‌های دانش‌بنیان در زمینه‌های فرهنگی، طبیعی، اقتصادی
موضوع ۹	برگزاری رویدادهایی در مورد آلودگی، خشکسالی و سلامت
سایر موضوعات	فراخوان فرصت‌های مطالعاتی و برپایی رویدادهایی برای توانمندسازی بانوان و همچنین شناسایی فرصت‌های کارآفرینی برای حل مشکلات اقتصادی

برای داده‌های مربوط به کارآفرینی اجتماعی نیز، مدل آزمایش شده خود با هایپیرامترهای به‌دست‌آمده را اجرا می‌کنیم که میزان وابستگی، $0/58$ و میزان سرگستگی، $7/8-$ است که اعداد مناسبی هستند. با توجه به آموزش‌های داده‌شده و نتایج به‌دست‌آمده از مدل حاصل از متون توییت کارآفرینی اجتماعی، مدل را با بهترین عملکرد خود اجرا کردیم و نتایج زیر به دست آمد.

بحث و نتیجه‌گیری

تویبتر به‌عنوان یکی از شبکه‌های اجتماعی محبوب، منبعی غنی از اطلاعات است و کاربران زیادی، احساسات، نظرات و اتفاقات روزمره خود را در آن، با یکدیگر به اشتراک می‌گذارند. پردازش این حجم از اطلاعات با استفاده از روش‌های سنتی، بسیار زمان‌بر و البته پریسک خواهد بود. لذا به کمک داده‌کاوی و متن‌کاوی، می‌توان این داده‌ها را به‌خوبی موردپردازش قرار داد و از نتیجه این پردازش‌ها، در تصمیم‌گیری مربوط به طرح‌ها و پروژه‌های مختلف استفاده کرد تا طرح‌ها مقبولیت بیشتری پیدا کنند و دغدغه اصلی مردم در آنها، در نظر گرفته شود. نوآوری اجتماعی نیز به‌عنوان یک پدیده اجتماعی، با هدف افزایش رفاه جامعه و حل معضلات اجتماعی پدید آمده است. در این پژوهش، توانستیم ابتدا با تحلیل احساسات، نشان دهیم که مردم نسبت به طرح‌های نوآوری اجتماعی و کارآفرینی اجتماعی، حسی مثبت دارند و سپس با کمک مدل‌سازی موضوعی LDA، از داده‌های تویبتر مرتبط با نوآوری اجتماعی و کارآفرینی اجتماعی اطلاعات مفیدی استخراج کنیم.

اطلاعات و بینش‌های مفیدی همچون: توجه به صنایع‌دستی روستایی، حمایت از کارآفرینان در جذب سرمایه، توجه به منطقه چابهار و جنوب کشور، افزایش رفاه زندگی مردم، و افراد معلول و توجه به معضل کودکان کار، موارد مهمی بودند که در تویبترهای کاربران دیده می‌شدند. با توجه به ابر کلمات و موضوعات استخراج‌شده، به‌راحتی می‌توان موارد پرتکرار تویبترهای کاربران را تشخیص داد و اطلاعات مفیدی را که به آنها اشاره شده، به دست آورد:

- در متون نوآوری اجتماعی، مواردی همچون: همایش‌های دانشجویی و برگزاری استارت‌آپ‌ویکندهایی به‌منظور شناسایی ایده‌های مختلف، دیده می‌شود که نشان می‌دهد پژوهش‌ها و مطالعات نظری در این زمینه، به‌خوبی در حال انجام است؛
- دغدغه‌ها و موضوعاتی همچون: خشکسالی، آلودگی و طبیعت، شناسایی شد که نشان می‌دهد مردم به این معضلات آگاه هستند و برای حل آنها با یکدیگر در تویبتر، گفت‌وگو می‌کنند؛

- در متون نوآوری اجتماعی، به لزوم ارتقای سطح رفاه در مناطق جنوب کشور در قالب مواردی همچون تأسیس دندان‌پزشکی اشاره شده است؛
- کودکان کار نیز یکی از موارد دیده‌شده در متون نوآوری اجتماعی هستند که نشان می‌دهد مردم سعی دارند به‌وسیله نوآوری اجتماعی، به کودکان کار کمک کنند. در نتیجه، می‌توان در طرح‌های مربوط به خدمت‌رسانی به کودکان کار از نوآوری اجتماعی بهره برد؛
- در داده‌های ما، موضوع توانمندسازی بانوان کاملاً به چشم می‌خورد که نشان می‌دهد تقویت بانوان و تغییر ذهنیت پیرامون این موضوع، در ایران اتفاق افتاده و مردم نیز نسبت به آن آگاهی دارند. بنابراین، می‌توان در طرح‌های مختلف، با توجه به سطح مقبولیت بالا نسبت به رهبری بانوان، از آنها کمک گرفت؛
- در مجموعه داده‌های ما، به کمک به افراد دارای معلولیت و افزایش کیفیت زندگی آنها اشاره شده است، مردم نیز در توییت‌ها، پیرامون این موضوع صحبت کرده و نسبت به آن آگاهی دارند؛
- در موضوعات به‌دست‌آمده، موضوع کمک و حل مشکلات اجتماعی و البته کارآفرینی منطقه چابهار به چشم می‌خورد که نشان می‌دهد مردم نسبت به این موضوع، آگاهی و دغدغه دارند و لازم است کارهای بیشتری در ارتباط با این موضوع انجام شود همچنین مردم از پتانسیل منطقه چابهار آگاه‌اند، به فکر مسئله اشتغال در این منطقه هستند و اگر مسئله اشتغال در چابهار به‌درستی حل شود، مسلماً نتایج مفیدی به دست خواهد آمد؛
- در موضوعات و موارد شناخته‌شده، موضوع صنایع‌دستی و کارآفرینی به چشم می‌خورد و با توجه به ظرفیت‌هایی که کشورمان در این حوزه دارد، می‌توان طرح‌های کارآفرینی بیشتری را- به‌خصوص برای بانوان روستایی، به‌منظور اشتغال‌زایی و حل مشکلات اقتصادی- ایجاد کرد؛
- یکی دیگر از موضوعات شناسایی‌شده- که جزء دغدغه‌های اصلی مردم نیز به‌شمار می‌رود- بحث اشتغال است که با توجه به موضوعات و

موارد پرتکرار، با شناسایی فرصت‌های کارآفرینانه در راستای حل مشکلات اقتصادی، می‌توان در بحث اشتغال نیز شاهد پیشرفت‌هایی بود؛

- از دیگر موضوعاتی که شناسایی شده، مشکل جذب سرمایه در پروژه‌های کارآفرینی است که سیاستگذاران می‌توانند این موضوع را مدنظر داشته باشند تا تسهیلات بیشتری در اختیار کارآفرینان- و به‌ویژه کارآفرینان اجتماعی- قرار گیرد.

تصمیم‌گیرندگان و مدیران سازمان‌ها و همچنین کارآفرینان اجتماعی، می‌توانند با توجه به موارد مطرح‌شده در بالا، طرح‌هایی را اجرا کنند که بیشتر مردم، نسبت به آنها دغدغه دارند تا در برنامه‌ریزی‌های خود، هدف‌گیری درستی داشته باشند. از این طریق، می‌توان مشارکت مردم در این‌گونه طرح‌ها را نیز بیشتر کرد که در نهایت، می‌تواند منجر به افزایش رفاه سطح کلی جامعه شود. برای پژوهش‌های آینده نیز می‌توان شبکه‌های اجتماعی دیگر را مورد‌پردازش قرار داد تا طیفی گسترده‌تر از کاربران، پوشش داده شوند. علاوه بر این، می‌توان از موضوعات استخراج‌شده این پژوهش، به‌عنوان داده اولیه برای استخراج داده‌های مرتبط با آن و پردازش‌های بیشتر برای بررسی هر چه عمیق‌تر هر موضوع استفاده کرد.

فهرست منابع

- چیت‌سازیان، علیرضا؛ و صادقی‌کیا، محمدعلی (۱۳۹۷). مشارکت مردمی در نوآوری اجتماعی، **مدیریت نوآوری**، ۲۰-۱.
- خدایاری، مریم؛ و نیلفروشان، هادی (۱۳۹۸). «کنکاشی بر مفهوم‌شناسی و ابعاد اصلی نوآوری اجتماعی». **شانزدهمین کنفرانس بین‌المللی مدیریت**.
- سعدآبادی، علی‌اصغر؛ پورعزت، علی‌اصغر؛ قانع‌راد، محمدمبین؛ مختارزاده گروسی، نیما؛ و الیاسی، مهدی (۱۳۹۵). «چرخه حیات نوآوری اجتماعی».
- سهرابی، بابک؛ رئیس، بابک؛ و خدایرست، فرشته (۱۳۹۵). «تحلیل نظرات کاربران وب‌سایت‌های تجارت اجتماعی براساس روش‌های متن‌کاوی و داده‌کاوی»، **مجله جهانی رسانه**.
- عباسی، مهدی (۱۳۹۵). «طبقه‌بندی توثیق‌های فارسی شبکه اجتماعی توییتر با استفاده از روش متن‌کاوی».
- عباسی، مهناز؛ موبیدی، فاطمه؛ و بازرگان، کیمیا (۱۳۹۷). «آنالیز احساسات داده‌های توییتر با استفاده از الگوریتم‌های داده‌کاوی». **پنجمین کنفرانس ملی علوم و مهندسی کامپیوتر و فناوری اطلاعات**.
- عرفان‌منش، محمدمبین؛ حسینی، الهه؛ و حبیبی، سحر (۱۳۹۷). «تحلیل توییتهای علمی مقالات در توییتر». **فصلنامه مطالعات ملی کتابداری و سازمان‌دهی اطلاعات**، ۱۰۷۰-۱۰۵۸.
- کشتکار، مهران (۱۳۹۷). «شناسایی ابعاد و مؤلفه‌های نوآوری اجتماعی در ایران»، **فصلنامه مدیریت نوآوری در سازمان‌های دفاعی**، ۲، ۱۵۱-۱۴۴.
- مرجانی، تیمور؛ و صدری، صدرالدین (۱۳۹۳). «توسعه کارآفرینی اجتماعی؛ آسیب‌ها، عوامل مؤثر و راهکارها». **مدیریت فرهنگ سازمانی**.
- میرزاآقازاده، میترا؛ سلوک، وحید؛ و باقرزاده، جمشید (۱۳۹۴). «ارزیابی تأثیر محبوبیت کاربر و هشتگ در محبوبیت پست‌های شبکه اجتماعی توییتر». **هفتمین کنفرانس بین‌المللی فناوری اطلاعات و دانش**.
- نداف، مهدی؛ جلدانی، شادی؛ و محمدی، سارا (۱۳۹۸). «بررسی نقش عوامل نهادی در کارآفرینی اجتماعی دانشجویان (شواهدی از دانشجویان مهندسی دانشگاه شهیدچمران اهواز)». **چهارمین کنفرانس ملی اقتصاد، مدیریت و حسابداری**.
- نظری، فائزه؛ نوروزی، کاوه؛ حسینی، سهراب؛ و صهبا، فرشید (۱۳۹۸). «مروری بر: داده‌کاوی متون شبکه‌های اجتماعی برای احساسات و تمایلات مصرف‌کننده برند». **چهارمین کنگره ملی تحقیقات بنیادین در مهندسی کامپیوتر و فناوری اطلاعات**، ۲۲-۱.
- Angelidou, M., & Psaltoglou, A. (2018). Digital social innovation in support of spatial planning. An investigation through nine initiatives in three smart city programmes. *Spatium*, 39, 7-16.
<https://doi.org/10.2298/SPAT1839007A>
- Bagheri, A., Saraei, M., & De Jong, F. (2013). Sentiment classification in Persian: Introducing a mutual information-based method for feature selection. *2013 21st Iranian Conference on Electrical Engineering, ICEE 2013*. <https://doi.org/10.1109/IranianCEE.2013.6599671>
- Ghavidel, H. A. (2014). *Impact of Topic Modeling on Rule-Based Persian Metaphor Classification and its Frequency Estimation*.

- Groot, A., & Dankbaar, B. (2014). Does Social Innovation Require Social Entrepreneurship? *Technology Innovation Management Review*, 4(12), 17–26. <https://doi.org/10.22215/timreview854>.
- Hosseini, P., Hosseini, P., & Broniatowski, D. A. (2020). Content analysis of Persian/Farsi tweets during COVID-19 pandemic in Iran using NLP. *ArXiv*, 1–16. <https://doi.org/10.18653/v1/2020.nlp-covid19-2.26>
- Howaldt, J., & Schwarz, M. (2017). Social Innovation and Human Development—How the Capabilities Approach and Social Innovation Theory Mutually Support Each Other. *Journal of Human Development and Capabilities*, 18(2), 163–180. <https://doi.org/10.1080/19452829.2016.1251401>
- Kaila, R.P. & Prasad, A. V. K. (2020). Informational Flow on Twitter - Corona Virus Outbreak – Topic. 11(3), 128–134.
- Kireyev, K., Palen, L., & Anderson, K. (2009). Applications of topics models to analysis of disaster-related twitter data. *NIPS Workshop on Applications for Topic Models: Text and Beyond*. http://www.umi.acs.umd.edu/~jbg/nips_tm_workshop/15.pdf
- Li, G., Dai, J. S., Park, E. M., & Park, S. T. (2017). A study on the service and trend of Fintech security based on text-mining: focused on the data of Korean online news. *Journal of Computer Virology and Hacking Techniques*, 13(4), 249–255. <https://doi.org/10.1007/s11416-016-0288-9>.
- lloyd, seth. (2013). *Quantum algorithms for supervised and unsupervised machine learning* (Issue x, pp. 1–26).
- Marques, P., Morgan, K., & Richardson, R. (2018). Social innovation in question: The theoretical and practical implications of a contested concept. *Environment and Planning C: Politics and Space*, 36(3), 496–512. <https://doi.org/10.1177/2399654417717986>
- Maylawati, D. S., Aulawi, H., & Ramdhani, M. A. (2018). The concept of sequential pattern mining for text. *IOP Conference Series: Materials Science and Engineering*, 434(1). <https://doi.org/10.1088/1757-899X/434/1/012042>
- Rossi, A., & Bocci, R. (2018). The Transformative Potential of Social Innovation. The Case of Wheat and Bread Value Chain in Tuscany. *Jrnl. of Soc. of Agr. & Food*, 24(3), 431–448. www.diversifood.eu
- Saura, J. R., & Bennett, D. R. (2019). A three-stage method for data text mining: Using UGC in business intelligence analysis. *Symmetry*, 11(4). <https://doi.org/10.3390/sym11040519>
- Saura, J. R., Palos-Sanchez, P., & Grilo, A. (2019). Detecting indicators for startup business success: Sentiment analysis using text data mining. *Sustainability (Switzerland)*, 11(3), 1–14. <https://doi.org/10.3390/su11030917>
- Sheikha, H. (2020). *text mining twitter social media for covid-19*.
- Shen, C. W., & Ho, J. T. (2018). Public opinion toward social business from a social media perspective. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10943 LNCS, 555–562. https://doi.org/10.1007/978-3-319-93803-5_52.

Sindhani, M., Parameswar, N., Dhir, S., & Ongsakul, V. (2019). Twitter analysis of founders of top 25 Indian startups. *Journal for Global Business Advancement*, 12(1), 117-144. <https://doi.org/10.1504/JGBA.2019.099918>.