Journal of Soft Computing and Information Technology (JSCIT)
Babol Noshirvani University of Technology, Babol, Iran
Journal Homepage: jscit.nit.ac.ir
Volume 13, Number 2, Summer 2024, pp. 37-54
Received: 05/01/2024, Revised: 10/15/2024, Accepted: 11/14/2024

Monitoring and Analyzing the Mentality Changes of Iranian Society in the Corona Pandemic based on Concept Drift Detection in Data Streams

Hoda Mashayekhi [1] and Fatemeh Jafarinejad [2*]
1- Associate Professor, Faculty of Computer Engineering, Shahrood University of Technology, Shahrood, Iran.
2*- Assistant Professor, Faculty of Computer Engineering, Shahrood University of Technology, Shahrood, Iran.
[1] hmashayekhi@shahroodut.ac.ir and [2*] jafarinejad@shahroodut.ac.ir

Corresponding author's address: Fatemeh Jafarinejad, Faculty of Computer Engineering, Shahrood University of Technology, Shahrood, Iran.

Abstract- In the investigation of unexpected events and their time-varying data, it is necessary to pay attention to the concept drifts to provide an authentic analysis. Although, the literature mostly deals with the technical/theoretical aspect of the drift methods, the practical application of these methods on real-world unstructured online data, and interactions of them with real-world events has received less attention. Among the unexpected events was the epidemic of corona disease (with an unknown nature, and high rate of infection and death). In the meantime, the concerns of the society were reflected in the virtual space throughout the mass of corona-related texts. In this research, we propose a method to investigate the fluctuations of the mentality in the online data stream during the early-days of the pandemic. Related posts on Telegram, Instagram, and Twitter and related Persian news during the first 100 days of the Corona outbreak were investigated. We analyze the features of the polarity of opinions, the view of officials' performance, and hope; using descriptive statistical methods as well as the proposed concept drift analysis method. Analyzing the content of key texts at the drift time points, we will discover the reasons for the change of negative concepts to positive factors and extract reliable solutions. We will illustrate that, in general, the published text during the investigated period move slowly from negative towards neutral and positive concepts. Specifically, for the hope, there is a positive hope throughout the entire period under review. Moreover, the period related to the Solar New Year holiday is a key period in terms of all three features. The review of key texts illustrates the effectiveness of strategies such as strengthening the national spirit, improving treatment facilities and justice in access to facilities from the authorities, and humoring in improvement of the mental health. The results clarify that the analysis of cyberspace data and people's mentality using concept drift leads to the extraction of useful information for more informed planning, more accurate information, and improving the mental health of society, that are important for the society to make major decisions and return to the peace.

Keywords- Coivid19, Data Mining, Data Streams, Concept Drift Detection.

مجله علمی رایانش نرم و فناوری اطلاعات
دانشگاه صنعتی نوشیروانی بابل
صفحه مجله: jscit.nit.ac.ir
جلد ۱۳، شماره ۲، تابستان ۱۴۰۳ ، صفحه ۳۷-۵۴
دریافت: ۱۴۰۳/۰۲/۱۲، بازنگری: ۱۴۰۳/۰۷/۲۴، پذیرش: ۱۴۰۳/۰۸/۲۴

پایش و تحلیل تغییرات ذهنیت افراد جامعه ایرانی در مواجهه با پاندمی کرونا بر مبنای تشخیص رانش مفهوم در جریان داده‌ها

هدی مشایخی[۱]، فاطمه جعفری نژاد*[۲]
۱- دانشکده مهندسی کامپیوتر، دانشگاه صنعتی شاهرود، شاهرود، ایران.
*۲- دانشکده مهندسی کامپیوتر، دانشگاه صنعتی شاهرود، شاهرود، ایران.
[1]hmashayekhi@shahroodut.ac.ir, [2*]jafarinejad@shahroodut.ac.ir

* نشانی نویسنده مسئول: فاطمه جعفری نژاد، شاهرود، دانشگاه صنعتی شاهرود، دانشکده مهندسی کامپیوتر.

چکیده- در بررسی وقایع غیرمترقبه و داده‌های متغیر با زمان، توجه به رانش مفهوم برای ارائه یک تحلیل معتبر ضروری می‌باشد. هرچند، پژوهش‌های مرتبط با رانش مفهوم، اکثرا به جنبه فنی و تئوری روش‌ها پرداخته و استفاده از این روش‌ها در داده‌های غیر ساخت‌یافتهٔ کاربردهای واقعی کمتر مورد توجه قرار گرفته است. هم‌چنین، پژوهش‌های اندکی بطور همزمان تغییر مفهوم در شبکه‌های اجتماعی و ارتباط آن با رویدادهای جهان واقع را در نظر گرفته‌اند. از جمله وقایع غیرمترقبه، اپیدمی بیماری کرونا (با نرخ بالای ابتلا، مرگ و میر، به همراه طبیعت ناشناخته) بود. در این بین، نگرانی‌های جامعه -به‌ویژه در روزهای ابتدایی- در فضای مجازی و در خلال انبوه متون مرتبط با کرونا منعکس شد. در این پژوهش، روشی برای بررسی تغییرات ذهنیت جامعه در جریان داده‌های برخط پیشنهاد می‌دهیم. دادگان این پژوهش، پست‌های شبکه‌های اجتماعی تلگرام، اینستاگرام و توئیتر و متون فارسی خبری مرتبط با کرونا در بازه ۱۰۰ روزه ابتدایی شیوع کروناست. درآنها ویژگی‌های قطبیت نظرات، دیدگاه نسبت به عملکرد مسئولین، و امید را رصد و تحلیل می‌کنیم. بدین منظور از روش‌های آماری توصیفی و روش پیشنهادی تحلیل رانش مفهوم استفاده می‌کنیم. سپس، با تحلیل محتوایی متون کلیدی در نقاط زمانی تغییر مفهوم، به کشف علت تغییر مفاهیم منفی به مثبت، و ارائه راهکارهای موثق به دادگان می‌پردازیم. یافته‌های این پژوهش نشان می‌دهند بطور کلی، مطالب منتشر شده در بازه مورد بررسی با شیب کندی به سمت مفاهیم خنثی و مثبت پیش می‌رود. در مورد ویژگی امید، امیدواری در کل بازه مورد بررسی وجود دارد. همچنین بازه مرتبط با تعطیلات سال نو شمسی یک بازه کلیدی از نظر هر سه ویژگی مورد بررسی است. بررسی متون کلیدی، نشان‌دهنده تأثیرگذاری راهکارهایی همچون تقویت روحیه ملی، ارتقای امکانات درمانی و عدالت در دسترسی به امکانات، از سمت مسئولین، طنزپردازی و توجه به وقایع مناسبتی به‌منظور بالا بردن بهداشت روان جامعه است. نتایج این پژوهش نشان می‌دهند که تحلیل داده‌های فضای مجازی و رصد تغییرات ذهنیت افراد در این جریان داده با استفاده از رانش مفهوم، منجر به استخراج اطلاعات مفید برای برنامه‌ریزی آگاهانه‌تر، اطلاع‌رسانی درست‌تر، و ارتقای سلامت روان جامعه می‌گردد؛ که در تصمیم‌گیری‌های کلان و بازگشت آرامش به جامعه راه‌گشاست.

واژه‌های کلیدی: ویروس کرونا، داده‌کاوی، جریان داده، تشخیص رانش مفهوم.

۱- مقدمه

از منظر سلامت و بهداشت روان، یک معضل اجتماعی را می‌توان امری دانست که بخش قابل توجهی از جامعه آن را ناخوشایند بدانند و معتقد باشند برای حل آن باید کاری کرد. با توجه به این تعریف، بی‌تردید شیوع یکبارهٔ کروناویروس یک معضل اجتماعی بود که نه تنها سلامت روان افراد را تحت تاثیر قرار داد، بلکه سبک زندگی را

دست‌خوش تغییرات نمود.

با توجه به گسترش رسانه‌های ارتباط جمعی در عصر ارتباطات، حجم عظیمی از دادگان باارزش در این رسانه‌ها وجود دارد که واکنش سریع افراد جامعه به اخبار منتشره را نشان می‌دهد. بررسی و تحلیل این دادگان، تحلیل‌گران بسیاری را در زمینه فهم مشکل، فهم واکنش مردم، یافتن علت پدیده‌های اثرپذیر بر واقعه، شناخت راهکارهای واقعی برون‌رفت از معضل و برگرداندن آرامش روانی به جامعه کمک می‌کند.

شناسایی اخبار غلط در زمینه بیماری [۱–۵]، توزیع اخبار مرتبط با کرونا در توئیتر [۶]، استفاده از شبکه‌های عمیق برای دسته‌بندی و تحلیل نظرات کاربران شبکه‌های اجتماعی [۷]، استفاده از روش‌های مدلسازی موضوعی به منظور کشف موضوعات پر اهمیت در گفتمان مردم [۸] یا کشف تکامل موضوعات در گفتمان مردمی [۴] از جمله این کاربردهاست.

در مورد شیوع ویروس کرونا، تحلیل دادگان مربوط به بازه زمانی ابتدایی وقوع این واقعه ناگهانی، می‌تواند بسیار سودمند باشد. تحلیل شبکه‌های اجتماعی و متون خبری در اپیدمی‌های مشابه نیز انجام شده است که از جمله آنها می‌توان به زیکا [۹]، ابولا [۱۰]، آنفولانزا [۱۱]، سارس [۱۲] و دیگر بیماری‌ها مانند ایدز [۱۳] و بیماری لایم در ایرلند [۱۴] اشاره نمود. همچنین سرویس روند آنفولانزای گوگل (GFT) با استفاده از اصطلاحاتی که در موتورجستجوی گوگل مورد جستجو قرار می‌گرفت، توانست شیوع سالانه بیماری آنفولانزا در ایالت متحده امریکا را مورد رصد و تحلیل قرار دهد [۱۵]. در اوایل شروع به کار، این روش دقت ۹۷ درصدی در پیش‌بینی شیوع آنفولانزا را برای خود ثبت نمود. البته با گذر زمان و رویداد رانش مفهوم در اصطلاحات و ادبیات مورد جستجو، و همچنین تغییر وضعیت آنفولانزا و درمان‌ها، به تدریج میزان خطای GFT افزایش یافته و در سال ۲۰۱۵ رسما این سرویس تعطیل شد.

استفاده از تکنیک‌های تشخیص رانش مفهوم در تحلیل متون و جریان دادگان مرتبط با کرونا در منابع اندکی مورد مطالعه قرار گرفته است. سوپرم و سو [۱۶] بر روی یک مجموعه داده چندزبانه از توئیت‌های مرتبط با کرونا کار کرده و ابزاری را برای تحلیل این جریان داده طراحی نمودند. استفاده از روش‌های مختلف تشخیص رانش مفهوم در مطالعه آنها حاکی از تغییرات سریع بحث‌های اجتماعی حول این بیماری پس از ژانویه ۲۰۲۰ بوده است. در مطالعه دیگری، پو و همکاران با استفاده از روش‌های تشخیص رانش مفهوم، محتوای درحال تکامل شبکه‌های اجتماعی را مورد مطالعه قرار دادند و به کسب اطلاعات ناشناخته و تفکیک حقایق از اطلاعات غلط در این رسانه‌های جمعی پرداختند [۱۷]. تشخیص رانش مفهوم به‌همراه روش‌های مدلسازی موضوعی در [۱۸] برای تحلیل و رصد گفتمان در کرونا مورد استفاده قرار گرفته است.

در این پژوهش، روشی برای تحلیل تغییر ذهنیت جامعه بر مبنای تشخیص رانش مفهوم ارائه می‌کنیم. برمبنای روش پیشنهادی و با بهره‌مندی از روش‌های آماری، به ارائه یک تحلیل کاربردی مستند به روش‌های هوش مصنوعی از نحوه تغییر ذهنیت و سلامت روانی افراد جامعه ایرانی در بازه ۱۰۰ روزه ابتدایی شیوع ویروس کرونا خواهیم پرداخت. سپس با شناسایی نقاط زمانی تغییر نگرش و بررسی متون و پست‌ها حول این نقاط کلیدی، به شناسایی علل تغییرات منفی و یا راهکار برون‌رفت از آنها به نفع تغییرات مثبت خواهیم پرداخت. از آنجا که این تحلیل موثق به دادگان است، راهکارهای واقعیِ مشاهده شده و تجربه شده در جامعه را ارائه خواهد داد. به بیان دیگر، این راهکارها به صورت عواملی که قبلا منجر به تغییر نگرش شده‌اند و بازخورد مثبتشان توسط دادگان ضبط شده است، استخراج می‌شود. با واکاوی این دادگان، راهکارهای مهم را استخراج کرده و به خواص و تأثیرگذاران حوزه سلامت پیشنهاد خواهیم داد. نوآوری این کار عبارتند از:

- استفاده از این روش‌های مبتنی بر تشخیص رانش مفهوم، برای بررسی داده‌های غیر ساخت‌یافته در یک کاربرد دنیای واقعی
- در نظر گرفتن توأمان تغییر مفهوم در شبکه‌های اجتماعی و ارتباط آن با رویدادهای جهان واقع
- پیشنهاد روشی برای تحلیل تغییر ذهنیت افراد جامعه مبتنی بر تشخیص رانش مفهوم در جریان داده، و شناسایی نقاط تغییر نگرش از مثبت به منفی یا برعکس.
- تحلیل و رصد جریان داده‌های متنی مرتبط با بیماری کرونا در فضای مجازی به منظور بررسی ویژگی‌های قطبیت، نگرش به عملکرد مسئولین و امید.
- استخراج محتوای مرتبط با موقعیت‌های رانش مفهوم و بهره‌مندی از آن جهت یافتن عوامل تغییر نگرش و همچنین برجسته کردن عملکردهای مثبت یا منفی.
- ارائه راهکارهای برخاسته از سوابق داده‌ای به منظور هدایت گفتمان جامعه به سمت ذهنیت مثبت منطقی و نه ناشی از سهل‌گیری بیماری.

ساختار این مقاله در ادامه بدین ترتیب است: در بخش دوم به مرور مفهوم تشخیص رانش مفهوم و تحقیقات انجام شده در این زمینه خواهیم پرداخت. بخش سوم به معرفی معماری پیشنهادی تشخیص رانش مفهوم در داده‌های شبکه‌های اجتماعی اختصاص دارد. در بخش چهارم به ارائه نتایج بدست آمده از اعمال روش پیشنهادی بر روی دادگان شبکه اجتماعی در ۱۰۰ روز ابتدایی واقعهٔ پاندمی کرونا

می‌پردازیم. در همین بخش ارزیابی و تحلیل نتایج بدست آمده نیز ارائه خواهد شد. در نهایت بخش پنجم به نتیجه‌گیری کلی و ارائه پیشنهاداتی برای کارهای آینده می‌پردازد.

۲- ادبیات پژوهش و کارهای مرتبط

آشنایی با مفاهیم جریان داده‌ها و یادگیری این نوع داده‌ها با مطالعه منابع مروری میسر است. برخی منابع، جریان داده و الگوریتم‌های کاوش در این زمینه را به صورت کلی توضیح می‌دهند، و برخی دیگر بر زمینه‌های خاصی تمرکز دارند.

جریان‌داده [۱۹] از حوزه‌هایی است که در سال‌های اخیر اهمیت کاربردی قابل توجهی پیدا کرده‌است. یک جریان‌داده، توالی نامحدودی از داده‌ها است. این داده‌ها از منابع مختلفی مانند شبکه‌های اجتماعی، تراکنش‌های بانکی، و شبکه‌های حسگر بی‌سیم تولید می‌شوند. می‌توان با تحلیل این‌گونه داده‌ها به کمک روش‌های داده‌کاوی اطلاعات مفیدی را استخراج نمود. در پژوهش‌های مختلف به جریان داده از منظر ارائه الگوریتمهای کاوش [۲۰، ۲۱]، یادگیری با ناظر [۲۲]، خوشه‌بندی [۲۳]، و رانش مفهوم [۲۴، ۲۵] پرداخته شده است. بطور کلی، مواردی نظیر برخط بودن، اندازه‌ی نامحدود داده‌ها، پیچیدگی در سازمان‌دهی، ماهیت پویا و افزایشی و عدم دسترسی به کل داده‌ها، از چالش‌های موجود در این حوزه است [۲۶].

رانش مفهوم در جریان داده، زمانی رخ می‌دهد که مفهوم اصلی موجود در داده‌ها در طول زمان تغییر یابد [۲۷]. اگر این رانش مفهوم با پدیدار شدن کلاس‌های جدید همراه شود، تکامل مفهوم اتفاق افتاده است. تغییر رفتار داده‌ها در طول زمان موجب شناسایی و پدیدار شدن رانش مفهوم می‌شود. با شناسایی این تغییرات، می‌توان تصمیمات بهتری در حوزه‌هایی مانند رشد جمعیت، تغییرات آب و هوایی، بازار سهام و غیره اتخاذ نمود.

با توجه به این موارد، تشخیص رانش مفهوم در جریان داده، بحثی پیچیده، بسیار اطلاع دهنده، و در عین حال چالش‌برانگیز است. فن و بیفت [۲۶] بر چالش‌هایی که با وجود حجم زیاد داده به وجود می‌آیند، تمرکز کرده‌اند. برخی رویکردهای دیگر بر تشخیص رانش مفهوم و سعی در یافتن مرزهای بهینه قطعه‌های داده در جریان داده تمرکز دارند [۲۸].

کرمپل و همکاران [۲۹] چالش‌های پژوهشی برای داده‌های در جریان را با تمرکز بر کاربردهای دنیای واقع توضیح می‌دهند. با توجه به پیچیدگی‌های بحث، اکثر کارهای انجام شده در زمینه رانش مفهوم، بیشتر به جنبه فنی روش‌ها پرداخته و استفاده از این روش‌ها برای بررسی داده‌های غیر ساخت‌یافته در کاربردهای واقعی کمتر مورد توجه قرار گرفته است. هم‌چنین پژوهش‌هایی که تغییر مفهوم را در شبکه‌های اجتماعی بررسی می‌کنند معمولا به رابطه بین این تغییر و رویدادهای جهان واقع نمی‌پردازند. به طور مثال پاتیل و همکاران [۳۰] به بررسی تغییر مفهوم در تحلیل نظرات در رسانه‌های اجتماعی می‌پردازند. آنها انواع تغییر مفهوم و تحلیل‌های مختلف را مرور می‌کنند. سوسی و شنتی [۳۱] روشی برای تشخیص و تحلیل تغییر مفهوم در جریان داده‌های شبکه‌های اجتماعی پیشنهاد می‌کنند. اوچیدا و یوشیدا [۳۲] به بررسی تغییر مفهوم در داده‌های کوید در ژاپن پرداختند و نشان دادند که پیش‌بینی تعداد موارد ابتلا بدون توجه به رانش مفهوم غیر دقیق خواهد بود. زو و ویلسن [۳۳] به بررسی تاثیر رانش مفهوم بر عملکرد پیش‌بینی در سیستم‌های هشدار مدارس متوسطه در دوران کرونا پرداختند. داک ورث و همکاران [۳۴] برای پیش‌بینی وضعیت بیماران مراجعه‌کننده به واحد اورژانس و تشخیص بیماران با ریسک بالا از روش‌های یادگیری و تغییر مفهوم استفاده کردند. آنها روش خود را برای دوران قبل از شیوع کرونا و بعد از آن ارزیابی کرده‌اند و استفاده از روش‌های تفسیر یادگیری ماشین را برای توضیح نتایج مد نظر قرار داده‌اند. مولر و سالاته [۳۵] به بررسی نظرات منعکس شده در فضای مجازی در مورد واکسن کرونا پرداختند. آنها نشان دادند با استفاده از مدل آموزش دیده با داده‌های تاریخی، بدون در نظر گرفتن رانش مفهوم، افت در احساسات کاربران در مورد واکسن قابل تشخیص نیست.

از طرف دیگر پژوهش‌های مختلفی بدون اشاره به تغییر مفهوم، به تحلیل داده‌های شبکه‌های اجتماعی در رابطه با پاندمی کرونا پرداخته‌اند. جنشر و کرن [۳۶] به بررسی نظرات کاربران در توئیتر در حین همه‌گیری کوید پرداخته و ارتباط آن را با الگوهای حرکتی افراد در مکان‌های مختلف بررسی کردند. عمیر و مسکیاری [۳۷] به بررسی قطبیت نظرات در باب واکسن کوید و ارتباط آن با مکان‌های جغرافیایی پست‌های شبکه اجتماعی پرداخته‌اند.

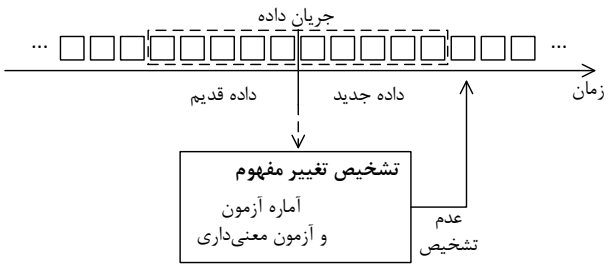
پژوهش‌هایی که به طور همزمان تغییر مفهوم در شبکه‌های اجتماعی و ارتباط آن با رویدادهای جهان واقع را در نظر بگیرند به ندرت مشاهده می‌شوند. بایر و همکاران [۳۸] به بررسی تأثیر سیاست‌های اعمال شده برای کنترل همه‌گیری بر روی شرایط بیماری مانند تعداد ابتلای کرونا پرداختند. به این منظور برای استخراج تغییرات قابل توجه در شرایط بیماری (مانند تعداد موارد ابتلا)، از روش‌های تغییر مفهوم استفاده کردند. بچینی و همکاران [۳۹] به بررسی عملکرد رده‌بندهای مختلف در تشخیص قطبیت نظرات کاربران در حضور تغییر مفهوم پرداختند. مجموعه داده مورد استفاده در پژوهش ذکر شده شامل نظرات کاربران در مورد موضوع واکسن در کشور ایتالیا است و رخدادهای جهان واقعی در این تحلیل

مد نظر قرار گرفته است. اما در این پژوهش هدف بررسی عملکرد رده‌بند بوده و رانش مفهوم در حاشیه مورد توجه قرار گرفته است که با دیدگاه پژوهش حاضر که تغییر مفهوم را مستقیما از روی سری زمانی داده‌ها استخراج می‌کند، متفاوت است. تحقیقات فوق نشان می‌دهد که توجه به رانش مفهوم در بررسی داده‌های متغیر با زمان برای ارائه یک تحلیل معتبر ضروری می‌باشد. با این وجود تحقیقاتی که به بررسی رابطه بین نظرات در فضای مجازی با رخدادهای دنیای واقع بپردازند محدود هستند. این تحقیقات می‌توانند چراغ راهی برای تأثیر متقابل این دو محیط بر یکدیگر بوده و راهنمای اتخاذ تصمیمات موثر برای تأثیرگذاری بر نظرات عمومی باشند. همچنین با بررسی عملکرد روش‌های تشخیص رانش مفهوم در چنین داده‌هایی، قابلیت‌ها و نقایص این روش‌ها بهتر آشکار خواهند شد.

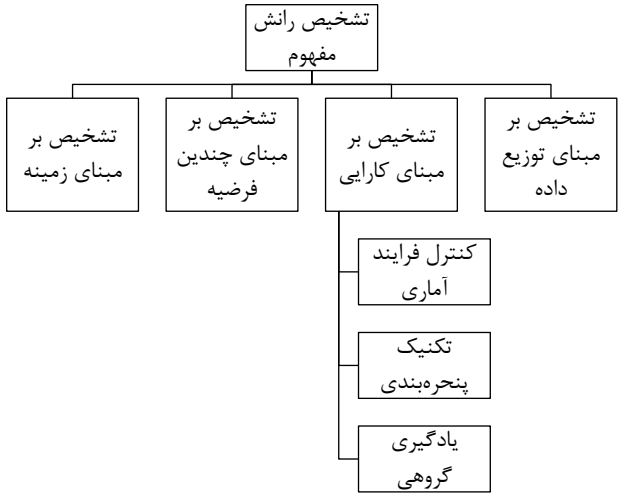
۲-۱- تشخیص رانش مفهوم

تشخیص رانش مفهوم در چند قدم کلی انجام می‌شود: (۱) استخراج داده: از آنجا که یک تک داده نمی تواند حاوی اطلاعات کافی برای تشخیص رانش مفهوم باشد، داده به صورت یک مجموعه یا بسته از جریان داده دریافت می‌شود. (۲) مدلسازی داده: در این مرحله ویژگی‌های حساس که در صورت رخداد رانش مفهوم بیشتر تحت تاثیر قرار می‌گیرند، استخراج می‌شوند. این مرحله نوعا شامل کاهش بعد یا کاهش اندازه نمونه است و اختیاری می‌باشد. (۳) محاسبه آمارهای آزمون: این مرحله نوعا شامل محاسبه عدم شباهت یا فاصله است. در این مرحله شدت رانش کمی سازی شده و آمارهای آزمون برای بررسی فرضیه تشکیل می‌شوند. این مرحله عملا چالش اصلی تشخیص رانش مفهوم را در بر می‌گیرد. (۴) آزمون فرضیه: در این مرحله از یک آزمون فرضیه خاص برای بررسی معناداری تغییر مشاهده شده در قدم سوم استفاده می‌شود. بدون این مرحله عملا آمارهای آزمون محاسبه شده در مرحله سوم بی معنا می‌شوند زیرا که امکان محاسبه بازه اطمینان رانش مفهوم وجود ندارد. به بیان دیگر باید مشخص شود که تغییر مشاهده شده بر اثر رانش مفهوم است و ناشی از نویز یا بایاس نمونه برداری نیست [۴۰]. مراحل کلی شرح داده شده در شکل ۱ نشان داده شده است [۴۱].

روش‌های تشخیص رانش مفهوم به چهار دسته کلی تشخیص بر مبنای توزیع داده، تشخیص بر مبنای کارایی، تشخیص بر مبنای چندین فرضیه، تشخیص بر مبنای زمینه[1] تقسیم می‌شوند [۴۱] (شکل ۲). روش اول از معیارهای فاصله برای تخمین شباهت بین توزیع داده‌ها در دو پنجره متفاوت استفاده کرده و اگر اختلاف معنادار باشد، رانش تشخیص داده می‌شود. این روش قابل اعمال به هر دو داده برچسب‌دار و غیر برچسب‌دار است، و در عین حال مستعد هشدارهای اشتباه می‌باشد. روش‌های مبتنی بر کارایی، خطای یادگیرنده را رصد کرده و افزایش در خطا را نشانه وقوع رانش مفهوم در نظر می‌گیرند. روش سوم، نتایج چند تشخیص‌دهنده رانش را به صورت موازی یا سلسله مراتبی ترکیب می‌کند. در نهایت روش چهارم تشخیص رانش مفهوم از داده‌های زمینه و سیستمی استفاده می‌کند. به طور مثال این روش‌ها می‌توانند از داده‌ها و روندهای گذشته برای تخمین خطا استفاده کنند.



شکل ۱: چارچوب کلی تشخیص رانش مفهوم.



شکل ۲: دسته‌بندی روش‌های تشخیص رانش مفهوم [۴۱].

بسیاری از مطالعات موجود بر دسته دوم تشخیص رانش مفهوم یعنی تشخیص بر اساس کارایی تمرکز کرده‌اند. این دسته خود به سه بخش کنترل فرایند آماری[2]، تکنیک پنجره‌بندی[3] و یادگیری گروهی تقسیم می‌شود. کنترل فرآیند آماری برای رصد تغییرات نرخ خطای برخط یادگیر(های) پایه و نظارت بر کیفیت فرآیند یادگیری استفاده می‌شود. فرض بر این است که در صورتی که کاهش عملکرد مدل از سطح آزمون معناداری فراتر رود، انحراف مفهوم رخ داده است.

یکی از روش‌های مطرح تشخیص رانش مفهوم در دسته کنترل فرایند آماری، الگوریتم DDM[4] [۴۲] است. DDM اولین روشی بود که سطح هشدار و سطح رانش را برای تشخیص رانش مفهوم مطرح

کرد. در این روش مرحله اول ذکر شده با پنجره‌های زمانی صورت می‌گیرد. وقتی نمونه داده جدیدی در دسترس قرار می‌گیرد، DDM نرخ خطای کلی برخط در پنجره را جهت زیاد شده آن به صورت معنا دار بررسی می‌کند. اگر بازه اطمینان نرخ تغییر خطای مشاهده شده به سطح هشدار برسد، DDM اقدام به ساخت یک یادگیر جدید می‌کند، در حالی که یادگیر قبلی همچنان برای پیش‌بینی استفاده می‌شود. اگر تغییرات به آستانه رانش برسد، یادگیر قبلی با یادگیر جدید جایگزین می‌شود. برای بدست آوردن نرخ خطای برخط، DDM به یک یادگیر برای پیش‌بینی نیاز دارد. این یادگیر یک مدل را در مرحله دوم از روی داده‌های آموزش ایجاد می‌کند. آمارهای آزمون در مرحله سوم همان نرخ خطای برخط است و آزمون فرضیه در مرحله چهارم با تخمین توزیع نرخ خطای برخط و محاسبه سطح هشدار و آستانه رانش انجام می‌شود.

در این مقاله از روش تشخیص رانش مفهوم HDDM [۴۳] استفاده می‌شود. این روش توسعه‌ای بر روش DDM است و از نامساوی هافدینگ برای برای تعیین ناحیه بحرانی رانش در مرحله چهارم استفاده می‌کند. در این روش کارآیی یادگیر پایه توسط نامساوی‌های احتمالاتی پایش می‌شود. در این روابط فرض بر این است که متغیرهای تصادفی، همگی مستقل، تک مقداره و کران‌دار هستند و از این طریق تضمین‌های تئوری برای تشخیص تغییرات ارائه می‌شود. HDDM برای هر دو نرخ مثبت کاذب و منفی کاذب حد ارائه می‌کند. این روش دو گونه مختلف دارد که در این مقاله از روش HDDM-A استفاده می‌کنیم.

باید توجه گردد که در یک مساله می‌توان از روش‌های مختلف تشخیص مفهوم استفاده نمود [۴۱]. از بین چهاردسته کلی روش‌های تشخیص رانش مفهوم، دسته اول (بر اساس توزیع داده) مستعد هشدارهای اشتباه است زیرا تغییرات در توزیع داده همواره منجر به کاهش عملکرد یادگیرنده نمی‌شود. روش‌های دسته سوم (بر اساس چندین فرضیه) دارای پیچیدگی در اجرا هستند. روش‌های دسته چهارم (مبتنی بر زمینه) نیز به دلیل محدودیت و وابستگی این روش‌ها به اطلاعات زمینه که بعضا به سختی قابل اکتساب است کمتر مورد استفاده قرار می‌گیرند. از طرف دیگر به دلیل در دسترس بودن برچسب داده‌ها، می‌توان از روش‌های کنترل فرایند آماری که یکی از متداول‌ترین روش‌های تشخیص رانش مفهوم هستند استفاده نمود. در مطالعات قبلی نشان داده شده است که روش HDDM در مقایسه با روش‌های دیگر تشخیص رانش مفهوم عملکرد خوبی از خود نشان می‌دهد [۴۴] و به همین دلیل در این مقاله برای بررسی داده‌های شبکه‌های اجتماعی و اخبار انتخاب شده است.

در روش HDDM، دو مقدار مختلف اطمینان متناسب با سطح *هشدار* (α_W) و سطح *رانش* (α_D) تعریف می‌شود. سه حالت مختلف برای جریان در نظر گرفته می‌شود که عبارتند از: حالت پایدار[۵] که تغییری وجود ندارد، هشدار[۶] که به نظر می‌رسد که یک رانش مفهوم احتمالی ممکن است مشاهده شود، و رانش[۷] که در آن رانش تشخیص داده شده است. بنابراین این روش یک مقدار آماری (میانگین) که از مقادیر مشاهده شده اخیر محاسبه شده است را رصد می‌کند، تا تشخیص دهد چه زمانی این مقدار از مقدار مورد انتظار به صورت معناداری فاصله گرفته است.

برای تشخیص رانش مفهوم، جریان داده را به صورت دنباله متغیر های تصادفی ... $X_1, \dots, X_n, \dots$ در نظر می‌گیریم. مساله ما یافتن یک افزایش معنادار در میانگین این داده‌ها است. اولین قدم پیش‌بینی یک نقطه برش (cut) در این دنباله است که نشان‌دهنده موقعیت تغییر وضعیت باشد. بعد از یافتن این نقطه، دنباله به صورت $X_1, X_2, \dots, X_{cut}, X_{cut+1}, \dots, X_n$ در نظر گرفته شده و می‌توان تست آماری رانش را به نحوی که در ادامه شرح داده می‌شود بر روی نمونه‌های $X_1, X_2, \dots, X_{cut}$ و $X_{cut+1}, \dots, X_n$ اجرا نمود. عبارت $\hat{X}_i + \epsilon_{\hat{X}_i} (1 \le i \le n)$ را در نظر بگیرید که در آن $\hat{X}_i$ میانگین متحرک است و $\epsilon_{\hat{X}_i}$ حد خطای متناظر آن است که از قضیه ۱ (نامساوی هافدینگ) بدست می‌آید:

قضیه ۱ (نامساوی هافدینگ) [۴۵]: فرض کنیم $X_1, \dots, X_n$ متغیرهای تصادفی مستقل با مقادیر کران‌دار $X_i \in [a_i, b_i], i \in \{1, \dots, n\}$ باشند. هم‌چنین داریم $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. آنگاه برای هر $\epsilon > 0$:

$$\Pr\{\overline{X} - \mathrm{E}[\overline{X}] \ge \epsilon\} \le \mathrm{e}^{-\frac{2\epsilon^2 \mathrm{n}^2}{\sum_{\mathrm{i}=1}^{\mathrm{n}} (\mathrm{b_i} - \mathrm{a_i})^2}} \qquad (۱)$$

با توجه به قضیه فوق، می توان خطای میانگین را با سطح معناداری حداکثر δ محاسبه نمود:

$$\epsilon_\delta = \sqrt{\frac{1}{2n} ln \frac{1}{\delta}} \qquad (۲)$$

وقتی تغییری در میانگین جمعیت رخ نمی‌دهد، $\hat{X}_i$ باید تقریبا ثابت باشد و مقدار $\hat{X}_i + \epsilon_{\hat{X}_i}$ رو به کاهش است. وقتی میانگین رو به افزایش باشد، تخمین میانگین یا همان $\hat{X}_i$ و مقدار $\hat{X}_i + \epsilon_{\hat{X}_i}$ باید افزایش یابد. بنابراین نقطه برش متناظر را می‌توان از مقدار کمینه $\hat{X}_i + \epsilon_{\hat{X}_i}$ استخراج نمود. سپس می‌توان از تست A برای تخمین وضعیت جریان (پایدار، هشدار و رانش) بر روی دنباله‌های متناظر $X_1, X_2, \dots, X_{cut}$ و $X_{cut+1}, \dots, X_n$ استفاده نمود. در این صورت اگر فرضیه صفر با با اندازه α_W رد شد، وضعیت جاری به هشدار مقداردهی می‌شود. و اگر با فرضیه صفر با اندازه α_D رد شد، وضعیت

برابر سطح رانش شده و تمامی شمارنده‌ها مقداردهی مجدد می‌شوند. اگر فرضیه صفر پذیرفته شود، وضعیت پایدار خواهد بود. برای تست A از قضیه فرعی زیر که از نامساوی هافدینگ استنباط شده‌است، استفاده می‌شود:

قضیه فرعی ۱ [۴۳]: اگر $X_1, \dots, X_n, X_{n+1}, \dots, X_{n+m}$ متغیرهای تصادفی مستقل با مقادیر در بازه $[a,b]$ باشند، و اگر $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ و $\bar{Z} = \frac{1}{n+m}\sum_{i=1}^{n+m} X_i$، آنگاه برای $\epsilon > 0$:

$$\Pr\{\bar{X} - \bar{Z} - (E[\bar{X}] - E[\bar{Z}]) \geq \epsilon\} \leq e^{-\frac{2\epsilon^2 n(n+m)}{m(b-a)^2}} \qquad (۳)$$

در این صورت قانون نپذیرفتن فرضیه صفر $H_0: E[\bar{X}] \leq E[\bar{Z}]$ در مقابل فرضیه دیگر $H_1: E[\bar{X}] > E[\bar{Z}]$ برابر $\bar{X} - \bar{Z} \geq \epsilon_\alpha$ خواهد بود که در آن:

$$\epsilon_\alpha = (b-a)\sqrt{\frac{m}{2n(n+m)} ln\frac{1}{\alpha}} \qquad (۴)$$

با توجه به مطالب بیان شده، برای اجرای تست آماری A، دو شمارنده $\hat{Z}_n$ و $\hat{X}_{cut}$ نگه‌داری می‌شوند. با هر مشاهده هر داده جدید، مقدار میانگین داده‌ها تا کنون $\hat{Z}_n$ و خطای آن $\epsilon_{\hat{Z}_i}$ به روز می‌شود. اگر جمع این میانگین و خطا کم‌تر از $\hat{X}_{cut} + \epsilon_{\hat{X}_{cut}}$ فعلی باشد، $\hat{X}_{cut}$ و خطای آن با مقادیر جدید به روز می‌شوند. هم‌چنین با ورود هر داده، تست آماری A با دو سطح α_D و α_W اجرا می‌شود و وضعیت جریان به یکی از حالات رانش، هشدار یا پایدار مقداردهی می‌شود. در صورتی که وضعیت رانش باشد، تمامی شمارنده‌ها بازنشانی می‌شوند. به این صورت نقاط رانش در جریان مشخص می‌شوند.

۳- روش پیشنهادی

در این بخش به معرفی معماری پیشنهادی جهت تشخیص رانش مفهوم و نقاط کلیدی رانش در دادگان شبکه‌های اجتماعی در بازه ۱۰۰ روزه ابتدایی شیوع کرونا با استفاده از روش HDDM می‌پردازیم. بدین منظور، ابتدا به معرفی دادگان مورد استفاده و ویژگی‌های آن می‌پردازیم. سپس در ادامه به تشریح معماری روش پیشنهادی خواهیم پرداخت.

۳-۱- معرفی دادگان

جامعه آماری این پژوهش، پیام‌های سه شبکه اجتماعی تلگرام، اینستاگرام و توئیتر و متون فارسی خبری مرتبط با کرونا در بازه ۱۰۰ روزه ابتدایی شیوع کروناست. در بررسی یک واقعه و اثرات آن، در نظر گرفتن یک بازه زمانی، با شروع از ابتدای وقوع واقعه دارای اطلاعات مفیدتری است و هر چه از تاریخ شروع واقعه دور می‌شویم بررسی، اطلاعات کمتری را بدست خواهد داد. در بازه زمانی مورد بررسی این پژوهش، به دلیل مواجهه جامعه با یک بیماری ناشناخته، و مبهم بودن منشا بیماری به همراه روش‌های پیشگیری و درمان آن، و هم‌چنین تصمیم‌های جدید و مختلف مسئولین برای کنترل شرایط بیماری، تغییرات زیادی در نظرات افراد در رابطه با بیماری مشاهده می‌شود. به همین دلیل انتخاب این داده‌ها، از یک سو زمینه مناسبی برای به کارگیری و ارزیابی روش‌های تشخیص رانش مفهوم فراهم می‌کند، و از سوی دیگر نسبت میان تغییرات در نظرات فضای مجازی نسبت به رخدادهای جهان واقع را بررسی می‌کند. همچنین، مطابق آنچه در ادامه در مورد نتایج بدست آمده از تشخیص رانش مفهوم خواهیم دید، تغییر تدریجی شیوع مفاهیم منفی، به مفاهیم خنثی یا مثبت در بازه مورد بررسی، می‌تواند حاکی از اهمیت و تا حدودی کافی بودن این بازه زمانی برای بررسی این واقعه باشد.

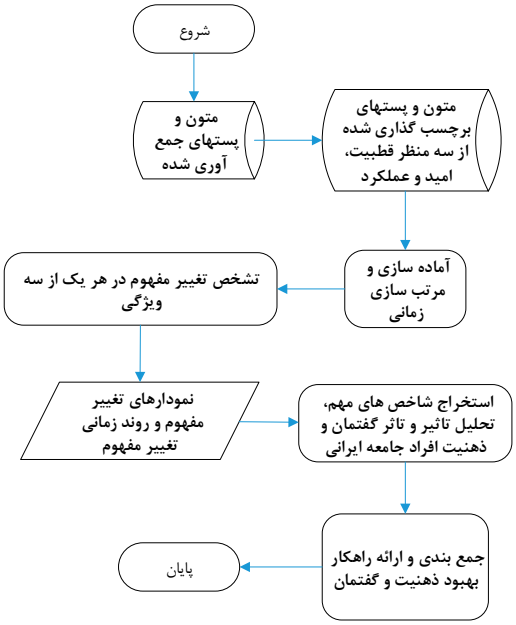
دادگان مرتبط با جامعه آماری مورد استفاده در این مقاله، دادگانی است که از سوی ستاد علوم شناختی در اختیار قرار گرفته است. این مجموعه بالغ بر ۲۲ میلیون پست و خبر منتشر شده در فضای مجازی با محوریت کرونا است که توسط آزمایشگاه شبکه‌های اجتماعی دانشگاه تهران در بازه زمانی ۱ بهمن ۱۳۹۸ تا ۱۰ اردیبهشت ۱۳۹۹ جمع‌آوری شده است. این مطالب از کانال‌های عمومی تلگرام، صفحات عمومی اینستاگرام و توئیتر و همچنین اخبار داخلی منتشر شده به صورت آنلاین استخراج شده است. برچسب‌گذاری موضوعی و هیجانی بصورت دستی برای حدود ۶۰۰۰ داده که بصورت تصادفی انتخاب شده‌اند توسط آزمایشگاه پردازش زبان طبیعی دانشگاه شهید بهشتی انجام شده است.

در مورد دادگان برچسب خورده، ۶۰ برچسب روی این دادگان لحاظ شده که می‌توان آنها را در چندین دسته‌بندی کلی برچسب‌های اجتماعی، سیاسی، هیجانی احساسی قرار داد. برخی از این ویژگی‌ها، ویژگی تک مقداره (بدین معنا که مفهوما در آنِ واحد تنها یک مقدار از میان مجموعه مقادیر دامنه خود را می‌توانند کسب کنند) و برخی چند مقداره هستند. ویژگی قطبیت بعنوان یک ویژگی تک‌مقداره با طول دامنه سه است که می‌تواند دارای یکی از مقادیر مثبت، منفی یا خنثی باشد. ویژگی عملکرد مسئولین ، یک ویژگی تک مقداره با طول دامنه دو است. عناصر این دامنه عبارتند از: عملکرد مثبت و عملکرد منفی. ویژگی امید نیز ویژگی تک مقداره با مقادیر مثبت و منفی است. در این مقاله، با استفاده از روش‌های تشخیص رانش مفهوم به تحلیل و رصد نحوه تغییر این سه ویژگی در دادگان مورد بررسی خواهیم پرداخت. جدول ۱ تعداد دادگان برچسب خورده برای هر یک از سه ویژگی را نشان می‌دهد. در مورد هر ویژگی، از تمامی نمونه‌های برچسبدار استفاده شده است. همچنین با توجه به اینکه

دادگان کلی (۲۲ میلیونی) دارای برچسب زمانی هستند، اما در دادگان مربوط به ویژگی‌های برچسب خورده (زیرمجموعه‌ای از کل دادگان ۲۲ میلیونی)، ستون زمان وجود ندارد، با یافتن اشتراک متون و پست‌های برچسبدار با متون دادگان کلی، به یافتن برچسب زمانی آنها می‌پردازیم. سپس به مرتب‌سازی زمانی دادگان مربوط به هریک از سه ویژگی می‌پردازیم. بدین ترتیب جریان داده مربوط به هر یک از ویژگی‌ها را بازسازی می‌کنیم.

جدول ۱: آمار مربوط به تعداد پست‌های برچسب دار مربوط به هر ویژگی در دادگان.

تعداد کل دادگان: بالغ بر ۲۰ میلیون متن			
تعداد کل دادگان برچسبدار: ۶۰۰۰			
ویژگی	قطبیت	عملکرد مسئولین	امید
تعداد با برچسب ویژگی	۳۷۵۷	۷۲۵	۷۸۵



شکل ۳: فلوچارت روش پیشنهادی تحلیل قطبیت، امید و ذهنیت جامعه ایران با استفاده از تشخیص رانش (تغییر) مفهوم.

۳-۲- معماری پیشنهادی

معماری پیشنهادی برای تشخیص رانش مفهوم سه ویژگی قطبیت، امید و عملکرد مسئولین در شکل ۳ نشان داده شده است. روال تشخیص رانش مفهوم، روشی با ناظر بوده و از دادگان برچسب‌گذاری شده استفاده می‌نماید. بنابراین در مورد سه ویژگی مذکور از برچسب مرتبط موجود در دادگان برچسب خورده استفاده می‌شود.

همانگونه که در فلوچارت شکل ۳ مشخص است، با داشتن متون و پست‌ها، برچسب زمانی آنها و برچسب مرتبط با ویژگی در آنها، به تشخیص نقاط زمانی رانش مفهوم برای هر یک از سه ویژگی با استفاده از روش HDDM می‌پردازیم. سپس به ترسیم نمودارهای مربوط به هر ویژگی می‌پردازیم. در نهایت، با بررسی دادگان حول نقاط رانش مفهوم (مثبت یا منفی) به تحلیل علل و ارائه راهکارهای مرتبط با دادگان خواهیم پرداخت. این راهکارها با توجه به منبع دادگان (متون خبری یا پست‌های شبکه‌های اجتماعی) به دو دسته راهکار مسئولین و راهکارهای مردمی دسته‌بندی می‌شود که در ادامه مقاله به تفصیل به آنها اشاره خواهد شد.

۴- نتایج

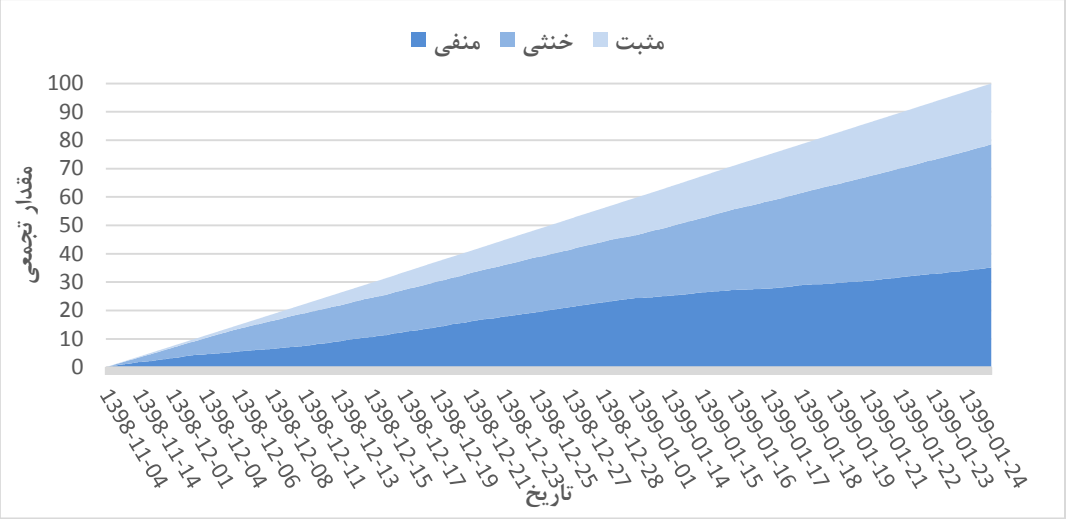
در این بخش به بررسی نتایج بدست آمده از روش‌های رانش مفهوم و تحلیل آنها در ویژگی‌های تک مقداره (قطبیت، امید و عملکرد) در داده‌های مورد بررسی در طول بازه ۱بهمن ۱۳۹۸ تا ۱۰ اردیبهشت ماه ۱۳۹۹ می‌پردازیم. لازم به ذکر است، به‌منظور استخراج اطلاعات مبتنی بر روش‌های رانش مفهوم در مورد سه ویژگی مورد بررسی از کتابخانه scikit-multiflow استفاده شده است.

۴-۱- بررسی رانش مفهوم قطبیت داده‌ها

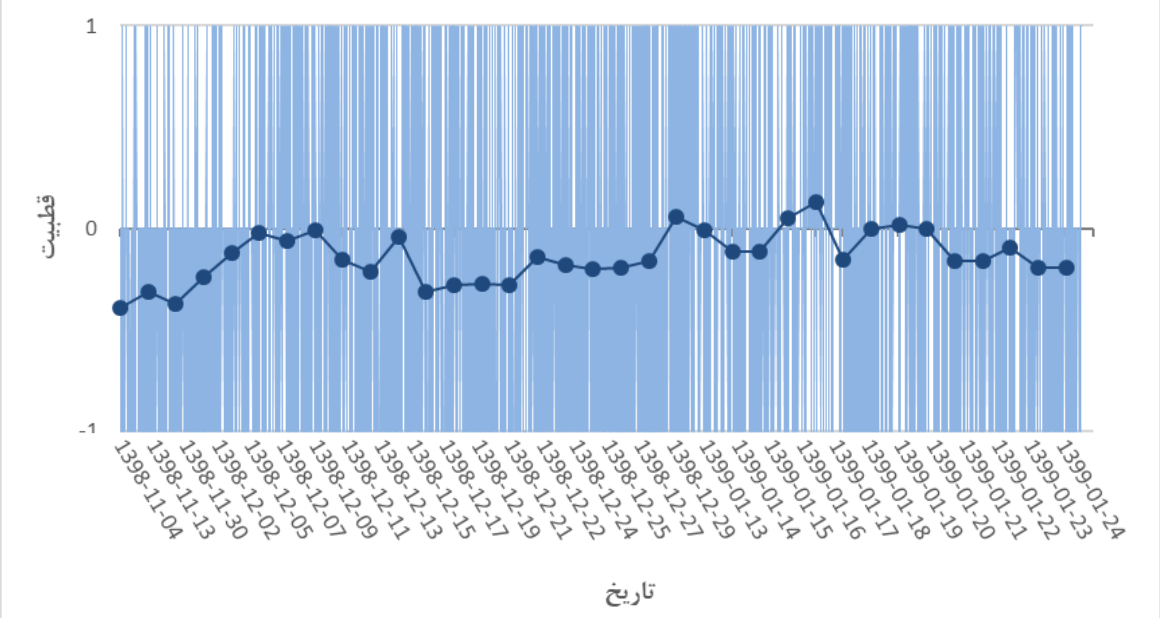
به منظور بررسی رانش مفهوم در ویژگی قطبیتِ متون خبری و پست‌های شبکه‌های اجتماعی، ابتدا نمودار تجمعی قطبیت داده‌ها در طول بازه مورد بررسی در شکل ۴-الف نمایش داده شده است. در این نمودار مجموع تعداد پست‌ها در هر برچسب قطبیت از ابتدای بازه مورد بررسی در طول زمان نشان داده شده است. همان‌طور که نشان داده می‌شود در انتهای بازه تعداد پست‌های خنثی از همه بیشتر بوده و بعد از آن پست‌های منفی بیشترین تعداد را دارند. در ابتدای شروع کرونا، قطبیت منفی و خنثی غالب است و این غلبه با شدت کمتر و بیشتر در طول دوره مشاهده می‌شود. در عین حال شیب رشد تعداد پست‌های مثبت و خنثی از پست‌های منفی بیشتر است. نمودار تغییرات قطبیت به‌همراه میانگین متحرک توالی‌های ۱۰۰ تایی داده‌ها در نمودار شکل ۴-ب نیز موید همین مطلب است.

در این نمودار و بطور کلی در نمودارهای تجمعی تغییرات هر ویژگی به‌همراه میانگین متحرک، ابتدا با خطوط آبی رنگ نمودار خطی ویژگی در هر نقطه زمانی نشان داده شده است. سپس با نقاط مشکی میزان میانگین متحرک نشان داده شده و بصورت خطی متصل شده است. در نمودار شکل ۴-ب، هر نقطه نشان دهنده میانگین قطبیت ۱۰۰ پست با در نظرگرفتن اعداد ۱-، ۰ و ۱ بترتیب برای قطبیت منفی، مثبت و خنثی در هر پست است.

همانطور که مشاهده می‌شود مطالب منتشر شده در رابطه با بیماری کرونا بیشتر در حیطه قطبیت منفی بوده اما در طول زمان با شیب کندی به سمت خنثی پیش می رود. بنابراین تغییر این ذهنیت و گرایش دادن قطبیت به سمت مثبت دارای اهمیت زیادی برای بهبود روحیه و رفتار جامعه است. نوسات زیاد موجود در نمودار را به دو صورت می‌توان توجیه کرد. اولی یک بحث اجتماعی و مقارن با نوسانات قطبیت دیدگاه افراد جامعه در مورد مسئله با توجه به حوادث و وقایع اجتماعی یا سیاسی پیرامونی است.



الف)



ب)

شکل ۴: الف) نمودار تجمعی قطبیت در طول بازه مورد بررسی.
ب) نمودار تغییرات قطبیت به‌همراه میانگین متحرک قطبیت بر روی گروه‌های ۱۰۰ تایی داده در طول زمان.

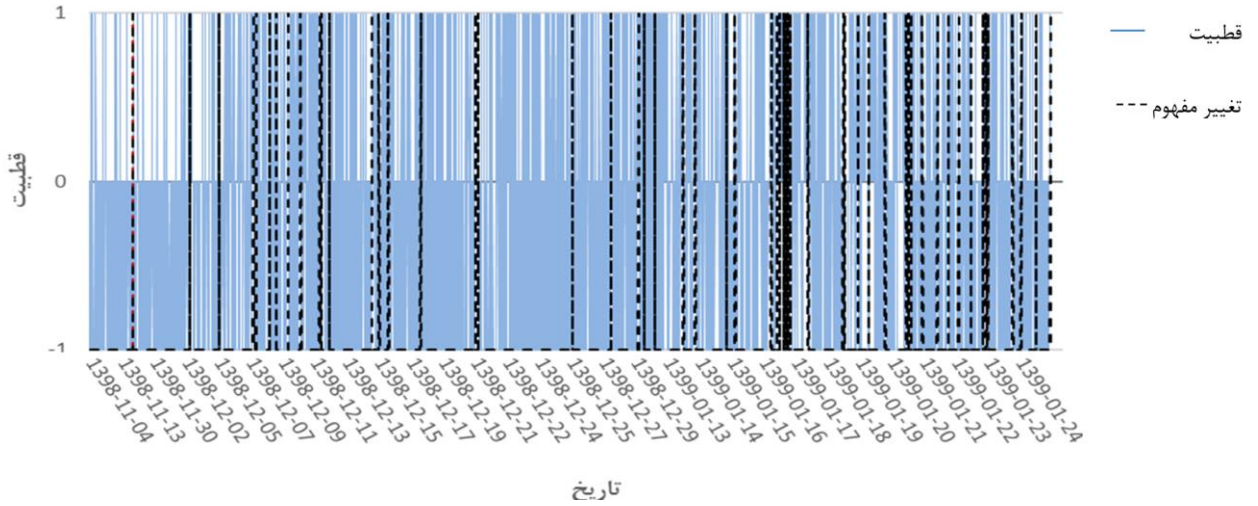
ثانیا از منظر تکنیکال و داده‌ای، ویژگی قطبیت در دادگان مورد استفاده داده برچسب خورده بیشتری (حدود ۴.۵ برابر) نسبت به دو ویژگی دیگر دارد. بنابراین نوسانات احساسی در این ویژگی بیشتر دیده می‌شود. این امر در ادامه و با ارائه نمودار میانگین متحرک تجمعی برای دو ویژگی دیگر مشهود خواهد بود.

بحث دیگر، پرداختن به نقاط مرتبط با نوسانات احساسی شدید و تحلیل وقایع مربوط به نقطه زمانی این نوع نوسان است. همانگونه که در شکل مشاهده می‌شود، مقارن با شروع سال نو ۱۳۹۹ شمسی و فراغ خاطر حاصل شده در خلال تعطیلات نوروزی، میانگین قطبیت (که تقریبا در تمام بازه زمانی منفی بود) به یکباره مثبت و بالای صفر می‌شود. در واقع می‌توان گفت قطبیت دیدگاه‌های افراد جامعه در این بازه زمانی به سمت خنثی و مثبت تمایل پیدا می‌کند. در ادامه با قرار گرفتن در پیک بیماری در اواخر فروردین، قطبیت نظرات مجدد به سمت منفی گرایش پیدا می‌کند.

در ادامه برای بررسی بیشتر رانش مفهوم در این داده‌ها، الگوریتم HDDM-A بر روی داده‌های دارای قطبیت اجرا شده است. این الگوریتم ۵۶ مورد تغییر را در این داده ها ثبت کرده است. تغییرات قطبیت نظرات به همراه موقعیت‌هایی که الگوریتم تشخیص رانش مفهوم به عنوان تغییر در مفهوم تشخیص داده است در شکل ۵ نشان داده شده‌اند. در این شکل و بطور کلی در نمودارهای مرتبط با رانش مفهوم در هر یک از سه ویژگی مورد بررسی، بتدا با خطوط آبی رنگ نمودار خطی ویژگی در هر نقطه زمانی نشان داده شده است. سپس نقاط مربوط به رانش مفهوم با نمودار خطی مشکی نقطه چین نمایش داده شد. همان طور که مشاهده می‌شود، در ابتدای ظهور بیماری، رانش مفهوم به کرات رخ داده است، اما با نزدیک شدن به تعطیلات سال نو، رانش مفهوم کمتری مشاهده می‌شود. همچنین حجم تغییرات مفهوم در بعد از تعطیلات عید که همزمان با اوج گیری بیماری و موج ابتلا بود بسیار زیاد شده است. در واقع می‌توان گفت در هنگام رخداد اجتماعی، نوسانات بیشتری در قطبیت رخ می‌دهد و قطبیت کلی نیز به سمت منفی گرایش پیدا می‌کند. این امر می‌تواند نشاندهنده لزوم توجه بیشتر به جهت‌دهی و تسکین افکار عمومی در مواقع اوج بحران باشد.

با بررسی موشکافانه محتوای منتشر شده قبل و بعد از موقعیت رانش مفهوم، می‌توان به دلیل تغییر در ذهنیت افراد و قطبیت آن پی برد. بنابراین رانش مفهوم ابزاری موثر برای درک و تحلیل دلیل رفتار افراد و تغییرات آن است. در ادامه این بخش در جدول ۲ تعدادی از این موارد را بررسی می‌کنیم. با توجه به غلبه قطبیت منفی و خنثی، تمرکز بیشتر بر روی مواردی است که قادر بوده‌اند تغییر را به سمت قطبیت مثبت سیر دهند. در این جدول (و جداول مشابه در مورد دو ویژگی دیگر)، در نقاط مهمی که تغییر مفهوم توسط الگوریتم HDDM شناسایی شده است، متون شبکه اجتماعی و قطبیت آنها مورد بررسی قرار گرفت.

تاریخ انتشار متون، نوع تغییر شناسایی شده توسط الگوریتم (بعنوان مثال از منفی به خنثی)، قطبیت هر یک از متون مورد بررسی در تاریخ مورد نظر، و همچنین نکات مهم این متون که باید مورد توجه قرار گیرد، بصورت مجزا در ستون‌هایی از جدول بیان شده است. در مورد ستون آخر، نکات صرفا با بررسی متون کلیدی در نقطهٔ تغییر مفهوم، بدست آمده‌اند.



شکل ۵: نمودار تغییرات قطبیت در مجموعه داده به همراه موقعیت‌های رانش مفهوم نشان داده شده با خط چین.

جدول ۲: نتایج بررسی رانش مفهوم در قطبیت مجموعه داده (مثبت: +، منفی: -، خنثی: ۰).

تاریخ	نوع تغییر	+- ۰	نمونه محتوا	نکته مورد توجه در تغییر مفهوم
۹۸/۱۱/۱۶	از خنثی و منفی به مثبت	۰	سلامتی کرونا که نگفت کروأنا گفت کرونا	- بازگویی عملکردهای اجتماعی در قالب طنز + دسترسی به امکانات درمان + همدلی و عواطف انسانی
		-	ویروس #کرونا رافقط چینی ها به آن مبتلا میشوند ،،چون هرنوع گوشتی رامصرف میکنند...	
		+	نخستین بخش ویژه برای کنترول ویروس کرونا در بیمارستان افغان–جاپان در کابل به بهره برداری سپرده شد.	
		+	وقتی هیولایی بین آغوش ها فاصله می اندازد و باید از راه دور عزیزانت در آغوش بکشی	
۹۸/۱۲/۰۵	از خنثی و منفی به مثبت	-	به عقیده من هم اگر در های حرم را ببندید به ساحت مقدس اسلام خیانت کرده اید	+ خاطره‌سازی و طنزپردازی
		+	در باب این ویروس جدید و راه‌های انتقال، یه خاطره تعریف کنم براتون	
۹۸/۱۲/۰۷	از خنثی به مثبت	+	اگه خدایی نکرده قرنطینه شدیم اینطوری محیط رو شاد کنیم	+ تقویت روحیه، پس از موج ناامنی و ترس با اعلام شیوع ویروس + اعتمادسازی و دسترسی به امکانات + تقویت عرق ملی
		+	امادگی ارتش برای ضدعفونی کردن شهرها	
		+	ایران عزیز ما همیشه در دامن خویش جوانان غیور و از خودگذشته پرورانیده	
۹۸/۱۲/۱۱	از منفی به مثبت	-	مجلس تا اطلاع بعدی تعطیل شد. عباسی، سخنگوی هیات رئیسه مجلس: به دلیل شیوع کرونا ویروس در کشور با تصمیم هیات رییسه، جلسات مجلس تا اطلاع ثانوی برگزار نمی شود. بودجه ۹۹ را نیز به صورت چند دوازدهم تصویب خواهیم کرد.	- به تعلیق درآمدن ساختارهای کلان و تاثیر کرونا بر اختلال و بی نظمی . + توصیه‌های بهداشتی + تقویت روحیه + همبستگی اجتماعی برای مبارزه با کرونا
		+	دلخوشیم به مدافعان غیوری که در هر میدانی، شجاعانه از هموطنان خود حمایت و مراقبت میکنند. #همه_با_هم_علیه_کرونا	
		+	...حتی الامکان اگرکارضروری نداشته باشیم تو خونه بمونیم و در صورتی که مجبور به حضور درمکان‌های عمومی هستیم حتما نکات بهداشتی...	
		+	کرونا رو با روحیه بالای خودمون شکست میدیم	
۹۸/۱۲/۱۲	از منفی به مثبت	-	ثاخت خوبه کجا۱۰ملیون صف خرید بود.کروناداره فلج میکنه مملکتو	- نگرانی‌های اقتصادی در مورد کرونا + هماهنگی بین سازمان‌ها در کمک به کمپین در خانه بمانیم.
		+	چکیده: کارت‌های بانکی مشتریان بانک ملی ایران که تاریخ انقضای آنها بهمن‌ماه بود، به‌صورت خودکار تمدید شد	
۹۸/۱۲/۲۵	از منفی به مثبت	-	... این طلاب در چنین لحظه حساسی خود را به کمک ما و بیماران رساندند،	- یادآوری شرایط قبلی - شخصیت‌سازی
		+	امسالو بخاطر کرونا نشد. ماهی بخریم به جاش نقاشیشو کشیدم بزارم سفره هفت سین.. شماهم اگه دوس داشتین بکشید	
۹۸/۱۲/۲۹ ۹۹/۰۱/۰۱	مثبت به منفی	+	حرف هایم را با تسلیت شروع کنم ما عزیزان زیادی را از دست دادیم	+ پیشنهاد مثبت - بیان منفی اخبار
		-	تصاویر ۶۵تن از متخصصان در اثرابتلا به کرونا با کمترین امکانات	
۹۹/۱/۱۶	خنثی به مثبت	۰	ازفعالیت شهرداری مازندران در پیشگیری از شیوع کرونا قدردانی کرد	+ قدردانی + دسترسی به امکانات
		+	جشنی متفاوت در معابر اصلی و فرعی هشتگرد با همکاری شهرداری	
۹۹/۰۱/۲۱	مثبت به خنثی	+	گرما قاتل ویروس کروناست؟	+ تقویت روحیه و شادسازی
		۰	پوتین: در صورت لزوم برای مبارزه با کرونا از منابع نظامی استفاده می‌کنیم ؟	

همانطور که مشاهده می‌شود، با وجود غلبه قطبیت منفی مخصوصا در ابتدا و اواسط دوره مورد بررسی، تلاش‌های مثبتی در راستای تغییر این رویه انجام شده است. محتوای اخبار نسبت به شبکه های اجتماعی بیشتر به اطلاعات جهان واقع نزدیک است، زیرا که شبکه‌های اجتماعی بیشتر منعکس کننده نظرات شخصی افراد بوده، اما اخبار به بازتاب رویدادها می‌پردازد. این امر می‌تواند منجر به دریافت اطلاعات اشتباه گردد [۱]. بنابراین ضروری است که این مهم در جامعه‌های با ضریب نفوذ بالای شبکه‌های اجتماعی باید مد نظر قرار گرفته و از جهت دهی مناسب محتوا در این بستر غفلت نشود. در این صورت می‌توان ذهنیت جامعه را به صورت منطقی و نه سهلگیری بیماری به سمت مثبت هدایت نمود.

۴-۲- بررسی رانش مفهوم در ویژگی عملکرد مسئولین

نمودار تجمعی عملکرد و میانگین توالی‌های ۱۰ تایی داده‌ها در طول بازه مورد بررسی در شکل ۶ نمایش داده شده است. همان طور که مشاهده می‌شود در بازه مربوط به اواسط بهمن‌ماه تا اواسط فروردین‌ماه، نگرش منفی به عملکرد وجود دارد. این بازه مربوط به بازه اعلام شیوع ویروس کرونا در ایران تا مهار نسبی کرونا در موج اول است. همانگونه که در شکل ۷ مشخص شده است، تغییر مفهوم در اواخر اسفندماه رخ داده است. اگر این موقعیت را در نمودار میانگین متحرک عملکرد بررسی کنیم، تغییرات سازگاری مشاهده می‌شود. میانگین دیدگاه به عملکرد در ابتدا منفی بوده و در اواخر اسفندماه مثبت شده است. با اتمام تعطیلات سال نو میانگین این دیدگاه مجدد منفی شده است. می‌توان مثبت شدن دیدگاه را به اعمال پروتکل‌ها در اواخر سال برای جلوگیری از شیوع کرونا و کاهش میزان شیوع بیماری و مرگ و میر ناشی از آن مرتبط دانست. بررسی متون قبل از زمان رانش مفهوم، نشان می‌دهد علت عمده دیدگاه منفی، علل اقتصادی یا ناهماهنگی دستگاه‌هاست. البته از اواخر اسفند عملکرد منفی کم‌کم میل به سوی مثبت دارد. که با بررسی اخبار می‌توانیم ببینیم که اختصاص بودجه جداگانه برای مبارزه با کرونا، کنترل ورود خروج اتومبیل‌ها به شهرها، اختصاص بسته حمایتی دولت به اقشارکم درآمد برای جبران خسارت های ناشی از بحران کرونا، راه‌اندازی سامانه سلامت برای جلوگیری از ازدحام جمعیت در مراکز بهداشتی، ضدعفونی معابر و مراکز شهر و توزیع بسته‌های بهداشتی، تعطیلی مراکز آموزشی در تغییر ذهنیت کنشگران شبکه‌های مجازی در ارتباط با عملکرد دولت موثر بوده است. نمونه‌ای از داده‌های قبل و بعد این تغییر را بررسی می‌کنیم.

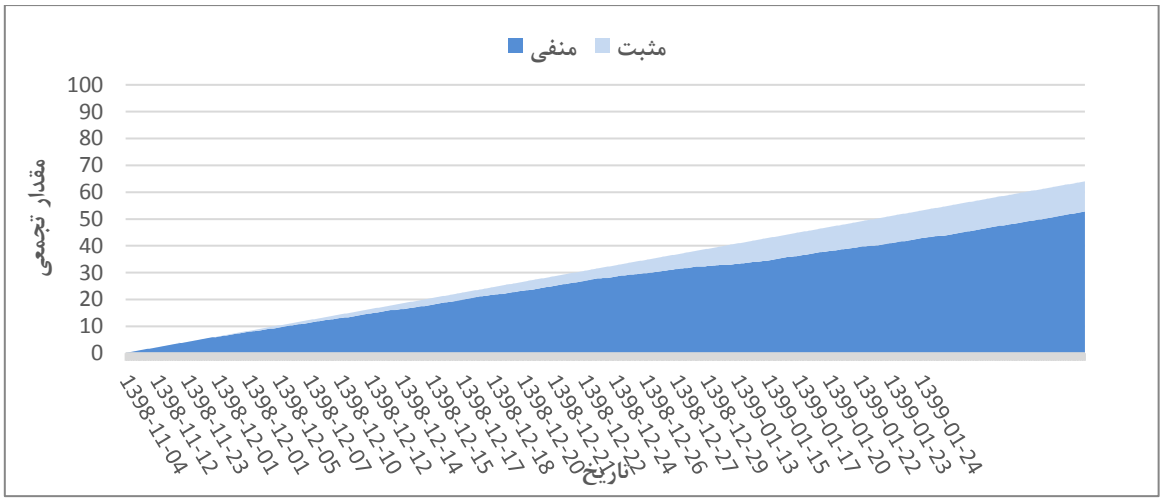
۴-۳- بررسی رانش مفهوم در ویژگی امید

در ادامه، به بررسی رانش مفهوم در ویژگی امید می‌پردازیم. نمودار تجمعی مربوط به این ویژگی و نمودار تغییرات میانگین توالی‌های ۱۰ تایی داده‌های امید در شکل ۸ آمده است.

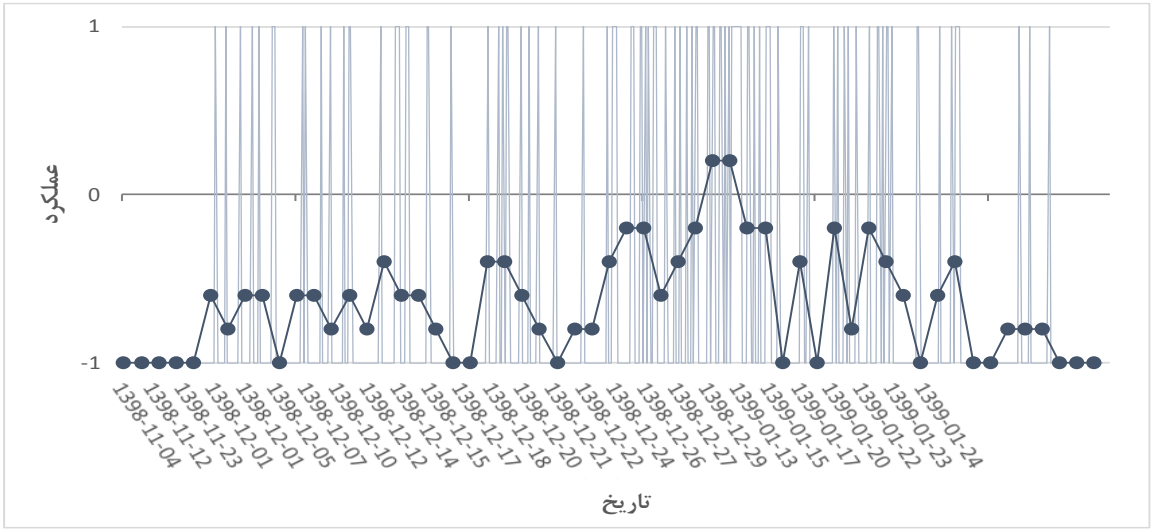
هیجانات منتج از ویروس کرونا در ابتدا عمدتا منفی بود و با به اشتراک گذاشتن این هیجانات در رسانه‌های جمعی در واقع نوعی همه‌گیری احساسات ناخوشایند و منفی در سطح جامعه شکل گرفت. اما همانطور که در نمودار مشاهده می‌کنیم سطح امید جامعه به سرعت رو به افزایش نهاده و تا قبل از موج بیماری در نیمه دوم فروردین ۱۳۹۹ مثبت بوده است. بعد از این موج میانگین امید در جامعه منفی شده است. بطور کلی، بر خلاف نمودار تجمعی میانگین دو ویژگی دیگر، این ویژگی در اکثر زمان‌ها مثبت بوده است. این نکته امیدوار کننده‌ای است که باید بر آن تاکید داشت که برخلاف آنچه در دیدگاه کاربران به مسئله کرونا و عملکرد مسئولین دیده می‌شود، افراد جامعه ایرانی امید دارند و به پشتوانهٔ آن امید توانستند این چالش غیر مترقبه را به خوبی پشت سر بگذارند.

موقعیت‌های تغییر مفهوم در داده های امید در شکل ۹ مشاهده میشود. تغییر مفهوم در اوایل بازه مورد بررسی و بعد از تعطیلات سال نو در اواخر فروردین‌ماه رخ داده است. اگر این موقعیت‌ها را در نمودار میانگین متحرک امید بررسی کنیم، تغییرات سازگاری مشاهده می‌شود. میانگین امید در ابتدا منفی بوده و کمی بعد از شیوع کرونا مثبت شده است. در اواخر فروردین‌ماه مجددا میانگین امید به سمت منفی گرایش پیدا کرده است

بررسی متون نمونه در زمان تغییر مفهوم در جدول ۴ نشان می‌دهد که مواردی مثل رفتارهای نامناسب افراد در جامعه در زمانی که همگان نیاز به همبستگی و مشارکت موثر را احساس می‌کنند، باعث کاهش امید شده است. همچنین طولانی شدن دوره بیماری و تعطیلی کسب و کارها و قرنطینه افراد موجب کاهش امید شده است. بنابراین لازم است که در چنین مواردی با افراد خاطی برخورد مناسب صورت گیرد، و همچنین در زمان طولانی شدن مواردی مثل پروتکل‌های حاوی قرنطینه، توضیحات مناسب به همراه دورنمای واقعی از آینده در اختیار مردم قرار گیرد. انتشار پیام‌های مثبت به همراه موفقیت‌ها در مبارزه با بیماری نیز در افزایش امید موثر نشان داده شده است.

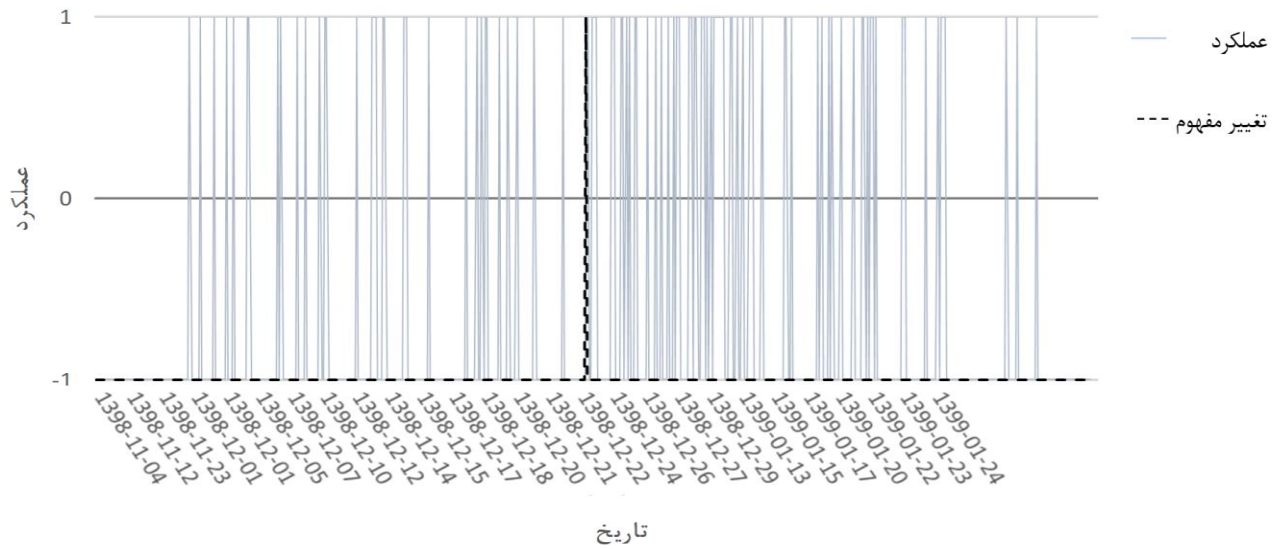


الف)



ب)

شکل ۶: الف) نمودار تجمعی ویژگی نگرش در مورد عملکرد در طول بازه مورد بررسی.
ب) نمودار تغییرات ویژگی نگرش به‌همراه میانگین متحرک بر روی گروه‌های ۱۰ تایی عملکرد.



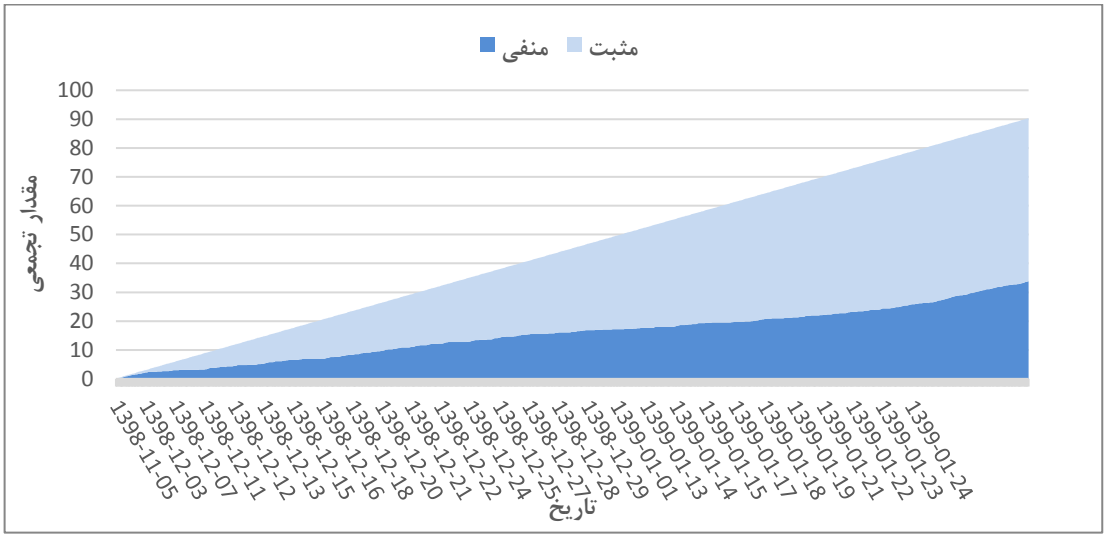
شکل ۷: نمودار تغییرات ویژگی نگرش در مجموعه داده به همراه موقعیت‌های رانش مفهوم نشان داده شده با خط چین.

جدول ۳: نمونه متون موجب رانش مفهوم ویژگی نگرش به عملکرد مسئولین در میان داده‌ها.

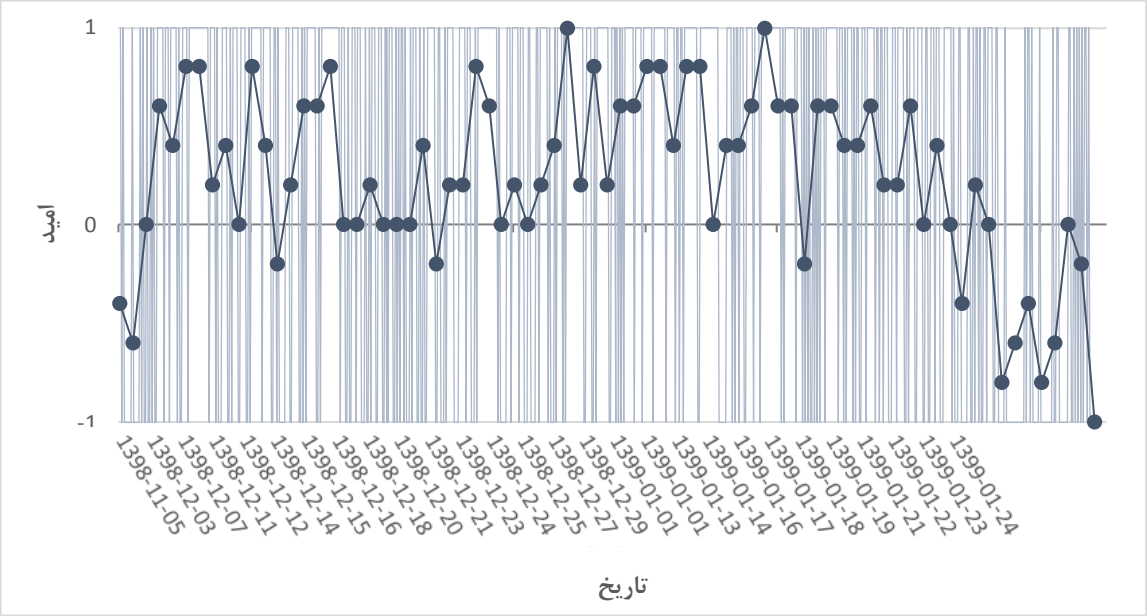
تاریخ تغییر	نوع تغییر	+۰	نمونه محتوا	موضوع مورد توجه در تغییر
۹۸/۱۲/۲۳	از منفی به مثبت	–	چکیده: برخلاف تصمیم مسئولان کشوری بر بازگشتن ساعات اداری به روال قبل از امروز شنبه، ادارات دولتی و خصوصی و همچنین بانک‌های جزیره امروز تعطیل بودند. متن: کد ویدیو دانلود فیلم اصلی**	- عدم هماهنگی دستگاه‌ها در تعطیلی یا بازگشایی + اختصاص بسته حمایتی دولت به اقشار کم درآمد + تعطیلی مراکز
		–	کرونا در ماکو یکی از شهروندان از عدم شفافیت در اطلاع رسانی انتقاد می‌کند!**	
		+	چکیده: رئیس سازمان برنامه و بودجه اعلام کرد: ۲ بسته حمایتی دولت به ۷ میلیون نفر از اقشار کم‌درآمد تعلق می‌گیرد.	
		+	با تشکیل جلسه فوق العاده کمیته حفاظت فنی و بهداشت کار، تعطیلی فاز۲ پالایشگاه آبادان با حدود ۳۰۰۰ نفر نیروی کار از ابتدای هفته آینده تصویب شد.**	

جدول ۴: نمونه متون موجب رانش مفهوم ویژگی امید در میان داده‌ها.

تاریخ تغییر	نوع تغییر	+–	نمونه محتوا	موضوع مورد توجه در تغییر
۹۸/۱۲/۰۴	از منفی به مثبت	–	چکیده: در حالیکه پرسنل اورژانس بدون داشتن تجهیزات حفاظت شخصی در شرایط کنونی با استرس و اضطراب انتقال کرونا دست و پنجه نرم میکنند ، جوان مینودشتی برای دریافت لایک، اورژانس گالیکش را سرکار گذاشت... آیا درست است به خاطر تهیه کلیپ شهری را بدون آمبولانس کنی؟ کرونا ترسناک نیست ترستاک این افراد در جامعه هستن	- رفتارهای نامناسب در جامعه + پیام‌های حاوی شعارهای مثبت + وطن دوستی
		+	به سلامتی و شادمانی بیاندیشیم.	
		+	ایران عزیز: با وجود زلزله کرونا تحریم و ... به امید خدا آنچنان زیبا می سازیمت که تا به حال چنین کشوری نبوده و در آینده نیز نخواهد بود	
۹۹/۰۱/۲۲	از مثبت به منفی	+	چکیده: برای نخستین بار تولید ماسک‌های فیلتردار پیشرفته یونی در وزارت دفاع ... این ماسک‌ها در مدت حداقل ۵۰ ساعت استفاده مستمر نیاز به شستشو ندارد	+ پیشرفت ملی در مبارزه با بیماری - طولانی شدن تعطیلی کسب و کارها و قرنطینه
		–	حدود یک ماهه که مردم در اکثر شهرها کسب و کار رو تعطیل کردن به امید اینکه کرونا رو شکست بدن هیچکسی دوست نداره که بیکار بشینه خونه	



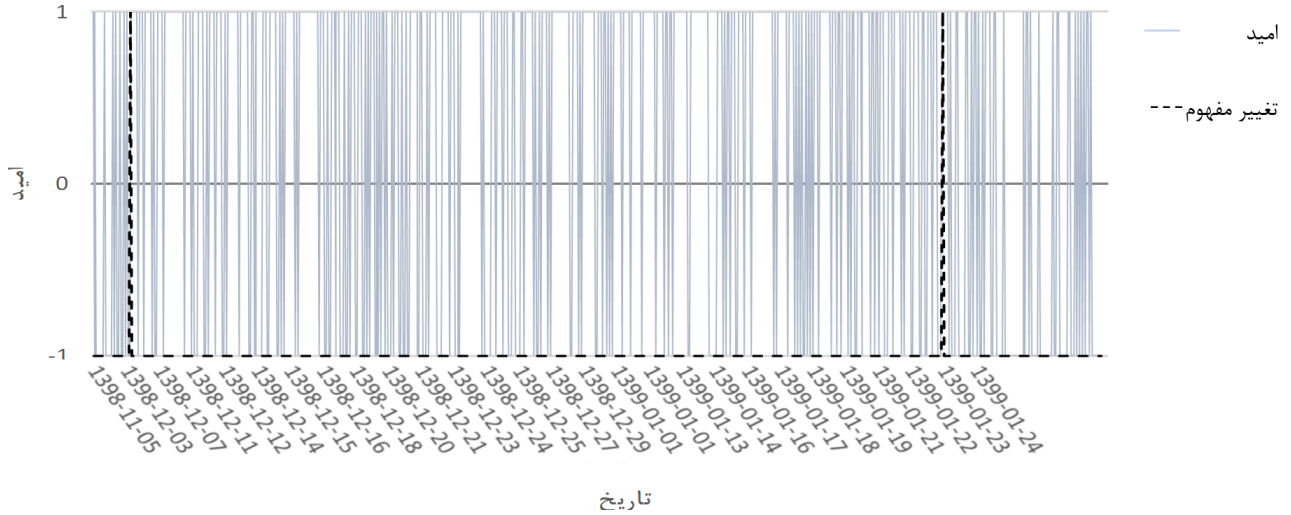
الف)



ب)

شکل ۸ الف) نمودار تجمعی امید در طول بازه مورد بررسی.

ب) نمودار تغییرات امید به‌همراه میانگین متحرک امید بر روی گروه‌های ۱۰ تایی داده در طول زمان.



شکل ۹: نمودار تغییرات امید در مجموعه داده به همراه موقعیت‌های رانش مفهوم نشان داده شده با خط چین.

۴-۴- بحث

در این مطالعه به بررسی و تحلیل جریان داده‌های متنی فارسی منتشر شده در زمینه کرونا در بازه ۱۰۰ روزه ابتدایی شیوع بیماری در ایران از نظر رانش مفاهیم پرداختیم. رانش مفهوم، ناظر بر عوض شدن توزیع آماری داده‌های یک کلاس یا ویژگی می‌باشد. در این راستا بعد از بررسی تئوری روش‌های تشخیص رانش مفهوم، با استفاده از روش HDDM-A به بررسی این امر در داده‌های مورد مطالعه جهت شناسایی تغییرات در قطبیت نظرات، دیدگاه نسبت به عملکرد مسئولین و امید پرداختیم.

با بررسی نمودارهای مرتبط با ویژگی قطبیت مشاهده شد که در ابتدای شروع کرونا مطالب منتشر شده در رابطه با بیماری کرونا بیشتر در حیطه قطبیت منفی بوده اما در طول زمان با شیب کندی به سمت خنثی پیش می رود. بنابراین تغییر این ذهنیت و گرایش دادن قطبیت به سمت مثبت دارای اهمیت زیادی برای بهبود روحیه و رفتار جامعه است. از آنجایی که شبکه‌های اجتماعی بیشتر منعکس‌کننده نظرات شخصی افراد هستند، وضعیت واقعی بیماری و اثرات آن می‌تواند به دلیل اطلاعات نادرست [۱۸] مورد غفلت قرار گیرد. این امر ممکن است بر روی ذهنیت و رفتار افراد جامعه اثر نامطلوبی داشته باشد. نمونه این موارد در بررسی محتوای مجازی قبل از شروع پیک بیماری در ماه آخر بازه مورد بررسی به چشم می‌خورد. همان‌طور که در میانگین داده‌ها مشاهده می‌شود، در روزهای پایانی اسفند ۹۸ شاهد غلبه قطبیت منفی در نظرات کاربران هستیم. این امر با توجه به ادامه‌دار شدن بیماری کرونا بعد از اعلام اولیه، شرایط جدید جامعه، و تغییر رفتار معمول افراد در روزهای پایانی سال که نوعا با روابط اجتماعی بالایی همراه بوده است، توجیه پذیر است.

همچنین در بازه مقارن با تعطیلات سال نو ۱۳۹۹ و فراغت حاصله، قطبیت نظرات نسبت به بازه قبل و بعد از آن مثبت‌تر شده و حجم تغییرات در رانش مفهوم در این بازه کمتر است. بنابراین قطبیت نظرات با وضعیت جاری بیماری در جامعه رابطه تنگاتنگی دارد و باید توجه کرد که گرچه ذهنیت منفی افراد خود موجب پیامدهای نامطلوب دیگری مثل استرس خواهد شد [۴۶] اما ذهنیت مثبت نیز در صورتی که ناشی از سهل‌گیری بیماری باشد، کاهش رعایت پروتکل‌ها و وخامت اوضاع بیماری در جامعه را سبب می‌شود.

در خصوص ویژگی دیدگاه نسبت به عملکرد مسئولین نیز، بررسی نمودارها نشان می‌دهد در بازه اعلام شیوع ویروس، نگرش منفی به عملکرد وجود دارد. با اعمال پروتکل‌های بهداشتی برای پیشگیری از بیماری و مهار نسبی کرونا، در اواخر اسفندماه یک تغییر مفهوم به نفع نگرش مثبت پیش آمده که تا پایان تعطیلات سال نو ادامه دارد. نکته حائز اهمیت اینکه افراد برای تصمیم‌گیری در باب رفتار خود در پدیده همه‌گیر کرونا، بیشتر به شبکه‌های اجتماعی و افراد خانواده و دوستان خود در این بسترها مراجعه می‌کنند [۴۷]. از آنجا که این نظرات می‌تواند با واقعیت آنچه در جامعه در باب بیماری وجود دارد متفاوت است، ذهنیت افراد ممکن است تحت تاثیر قرار گرفته و به صورت غیر واقعی به سمت مثبت یا منفی پیش رود. با توجه به تاثیر این ذهنیت بر رفتار افراد، لازم است بر اقدامات مناسب اطلاع‌رسانی و آگاهی بخشی در بسترهای مختلف در جامعه تاکید گردد.

روش تغییر مفهوم در ویژگی امید نیز (مانند دو ویژگی قبل) بازه مرتبط با پایان تعطیلات سال نو را بعنوان یک نقطه کلیدی تشخیص

داده است. میانگین امید در ابتدا منفی بوده و کمی بعد از شیوع کرونا و بالارفتن تعاملات اجتماعی مثبت شده است. در اواخر فروردین‌ماه مجددا میانگین امید به سمت منفی گرایش پیدا کرده است. این امر نشان‌دهنده لزوم توجه بیشتر به جهت‌دهی و تسکین افکار عمومی در مواقع اوج بحران باشد.

بررسی محتوایی متون حول نقاط رانش مفهوم در این جریان داده‌ها نشان می‌دهد راهکارهایی مانند همراهی و همدردی دولت با قشر محروم و وحدت رویه و هماهنگی بین دستگاه‌ها قبل و بعد از اتخاذ یک تصمیم می‌تواند تأثیر زیادی بر نگرش مردم در مورد عملکردها داشته باشد. همچنین، با بررسی موشکافانه‌تر و تحلیل محتوایی متون حول نقاط تغییر مفهوم در جدول ۲ برای ویژگی قطبیت (که بیشترین تعداد نقاط تغییر را دارد) می‌توان به راهکارهایی برای ایجاد تغییرات مثبت رسید، که بر مبنای منشا متن (خبر یا شبکه‌های اجتماعی) به دو دسته راهکار مسئولین و راهکارهای متنی طبقه‌بندی می‌شوند.

مسئولین:

تقویت روحیه ملی، تاکید بر جهادگری افراد و سازمان‌ها، اعتمادسازی به تلاش‌های مجموعه حاکمیت، ارتقای امکانات درمانی و غیر درمانی مورد نیاز، عدالت در دسترسی به امکانات، برگزاری مراسمات جمعی از جمله جشن‌های و مراسمات عمومی، شعارسازی سازنده، خودداری از بیان منفی اخبار و ترکیب خبرهای نامساعد با آموزه‌های ملی، مذهبی و جهادی، شخصیت سازی از افراد عام جامعه که به صورت موثر در تقابل با بیماری قرار دارند، بیان موثر و مکرر مسائل بهداشتی، توجیه افکار عمومی در باب تصمیمات مهم حاکمیتی

مردم:

پیام های امیدبخش، طنز پردازی، اشتراک گذاری تجربیات شخصی و خاطرات مثبت، توجه به وقایع مناسبتی، تقویت روحیه از طریق انتشار شعارها، شخصیت سازی، بازنشر پیام‌های مثبت

بنابراین با توجه به دادگان حول شاخص‌های زمانی رانش مفهوم، به شناسایی راهکارهای برخاسته از متون به‌منظور حفظ جنبه‌های مثبت و پرهیز از جنبه‌های منفی پرداختیم.

۵- نتیجه‌گیری

ایجاد پاندمی کرونا در جهان، سلامت روان افراد را دستخوش تغییرات شدید نموده است. در این مقاله روشی برای تحلیل تغییر ذهنیت افراد مبتنی بر تشخیص رانش مفهوم در جریان داده معرفی گردید. با بررسی متون فارسی منتشره، به پایش و تحلیل ذهنیت افراد جامعه ایرانی در بازه ۱۰۰ روزه ابتدایی شیوع بیماری در ایران پرداختیم. بررسی‌های آماری نشان داد با توجه به ناشناخته بودن طبیعت بیماری در ابتدای بازه مورد بررسی، مطالب منتشره عمدتا بار منفی داشته؛ اما رفته رفته به سمت خنثی و مثبت پیش می‌رود. روش‌های تشخیص رانش مفهوم بازه مقارن با تعطیلات سال نو ۱۳۹۹ و فراغت حاصله را بازه کلیدی برای تغییرات ذهنیت در هر سه ویژگی قطبیت، دیدگاه به عملکرد مسئولین و امید به سمت بار مثبت نشان می‌دهد. این امر نشان‌دهنده لزوم توجه بیشتر به جهت‌دهی و تسکین افکار عمومی در مواقع اوج بحران باشد. همچنین بررسی محتوایی متون حول سایر نقاط رانش مفهوم در این جریان داده‌ها، منجر به یافتن دو دسته راهکار در سطح مسئولین و راهکارهای مردمی برای ارتقای سلامت جامعه شد.

برخی از پیشنهاداتی که برای ادامه بررسی‌های بین رشته‌ای حول این مسئله می‌توان انجام داد عبارتند از: استفاده از مفاهیم خوشه‌بندی یا خوشه‌بندی پویا در نقاط قبل و بعد تغییر برای بهبود تحلیل معنایی تغییرات، ترکیب اطلاعات داده‌های برچسب‌دار با پرسشنامه‌ها برای تحلیل‌های جامعه‌شناسی و روانشناسی عمیقتر و دقیتر از موضوع جهت ارائه راهکارهای بلند مدت، استفاده از سایر ویژگی‌های دادگان شبکه اجتماعی مثل منتشر کننده پست، تعداد دنبال کنندگان، تعداد بازنشرها و غیره در جهت بررسی دقیقتر تاثیر پیام‌ها بر جامعه، بررسی اخبار خبرگزاری‌ها و وجود انتشار همزمان در شبکه‌های اجتماعی، پیش بینی قطبیت با توجه به سابقه موضوعات مطرح شده در روزهای اخیر، تحلیل سری زمانی داده‌ها، و بررسی بایاس کلمات نسبت به هم در مقایسه با مجموعه داده استاندارد دیگر.

سپاسگزاری

دادگان مورد استفاده در این مقاله، دادگانی است که از سوی ستاد علوم شناختی، در فراخوان ملی شناخت و تحلیل گفتمان و رفتار جامعه ایرانی در بحران کرونا، در اختیار تیم‌های شرکت کننده قرار گرفته است. این مجموعه دادگان توسط آزمایشگاه شبکه‌های اجتماعی دانشگاه تهران جمع‌آوری شده است. همچنین آزمایشگاه پردازش زبان طبیعی دانشگاه شهید بهشتی به برچسب‌گذاری موضوعی و هیجانی بخشی از این دادگان پرداختند. نویسندگان بر خود لازم می‌دانند مراتب تشکر صمیمانه خود را از این سه نهاد اعلام دارند.

مراجع

[1] P. Bastani and M. A. Bahrami, "COVID-19 Related Misinformation on Social Media: A Qualitative Study from Iran.," J. Med. Internet Res., Apr. 2020.

[2] S. N. Saleh, C. U. Lehmann, S. A. McDonald, M. A. Basit, and R. J. Medford, "Understanding Public Perception of Coronavirus Disease 2019 (COVID-19) Social Distancing on Twitter," Infect. Control Hosp. Epidemiol., Vol. 2019, pp. 1–8, Aug. 2020.

[3] H. Sha, M. Al Hasan, G. Mohler, and P. J. Brantingham, "Dynamic Topic Modeling of the COVID-19 Twitter Narrative Among U.S. Governors and Cabinet Executives," arxive, No. 2, pp. 2–7, 2020.

[4] C. Ordun, S. Purushotham, and E. Raff, "Exploratory Analysis of Covid-19 Tweets using Topic Modeling, UMAP, and DiGraphs," arxive, May 2020.

[5] T. Wang, K. Lu, K. P. Chow, and Q. Zhu, "COVID-19 Sensing: Negative Sentiment Analysis on Social Media in China via BERT Model," IEEE Access, Vol. 8, pp. 138162–138169, 2020.

[6] H. W. Park, S. Park, and M. Chong, "Conversations and Medical News Frames on Twitter: Infodemiological Study on COVID-19 in South Korea," J. Med. Internet Res., Vol. 22, No. 5, 2020.

[7] K. Chakraborty, S. Bhatia, S. Bhattacharyya, J. Platos, R. Bag, and A. E. Hassanien, "Sentiment Analysis of COVID-19 Tweets by Deep Learning Classifiers—A Study to Show How Popularity is Affecting Accuracy in Social Media," Appl. Soft Comput. J., Vol. 97, No. January, p. 106754, Dec. 2020.

[8] H. Jelodar, Y. Wang, R. Orji, and S. Huang, "Deep Sentiment Classification and Topic Discovery on Novel Coronavirus or COVID-19 Online Discussions: NLP Using LSTM Recurrent Neural Network Approach," IEEE J. Biomed. Heal. informatics, Vol. 24, No. 10, pp. 2733–2742, 2020.

[9] S. Vijaykumar, G. Nowak, I. Himelboim, and Y. Jin, "Virtual Zika Transmission after the first U.S. Case: Who Said What and How it Spread on Twitter," Am. J. Infect. Control, Vol. 46, No. 5, pp. 549–557, 2018.

[10] H. Liang, I. C. H. Fung, Z. T. H. Tse, J. Yin, C. H. Chan, L. E. Pechta, B. J. Smith, R. D. Marquez-Lameda, M. I. Meltzer, K. M. Lubell, and K. W. Fu, "How Did Ebola Information Spread on Twitter: Broadcasting or Viral Spreading?," BMC Public Health, Vol. 19, No. 1, pp. 1–11, 2019.

[11] M. J. Paul, M. Dredze, and D. Broniatowski, "Twitter Improves Influenza Forecasting," PLoS Curr., pp. 1–12, 2014.

[12] T. R. Berry, J. Wharf-Higgins, and P. J. Naylor, "SARS Wars: An Examination of the Quantity and Construction of Health Information in the News Media," Health Commun., Vol. 21, No. 1, pp. 35–44, 2007.

[13] P. Breen, J. Kelly, T. Heckman, and S. Quinn, "Mining Pre-Exposure Prophylaxis Trends in Social Media," in 2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA), 2016, pp. 214–221.

[14] J. S. P. Tulloch, R. Vivancos, R. M. Christley, A. D. Radford, and J. C. Warner, "Mapping Tweets to a Known Disease Epidemiology; a Case Study of Lyme Disease in the United Kingdom and Republic of Ireland," J. Biomed. Informatics X, Vol. 4, No. September, p. 100060, 2019.

[15] J. Ginsberg, M. H. Mohebbi, R. S. Patel, L. Brammer, M. S. Smolinski, and L. Brilliant, "Detecting Influenza Epidemics using Search Engine Query Data," Nature, Vol. 457, No. 7232, pp. 1012–1014, 2009.

[16] A. Suprem and C. Pu, "EDNA-Covid : A Large-Scale Covid-19 Tweets Dataset Collected with the EDNA Streaming Toolkit" arXiv preprint arXiv:2010.04084 (2020)

[17] C. Pu, A. Suprem, and R. A. Lima, "Challenges and Opportunities in Rapid Epidemic Information Propagation with Live Knowledge Aggregation from Social Media," In 2020 IEEE Second International Conference on Cognitive Machine Intelligence (CogMI) (pp. 131-140). IEEE.

[18] F. Jafarinejad, M. Rahimi, and H. Mashayekhi, "Tracking and Analysis of Discourse Dynamics and Polarity during the Early Corona Pandemic in Iran," J. Biomed. Inform., Vol. 121, p. 103862, 2021.

[19] C. Aggarwal, "Data Streams: Models and Algorithms," Vol. 31. 2007.

[20] G. Ditzler, M. Roveri, C. Alippi, and R. Polikar, "Learning in Nonstationary Environments: A Survey," IEEE Comput. Intell. Mag., Vol. 10, No. 4, pp. 12–25, 2015.

[21] H.-L. Nguyen, Y.-K. Woon, and W. K. Ng, "A Survey on Data Stream Clustering and Classification," Knowl. Inf. Syst., Vol. 45, Dec. 2014.

[22] V. Lemaire, C. Salperwyck, and A. Bondu, "A Survey on Supervised Classification on Data Streams," Lect. Notes Bus. Inf. Process., Mar. 2015.

[23] J. A. Silva, E. R. Faria, R. C. Barros, E. R. Hruschka, A. C. P. L. F. de Carvalho, and J. Gama, "Data Stream Clustering: A Survey," ACM Comput. Surv., Vol. 46, No. 1, Jul. 2013.

[24] J. Gama, I. Žliobaitun, A. Bifet, M. Pechenizkiy, and A. Bouchachia, "A Survey on Concept Drift Adaptation," ACM Comput. Surv., Vol. 46, No. 4, Mar. 2014.

[25] G. I. Webb, R. Hyde, H. Cao, H. L. Nguyen, and F. Petitjean, "Characterizing Concept Drift," Data Min. Knowl. Discov., Vol. 30, No. 4, pp. 964–994, 2016.

[26] W. Fan and A. Bifet, "Mining Big Data: Current Status, and Forecast to the Future," ACM SIGKDD Explor. Newsl., Vol. 16, pp. 1–5, Jan. 2014.

[27] I. Žliobaitė, "Learning Under Concept Drift: an Overview." Technical report, Vilnius University, 2010.

[28] A. Haque, L. Khan, and M. Baron, Semi Supervised Adaptive Framework for Classifying Evolving Data Stream, Vol. 9078. 2015.

[29] G. Krempl, I. Žliobaite, D. Brzeziński, E. Hüllermeier, M. Last, V. Lemaire, T. Noack, A. Shaker, S. Sievi, M. Spiliopoulou, and J. Stefanowski, "Open Challenges for Data Stream Mining Research," SIGKDD Explor. Newsl., Vol. 16, No. 1, pp. 1–10, Sep. 2014.

[30] M. A. Patil, S. Kumar, S. Kumar, and M. Garg. "Concept drift Detection for Social Media: A survey". In 2021 3rd international conference on advances in computing, communication control and networking (ICAC3N) (pp. 12-16). IEEE, Dec. 2021.

[31] E. Susi, and A. P. Shanthi, "Sentiment Drift Detection and Analysis in Real Time Twitter Data Streams." Computer Systems Science & Engineering, Vol. 45, No. 3, pp. 3231-3246, 2023.

[32] T. Uchida, and K. Yoshida. "Concept Drift in Japanese COVID-19 Infection Data." Procedia Computer Science, Vol. 207, pp. 380-387, 2022.

[33] Y. Xu, and K. Wilson, "Early Alert Systems During a Pandemic: A Simulation Study on the Impact of Concept Drift." LAK21: 11th International Learning Analytics and Knowledge Conference. 2021.

[34] C. Duckworth, F. P. Chmiel, D. K. Burns, Z. D. Zlatev, N. M. White, T. W. Daniels, ... and M. J. Boniface, "Using Explainable Machine Learning to Characterise Data Drift and Detect Emergent Health Risks for Emergency Department Admissions During COVID-19." Scientific reports 11.1 (2021): 1-10.

[35] M. Müller, and M. Salathé, "Adressing Machine Learning Concept Drift Reveals Declining Vaccine Sentiment During the COVID-19 Pandemic." arXiv preprint arXiv:2012.02197, 2020.

[36] M. Jantscher, and R. Kern. "Causal Investigation of Public Opinion during the COVID-19 Pandemic via Social Media Text." In Proceedings of the Thirteenth Language Resources and Evaluation Conference, pp. 211-226, 2022.

[37] A. Umair, and Masciari, E. "Sentimental and Spatial Analysis of Covid-19 Vaccines Tweets." Journal of Intelligent Information Systems, Vol. 60, No. 1, pp. 1-21, 2023.

[38] L. Baier, N. Kühl, J. Schöffer, and G. Satzger, "Utilizing concept drift for measuring the effectiveness of policy interventions: The

case of the COVID-19 pandemic." arXiv preprint arXiv:2012.03728 (2020).

[39] A. Bechini, A. Bondielli, P. Ducange, F. Marcelloni, and A. Renda. "Addressing Event-Driven Concept Drift in Twitter Stream: a Stance Detection Application." IEEE Access, Vol. 9, pp. 77758-77770, 2021.

[40] J. Lu, A. Liu, F. Dong, F. Gu, J. Gama, and G. Zhang, "Learning Under Concept Drift: A Review," IEEE Trans. Knowl. Data Eng., pp. 1–1, 2018.

[41] F. Bayram, B. S. Ahmed, & A. Kassler, A. "From Concept Drift to Model Degradation: An Overview on Performance-Aware Drift Detectors." Knowledge-Based Systems, Vol. 245, 108632, 2022.

[42] J. Gama, P. Medas, G. Castillo, and P. Rodrigues, "Gama J., Medas P., Castillo G., Rodrigues P. (2004) Learning with Drift Detection. Advances in Artificial Intelligence – SBIA 2004. Lecture Notes in Computer Science, vol 3171. Springer, Berlin, Heidelberg. 2004, pp. 286–295.

[43] I. Frías-Blanco, J. d. Campo-Ávila, G. Ramos-Jiménez, R. Morales-Bueno, A. Ortiz-Díaz, and Y. Caballero-Mota, "Online and Non-Parametric Drift Detection Methods Based on Hoeffding's Bounds," IEEE Trans. Knowl. Data Eng., Vol. 27, No. 3, pp. 810–823, 2015.

[44] R. S. M. Barros and S. G. T. C. Santos, "A Large-Scale Comparison of Concept Drift Detectors," Inf. Sci. (Ny)., Vol. 451–452, February 2019, pp. 348–370, 2018.

[45] W. Hoeffding, "Probability Inequalities for Sums of Bounded Random Variables," *The collected works of Wassily Hoeffding*, 1994, pp. 409-426.

[46] A. Shokri, G. Moradi, B. Piroozi, et al. "Perceived Stress due to COVID-19 in Iran: Emphasizing the Role of Social Networks." Med J Islam Repub Iran, pp. 34-55, 2020.

[47] T. Paykani, G. D. Zimet, R. Esmaeili, A. R. Khajedaluee, and M. Khajedaluee, "Perceived Social Support and Compliance with Stay-at-Home Orders During the COVID-19 Outbreak: Evidence from Iran.," BMC Public Health, Vol. 20, No. 1, p. 1650, Nov. 2020.

پاورقی‌ها:

[1] Contextual-based detection
[2] Statictical process control
[3] Windowing technique
[4] Drift Detection Method
[5] Stable
[6] Warning
[7] Drift