

Extraction and Visualization of Data Mining Knowledge Map in Covid-19 Research

Roya Abadeh 

Senior Expert in Information Technology Management, Faculty of Management and Economics, Tarbiat Modares University, Tehran, Iran

Alireza Hassanzadeh 

Professor, Department of Information Technology Management, Faculty of Management and Economics, Tarbiat Modares University, Tehran, Iran

Shaban Elahi 

Professor, Department of Management, Faculty of Administrative Sciences and Economics, Vali-e- Asr University (AJ) Rafsanjan, Rafsanjan, Iran

rrrrr rrt

The most important objectives of this research are to evaluate and illustrate the knowledge map of data mining in covid-19 research, to identify the various components and relationships between data mining concepts in different fields of covid-19 studies. The research method is applied and descriptive. In this research, we utilize the content analysis and text mining techniques. One of the findings of the current research is that a group of researches was related to the "structure and human nature of the disease". Another group of research looked at the set of factors related to adulthood that could be the basis for the absorption of the Covid-19 virus. In the co-authoring section, the highest amount of cooperation was between America and India. The data mining map of Covid-19 is based on thematic modeling including the structure and human nature of the disease, the

– Corresponding Author: ar_hassanzadeh@modares.ac.ir

How to Cite: Abadeh, R., Hassanzadeh, A., Elahi, Sh. (2024). Extraction and Visualization of Data Mining Knowledge Map in Covid-19 Research, *Journal of Business Intelligence Management Studies*, 13(49), 255-281. DOI: 10.22054/ims.2024.77278.2432

background and structure of the virus, pandemic prevention, and computer science and artificial intelligence. In this research, according to the use of information technology tools and text mining techniques, a large amount of research has been studied. The presented knowledge map deals with the thematic separation of the researches and information discovery in each collection, so it enables the possibility of comparison between specific thematic groups. In this research, the components of data mining concepts in different fields of Covid-19 studies were examined in six dimensions. In the sixth dimension, the data mining map of Covid-19 was presented based on thematic modeling.

1. Introduction

What the experience of health officials during the Corona crisis has proven is that it is not possible to manage the disease by relying on traditional methods. Modern information technologies should be used for the optimal management of this disease. Data mining and data analysis can greatly help in preventing, identifying and solving crises. One of the ways that helps researchers to achieve their research goals in their specialized field is to have an overview of the scientific framework of the field in question and previous studies. In this regard, this research has drawn a map of knowledge in the field of Corona data mining by monitoring authors and important and influential works, text mining of studies and specifying thematic clusters formed over time and visualizing information in the field of data mining of Corona. Conducting research in different subject groups provides the possibility of comparing and discovering connections between different fields of knowledge. The provided knowledge map can be used in evidence-based studies in covid-19 researches.

2. Literature Review

Considering the large amount of data related to Covid-19 that are produced from various sources such as scientific articles, reports of health organizations, clinical and laboratory data, social networks and media, the need for data mining is felt to extract useful and reliable knowledge (Chen et al., 2020). Data mining can help identify risk factors, diagnose and predict disease, evaluate the effects of treatments and vaccines, analyze social and economic behaviors, and provide solutions (Lee et al., 2020). One of the ways that helps

researchers to achieve their research goals in their specialized field is to have a general understanding of the scientific framework of the field in question. In this regard, visualizing information or drawing a map and drawing the scientific structure of that field seems necessary. In a knowledge map that is drawn based on the scientific-research outputs of scientists of a scientific field, influential authors, thematic clusters formed over time and important and influential works are determined and introduced. In knowledge maps, the appearance of new areas and the cessation of some saturated scientific areas can be clearly observed and studied. Anderson (2021) used the text mining clustering approach to evaluate the different foci of a large number of abstracts related to Covid-19. Thakur and Kumar (2021) highlight the applied text mining techniques on published scientific articles by carefully reviewing the studies conducted in the last ten years.

3. Methodology

This research uses applied methodology in terms of purpose. The statistical population of this research is all the researches published in reliable publications that are in the Scopus database. The statistical sample of this research is all scientific documents registered on the topic of data mining of Covid-19 in the years 2020, 2021 and 2022 in the Scopus database. The number of examined documents is 1749, which includes conference and research articles. In this research, all scientific documents where data mining and covid-19 were simultaneously one of its keywords were extracted. Then, by using appropriate algorithms and, text mining software, knowledge map drawing was done.

3. Results

As a result of topic analysis or Topic Modeling with the help of LDA or Latent Dirichlet Allocation algorithm, four topic keyword sets are displayed as described as follows:

Group 1: includes words that focus on data analysis related to the covid-19 crisis.

Group 2: includes words that seem to focus on machine learning models used to study covid-19.

Group 3: includes words about public information and social aspects of the covid-19 crisis.

Group 4: includes the words also includes data mining and

modeling but seems to focus more on the results obtained from these methods in studying different aspects of the crisis.

In the second part of this study, the co-occurrence map analysis of key words in data mining research on Covid-19 shows the conceptual structure of the relationship between concepts or words in a set of publications.

4. Discussion

In order to identify the components and relationships between data mining concepts in different areas of Covid-19 studies, they were identified using text mining and bibliometric methods. As a result of the process of text mining and the application of different algorithms, it can be said that data mining is one of the most widely used techniques in the following 4 areas in the research of Covid-19: 1. Understanding public reactions or crisis impact, 2. Predicting outcomes or analyzing patient data, 3. How to disseminate information through the media and public reactions, and 4. Studying different aspects of the crisis.

As a result of checking the co-occurrence map of keywords, thematic modeling was divided into the following 4 sub-categories: The first category is in the field of social media, machine learning, and in other words around concepts related to artificial intelligence. Another category is related to topics and concepts such as public health, vaccination, psychology, people's thoughts, communication, fear and emotions. The collection of this cluster can be known as a pandemic. The third category of studies can be considered as "structure and human nature of the disease". The fourth category of studies can be considered related to adulthood. The most widely used data mining method in covid-19 research with the highest reproducibility rate is Classification algorithm with 221 repetitions.

5. Conclusion




The components of data mining concepts in different fields of Covid-19 studies were examined in six dimensions. In the sixth dimension, the data mining map of Covid-19 was presented based on thematic modeling.

Keywords: Knowledge Map, Data Mining, Covid-19, Text Mining, Visualization.



استخراج و مصورسازی نقشه دانش داده کاوی در تحقیقات

کووید-۱۹

- کارشناس ارشد مدیریت فناوری اطلاعات، دانشکده مدیریت و اقتصاد، دانشگاه تربیت مدرس، تهران، ایران  **رؤیا آباده**
- استاد، گروه مدیریت فناوری اطلاعات، دانشکده مدیریت و اقتصاد، دانشگاه تربیت مدرس، تهران، ایران  **علیرضا حسن زاده***
- استاد، گروه مدیریت، دانشکده علوم اداری و اقتصاد، دانشگاه ولی عصر (عج) رفسنجان، رفسنجان، ایران  **شعبان الهی**

چکیده

مهم‌ترین اهداف این تحقیق عبارت از ارزیابی و مصورسازی نقشه دانش داده کاوی در تحقیقات کووید-۱۹، شناسایی اجزا و روابط مختلف میان مفاهیم داده کاوی در حوزه‌های مختلف مطالعات کووید-۱۹ می‌باشد. این تحقیق از نظر هدف کاربردی و از نوع توصیفی، تحلیلی با بهره‌مندی از روش تحلیل محتوا و فنون متن کاوی انجام خواهد شد. یکی از یافته‌های تحقیق حاضر این است که دسته‌ای از تحقیقات به «ساختار و ماهیت انسانی بیماری» مرتبط بود. دسته‌ای دیگر از تحقیقات به مجموعه عوامل مربوط به بزرگ‌سالی می‌پرداخت که می‌توانست زمینه‌ساز جذب ویروس کووید-۱۹ باشد. در بخش هم‌نویسندگی بیشترین میزان همکاری میان امریکا و هند بوده است. نقشه داده کاوی کووید-۱۹ بر اساس مدل‌سازی موضوعی مشتمل بر ساختار و ماهیت انسانی بیماری، زمینه و ساختار ویروس، پیشگیری پاندمیک و علوم کامپیوتری و هوش مصنوعی است. در این تحقیق با توجه به استفاده از ابزارهای فناوری اطلاعات و تکنیک متن کاوی حجم وسیعی از تحقیقات مورد مطالعه قرار گرفته است. نقشه دانش ارائه شده به تفکیک موضوعی تحقیقات و کشف اطلاعات در هر مجموعه می‌پردازد لذا امکان مقایسه بین گروه‌های موضوعی خاص را امکان‌پذیر می‌کند. در این تحقیق، اجزای مفاهیم داده کاوی در حوزه‌های مختلف مطالعات کووید-۱۹ در شش بعد مورد بررسی قرار گرفت. در بعد ششم نقشه داده کاوی کووید-۱۹ بر اساس مدل‌سازی موضوعی ارائه گردید.

کلیدواژه‌ها: نقشه دانش، داده کاوی، کووید ۱۹، متن کاوی، مصورسازی.

مقدمه

پایگاه داده‌ها در حوزه سلامت حاوی میزان وسیعی از داده‌های بالینی است که کشف ارتباطات و الگوها در آن می‌تواند به دانش جدید پزشکی بیانجامد. امروزه با توجه به ظهور نظام‌های اطلاعات یکپارچه و رشد فناوری اطلاعات، این مهم بیش از پیش نمایان شده است. داده‌کاوی از جمله پیشرفت‌های فناوری در راستای مدیریت داده‌ها است و استفاده گسترده از سیستم‌های اطلاعات و پایگاه‌های داده، ادغام آن را با شیوه‌های سنتی به یک الزام تبدیل کرده است (مقدسی و همکاران، ۱۳۹۱). آنچه تجربه متولیان سلامت در هنگام بحران کرونا ثابت کرده است این بوده که مدیریت بیماری با اتکا به روش‌های سنتی امکان‌پذیر نبوده و باید از فناوری‌های نوین اطلاعاتی برای مدیریت بهینه‌ی این بیماری استفاده کرد (محمدزاده و همکاران، ۲۰۲۱). متدهای داده‌کاوی به‌عنوان زیرمجموعه‌ای از ابزارهای فناوری اطلاعات یکی از زمینه‌های امیدوارکننده و دائماً در حال تحول در عرصه تجزیه و تحلیل داده‌ها می‌باشد که پس از شیوع کووید-۱۹ به‌طور گسترده‌ای در تحقیقات دانشمندان و محققین مورد استفاده قرار گرفته است. کنسرسیوم غنی از تکنیک‌ها که در حوزه داده‌کاوی قرار می‌گیرند، باعث شده تا داده‌کاوی را به نیرویی قدرتمند برای دانشمندان جهت کشف الگوهای پنهان تبدیل کند. داده‌کاوی و تحلیل داده می‌تواند به میزان زیادی در جلوگیری، شناسایی و حل بحران کمک کند.

یکی از راه‌هایی که پژوهشگران را برای رسیدن به اهداف پژوهشی در حوزه تخصصی خود کمک می‌کند، داشتن درک و نمایی کلی از چارچوب علمی حوزه موردنظر و مطالعات پیشین است (سهیلی و همکاران، ۲۰۱۴). در این راستا این پژوهش با پایش نویسندگان و آثار مهم و تأثیرگذار، متن‌کاوی مطالعات و مشخص نمودن خوشه‌های موضوعی شکل گرفته در طول زمان و مصورسازی اطلاعات حوزه داده‌کاوی کرونا به ترسیم نقشه دانش در آن حوزه پرداخته است. با توجه به گستردگی بحث علم داده و داده‌کاوی و پیشرفت روزافزون آن عموماً بخشی از مطالب خاص و پژوهش‌ها می‌تواند در آن تکرار شود، این نقشه به‌عنوان مرجعی برای نمایش متدهای بکار رفته و میزان فراوانی و

تأثیرگذار بودن آن در روند رشد تحقیقات کووید-۱۹ می تواند راه گشای تحقیقات آینده باشد و مانع از تکرار تحقیقات مشابه گردد. بیان خاستگاه و روش های داده کاوی انجام شده در پر استنادترین مقالات می تواند زمینه ای برای انجام آن تحقیقات در کشور ایران فراهم نماید.

استخراج اطلاعات روشی به منظور پیدا کردن اطلاعات نیمه ساخت یافته و ساخت یافته متن است. استخراج اطلاعات برنامه مهمی در متن کاوی است. متن کاوی حتی فراتر از دستیابی به اطلاعات است که به کاربران در تجزیه و تحلیل و خلاصه سازی اطلاعات و سهل کردن تصمیمات کمک می کند (شفیعی نیکابادی و همکاران، ۱۴۰۲)

مصورسازی داده یکی از مراحل تجزیه و تحلیل داده ها در «علم داده»^۱ محسوب می شود. «ویتالی فریدمن»^۲، مصورسازی داده و ویژگی ها آن را این طور توصیف می کند: «هدف اصلی مصورسازی داده یا تجسم آن ها، برقراری ارتباط واضح و مؤثر از طریق ابزارهای گرافیکی است. این بدان معنا نیست که لزوماً مصورسازی داده باعث ایجاد یک تصویر زیبا شود بلکه درک اطلاعات به شیوه ساده و راحت منظور این روش توصیفی محسوب می شود. به طور مؤثر، هم فرم زیباشناختی و هم عملکرد باید دست به دست هم دهند و با برقراری ارتباط با جنبه های اصلی آن به روشی بصری، اطلاعات نهفته در داده های نسبتاً پراکنده و پیچیده را ارائه دهند. مصورسازی داده که بدون هدف تولید شده و فقط جاذبه های بصری داشته باشد، منظور نظر مصورسازی داده نخواهد بود». در قرن حاضر، با توجه به حجم بسیار زیاد اطلاعات و لزوم به شناخت رفتار آن ها در زمان کوتاه، مصورسازی داده به یک حوزه فعال تحقیق، تدریس و توسعه تبدیل شده است، به طوری که این تکنیک، مصورسازی علمی و اطلاعاتی را متحد کرده است. پروفیسور «ادوارد توفت»^۳ توضیح می دهد که کاربرانی که از روش های نمایش اطلاعات استفاده می کنند، وظایف تحلیلی خاصی مانند انجام مقایسه را رهبری می کنند. طراحی گرافیکی اطلاعات باید از

1. Data Science
2. Vitaly Friedman
3. Edward Tufte

وظیفه این گروه که همان تحلیل داده‌ها است، پشتیبانی کند. (توفت، ۲۰۰۱) مصورسازی داده‌ها با در نظر گرفتن ارتباطات، علم داده و طراحی می‌تواند بینش‌های کلیدی را از تحقیقات صورت گرفته به نحوی گیرا و پرمعنی استخراج کند. این تحقیق با الگو گرفتن از هستان‌شناسی تحقیقات داده‌کاوی باهدف توصیف و رسمی کردن موجودیت‌ها از مطالعات داده‌کاوی کووید-۱۹ و کشف دانش از آن صورت گرفته است که دربرگیرنده موجودیت‌های پایه (مانند نوع داده‌ها، وظیفه داده‌کاوی، الگوریتم داده‌کاوی، نویسنده، کشور مبدأ و غیره) است. تمام موجودیت‌های داده‌کاوی تعریف شده و سازمان‌دهی شده در قالب یک هستان‌شناسی، می‌توانند ستون فقرات سیستم‌های داده‌کاوی خودکار باشند. (پانوو و همکاران، ۲۰۰۹) انجام تحقیق در گروه‌های موضوعی مختلف امکان مقایسه و کشف ارتباطات بین حوزه‌های مختلف دانش را فراهم می‌کند. اطلاعات نقشه دانش ارائه شده می‌تواند در مطالعات مبتنی بر شواهد در تحقیقات کووید-۱۹ مورد استفاده قرار گیرد.

اهداف تحقیق عبارت‌اند از: ارزیابی و مصورسازی نقشه دانش داده‌کاوی در تحقیقات کووید-۱۹، شناسایی اجزا و روابط مختلف میان مفاهیم داده‌کاوی در حوزه‌های مختلف مطالعات کووید-۱۹، تبیین پیاده‌سازی نقشه دانش تحقیقات داده‌کاوی کووید-۱۹

پیشینه پژوهش

داده‌کاوی یکی از شاخه‌های علوم داده است که با استفاده از روش‌های آماری، ریاضی و هوش مصنوعی، الگوها، روابط و قواعدی را از داده‌های بزرگ و پیچیده استخراج می‌کند (هان و همکاران، ۲۰۱۲). داده‌کاوی شامل مراحل مختلفی است که از جمله می‌توان به انتخاب و پیش‌پردازش داده‌ها، انتخاب و اعمال تکنیک‌ها و الگوریتم‌های مناسب، ارزیابی و تفسیر نتایج و ارائه دانش از داده‌ها اشاره کرد (فایض و همکاران، ۲۰۱۶). داده‌کاوی دارای چندین روش و تکنیک است که بر اساس نوع داده‌ها و هدف داده‌کاوی می‌توان از آن‌ها استفاده کرد. از جمله این روش‌ها می‌توان به خوشه‌بندی (Clustering)، دسته‌بندی (Classification)، رگرسیون (Regression)، شبکه‌های عصبی (Neural Networks)،

ماشین‌های بردار پشتیبان (Support Vector Machines)، قواعد انجمنی (Association Rules)، تحلیل مؤلفه‌های اصلی (Principal Component Analysis) و تحلیل مسیر (Path Analysis) اشاره کرد (ونگ و همکاران، ۲۰۱۹).

داده کاوی در زمینه‌های مختلفی از جمله پزشکی، تجارت، آموزش و مهندسی کاربرد دارد و می‌تواند به تصمیم‌گیری‌های مبتنی بر داده کمک کند و ارزش افزوده‌ای برای سازمان‌ها و افراد ایجاد کند (ونگ و همکاران، ۲۰۱۹). یکی از زمینه‌هایی که نیاز به داده کاوی دارد، تحقیقات کووید-۱۹ است. کووید-۱۹ یک بیماری عفونی است که توسط ویروس جدیدی به نام SARS-CoV-2 ایجاد می‌شود و از اواخر سال ۲۰۱۹ شیوع گسترده‌ای در سراسر جهان داشته است (وزارت بهداشت و درمان، ۲۰۲۰). برای مقابله با این بیماری، نیاز به تحقیقات گسترده و مستمر در زمینه‌های مختلفی از جمله علوم پایه، علوم بالینی، علوم اجتماعی و علوم داده وجود دارد (کوهن و همکاران، ۲۰۲۰). با توجه به حجم زیاد داده‌های مربوط به کووید-۱۹ که از منابع مختلفی مانند مقالات علمی، گزارش‌های سازمان‌های بهداشتی، داده‌های بالینی و آزمایشگاهی، شبکه‌های اجتماعی و رسانه‌ها تولید می‌شوند، نیاز به داده کاوی برای استخراج دانش مفید و قابل استناد از این داده‌ها احساس می‌شود (چن و همکاران، ۲۰۲۰). داده کاوی می‌تواند به شناسایی عوامل خطر، تشخیص و پیش‌بینی بیماری، ارزیابی اثرات درمان‌ها و واکسن‌ها، تحلیل رفتارهای اجتماعی و اقتصادی و ارائه راهکار کمک کند (لی و همکاران، ۲۰۲۰). با این حال، داده کاوی در تحقیقات کووید-۱۹ با چالش‌هایی روبرو است که نیاز به حل و فصل دارند. یکی از این چالش‌ها کیفیت و قابلیت اطمینان داده‌ها است که ممکن است به دلیل ناهمگونی، ناقص بودن، نادرست بودن یا تغییرپذیری داده‌ها کاهش یابد (کوهن و همکاران، ۲۰۲۰). چالش دیگر مربوط به انتخاب و اعمال تکنیک‌ها و الگوریتم‌های مناسب برای داده کاوی است که باید با توجه به ماهیت و هدف داده‌ها انجام شود (چن و همکاران، ۲۰۲۰). چالش سوم مربوط به ارزیابی و تفسیر نتایج داده کاوی است که باید با رویکردی علمی، اخلاقی و اجتماعی صورت گیرد (لی و همکاران، ۲۰۲۰). برای رفع این

چالش‌ها، نیاز به ایجاد یک نقشه مفهومی دانش داده‌کاوی در تحقیقات کووید-۱۹ وجود دارد. نقشه مفهومی دانش یک ابزار بصری است که مفاهیم، اصطلاحات، تعاریف، روش‌ها، الگوریتم‌ها، معیارها، مثال‌ها و منابع مربوط به یک حوزه دانش را به صورت ساختاریافته و منطقی نشان می‌دهد (نواک و همکاران، ۲۰۰۶). نقشه مفهومی دانش می‌تواند به چندین منظور مورداستفاده قرار گیرد، از جمله: یادگیری و آموزش، ارزیابی و اندازه‌گیری، ارتباط و همکاری، خلاقیت و نوآوری، مدیریت و برنامه‌ریزی و تصمیم‌گیری (کینشین و همکاران، ۲۰۰۸).

در کتاب «ترسیم نقشه علم» (سهیلی و همکاران، ۱۳۹۸) عنوان گردیده که یکی از راه‌هایی که پژوهشگران را برای رسیدن به اهداف پژوهشی در حوزه تخصصی خود کمک می‌کند، داشتن درک و نمایی کلی از چارچوب علمی حوزه موردنظر است. در این راستا دیداری‌سازی اطلاعات یا ترسیم نقشه و ترسیم ساختار علمی آن حوزه، ضروری به نظر می‌رسد، در یک نقشه علمی که بر اساس برون‌دادهای علمی-پژوهشی دانشمندان یک حوزه علمی ترسیم می‌شود، نویسندگان تأثیرگذار، خوشه‌های موضوعی شکل گرفته در طول زمان و آثار مهم و تأثیرگذار، تعیین و معرفی می‌شوند. از دیگر ویژگی‌های دیداری‌سازی اطلاعات که حاصل آن ترسیم نقشه‌های علمی است، امکان مطالعه تاریخ علم است. در نقشه‌های علمی به وضوح ظهور حیطه‌های جدید و توقف برخی حیطه‌های علمی اشیاع‌شده قابل ملاحظه و مطالعه است. به بیانی ساده نقشه علمی به تصویر کشیدن نتایج برآمده از تجزیه و تحلیل انتشارات یک حوزه علمی از زوایای مختلف و ارائه نگرشی کلی از آن حوزه است. همان‌گونه که تاریخ هر شهر، کشور و منطقه را افراد، رویدادها و حوادث برجسته و غیره آن منطقه شکل می‌دهند، تاریخ علم هر حوزه علمی، کشور و یا دانشگاه را نیز نویسندگان و مقالات برجسته آن‌ها تشکیل می‌دهند.

مقاله «کاربرد روش‌های متن‌کاوی در تحقیقات نوآوری: وضعیت فعلی، الگوهای تکامل و اولویت‌های توسعه» (آنتون و همکاران، ۲۰۲۰) به دنبال ارائه اولین بررسی سیستماتیک کاربردهای متن‌کاوی در تحقیقات نوآوری بوده است. مجموعه‌ای از ۱۲۴

مقاله که در مورد یک موضوع مرتبط با نوآوری در سبدهی از ۱۰ مجله برتر مدیریت نوآوری و ۸ مجله برتر مدیریت عمومی منتشر شده است، شناسایی و بررسی می‌شود. بر اساس تجزیه و تحلیل دستی و محاسباتی سیستماتیک این ۱۲۴ مقاله، وضعیت و تکامل کاربردهای متن کاوی در تحقیقات نوآوری مستند گردیده و اولویت‌های مفهومی، روش‌شناختی و زمینه‌ای را برای کاربردهای متن کاوی آینده استخراج می‌کند. این گونه استدلال گردیده است که کاربردهای متن کاوی در تحقیقات نوآوری، اکنون به مرحله بلوغ رسیده‌اند و نیاز به نمایش و مطالعات موردی متن کاوی در تحقیقات نوآوری به سرعت در حال کاهش است. پیشینه مفهومی مورد نیاز برای توسعه طرح‌های تحقیقاتی که کاربردهای متن کاوی در تحقیقات نوآوری را به سطح بعدی می‌رساند، در این مقاله در اختیار محققان قرار گرفته است.

در مقاله «استفاده از متن کاوی برای به دست آوردن بینش از ادبیات COVID-19» (اندرسون، ۲۰۲۱) با استفاده از رویکرد خوشه‌بندی متن کاوی، در این مطالعه توانسته است کانون‌های متفاوت تعداد زیادی از چکیده‌های مرتبط با کووید-۱۹ را ارزیابی کند. این مقاله یک رویکرد خوشه‌بندی متن را پیشنهاد می‌کند تا به محققان در محدود کردن دامنه مقالات تحقیقاتی که نیاز به هضم دارند کمک کند. نتایج تجزیه و تحلیل خوشه‌بندی متن همچنین بینشی را در مورد موضوعات تحقیقاتی در حال ظهور ارائه می‌دهد. این تجزیه و تحلیل بیان می‌کند یک مفهوم عملی این پژوهش این است که باعث کاهش زمان مورد نیاز یک محقق برای غربال کردن مقالات، جهت یافتن مقالات مرتبط با تحقیق خود خواهد شد، محدودیت مشاهده شده در تحقیق بازه سال‌های تحقیق می‌باشد که فقط شامل داده‌های سال ۲۰۲۰ می‌باشد.

در مقاله «کاربرد تکنیک‌های واکاوی متن در مقالات تحقیقی: روش‌ها و ابزارها» (تاکور و کومار، ۲۰۲۱) با بررسی دقیق مطالعات انجام شده در ده سال گذشته، تکنیک‌های متن کاوی کاربردی بر روی مقالات علمی منتشر شده را برجسته می‌کند. یافته‌های این مطالعه نشان‌هایی را در مورد ابزارها و تکنیک‌های متن کاوی ترجیحی، حوزه‌های اصلی

که مطالعات در آن انجام شده و حجم نمونه ترجیحی توسط محققان در حین متن کاوی ادبیات علمی منتشر شده ارائه می کند. ابزارهای پرکاربرد معرفی شده عبارتند از R- KHcoder3, MALLET, Rapid miner, Python, SAS Enterprise, Packages Latent Dirichlet تکنیک های متن کاوی دارای فراوانی معرفی شده عبارتند از Support Vector, Association rules k-Means clustering, Allocation (LDA) Machine, Natural Language Processing, تحقیقات صورت گرفته مربوط به مقالات منتشر شده در پایگاه های داده اسکوپوس، گوگل اسکولار، IEEE، Europe PMC و ACM در بازه زمانی سال های ۲۰۱۹ تا ۲۰۲۰ می باشد.

در مقاله با عنوان «استفاده از تکنیک های داده کاوی برای مبارزه و کنترل اپیدمی ها: مطالعه مروری» (صفدری و همکاران، ۲۰۲۱) به بررسی مقالات منتشر شده برای تعیین محبوب ترین روش های داده کاوی و شکاف دانش در تکنیک های داده کاوی برای کنترل اپیدمی ها پرداخته اند. پایگاه های داده Web of Science، Scopus و PubMed برای جستجوی سیستماتیک انتخاب شده اند. پس از ارزیابی مقالات استخراج شده نتیجه فوق از تحقیقات حاصل گردیده است: از ۳۳۵ استناد بازیابی شده، ۵۰ مقاله به عنوان مقالات واجد شرایط از طریق بررسی محدود تعیین شده اند. نتایج بررسی نشان داده است که مورد علاقه ترین روش داده کاوی مربوط به پردازش زبان طبیعی (۲۲٪) و رایج ترین رویکرد پیشنهادی آشکار کردن ویژگی های بیماری (۲۲٪) بوده است. در مورد بیماری ها، بیشترین بیماری مورد توجه کووید-۱۹ بوده است. مطالعات نشان می دهد که استفاده از تکنیک های یادگیری تحت نظارت (۹۰٪) غالب است. در مورد دامنه مراقبت های بهداشتی، بیماری های عفونی (۳۶٪) شایع ترین هستند و به طور دقیق توسط رشته اپیدمیولوژی دنبال می گردند. رایج ترین نرم افزارهای مورد استفاده در مطالعات SPSS (۲۲٪) و R (۲۰٪) بوده است. نتایج نشان داده است که تحقیقات ارزشمندی با استفاده از قابلیت های روش های کشف دانش برای درک ابعاد ناشناخته بیماری ها در همه گیری ها انجام شده است. این مطالعه تنها بر روی ۵۰ مقاله که دارای ویژگی های مدنظر محقق (طی فرآیند PRISMA)

بوده، صورت گرفته و روش تحقیق مطالعه از نوع مروری بوده است.

روش^۱

از آنجایی که هدف این پژوهش استخراج و مصورسازی نقشه دانش داده کاوی در تحقیقات کووید-۱۹ و به زبانی دیگر ایجاد رابط بصری با دانش علمی جهان در زمینه داده کاوی در تحقیقات کووید-۱۹ می باشد، این تحقیق از نظر هدف کاربردی و از نوع توصیفی، تحلیلی با بهره‌مندی از روش تحلیل محتوا و فنون متن کاوی انجام شده است. جمع آوری اطلاعات به صورت کیفی-کمی (ترکیبی) می باشد. جامعه پژوهش، پژوهش‌های مرتبط با داده کاوی در حوزه کووید-۱۹ در سراسر جهان در بازه زمانی سال‌های ۲۰۲۰ تا ۲۰۲۲ میلادی در پایگاه داده Scopus است.

جامعه آماری این پژوهش را می توان کلیه پژوهش‌های منتشر شده در نشریات معتبر و در پایگاه داده Scopus جهت جستجو دانست. نمونه آماری این تحقیق تمام سندهای علمی ثبت شده با موضوع داده کاوی کووید-۱۹ در سال‌های ۲۰۲۰، ۲۰۲۱ و ۲۰۲۲ در پایگاه داده اسکوپوس است. تعداد سندهای مورد بررسی ۱۷۴۹ می باشد که شامل مقالات کنفرانسی و پژوهشی است.

با توجه به ماهیت تحقیق که کیفی است و از روش متن کاوی و بیلیومتریك استفاده شده است، لذا ابزار اندازه گیری در این تحقیق به صورت مشهود وجود ندارد. گردآوری اطلاعات از طریق سایت اسکوپوس به عنوان یکی از نمایه‌های استنادی معتبر و با استفاده از موتور جستجوی این سایت بوده است. سایت اسکوپوس قابلیت جستجوی اسناد علمی را در حیطه حوزه‌های علمی مختلف، کشورها، مراکز پژوهشی و دانشکده‌ها و از این قبیل را در اختیار پژوهشگران قرار می دهد. در این تحقیق تمامی اسناد علمی که داده کاوی و کووید-۱۹ به طور هم‌زمان یکی از کلیدواژه‌های آن بوده است استخراج گردید. سپس با به کارگیری الگوریتم‌ها و نرم‌افزارهای مناسب به متن کاوی و ترسیم نقشه دانش پرداخته

شد. نمودار مراحل انجام تحقیق به شرح شکل شماره (۱) می باشد:

شکل ۱. نمودار مراحل انجام تحقیق



الگوریتم‌ها و نرم افزارهای مورد استفاده:

متن کاوی متون با استفاده از نرم افزار Orange صورت گرفته است. فرآیند متن کاوی در دو فاز Text preprocessing و Content analysis صورت پذیرفته است. در فاز اول از الگوریتم‌های Normalization and noise reduction و در فاز دوم از الگوریتم‌های Dictionary-based techniques استفاده گردید. از نرم افزار VOSviewer جهت تحلیل پراستنادترین آثار، ترسیم ماتریس استنادات، خوشه بندی و به تصویر کشیدن روابط میان نویسندگان استفاده گردیده است.

در مرحله اول جهت دستیابی به مفاهیم خوشه‌های موضوعی و تنوع الگوریتم‌های داده کاوی به مرور سیستماتیک چکیده پژوهش‌های داده کاوی کووید-۱۹ توسط فرآیند

متن کاوی زبان طبیعی توسط برنامه orange پرداخته شد. چکیده مقالات موجود در فایل استخراج شده از پایگاه داده اسکوپوس توسط برنامه اکسل و ابزار kutools به ۱۷۴۹ فایل متنی تبدیل گردید فایل های متنی به برنامه انتقال داده شد. پیش پردازش داده ها در نرم افزار اورنج^۱ شامل حذف کلمات بی اهمیت stop words، ریشه یابی کردن و لمتایز کردن lemmatization بود. تحلیل موضوعی یا Topic Modeling با کمک الگوریتم LDA یا Latent Dirichlet Allocation یکی از روش های محبوب در پردازش زبان طبیعی است که برای استخراج موضوعات مخفی در مجموعه ای از متون استفاده می شود. در این روش، هر سند به عنوان ترکیبی از موضوعات در نظر گرفته می شود و هر موضوع به عنوان توزیعی از کلمات.

روند کار الگوریتم LDA به شرح زیر است:

۱. پیش پردازش: ابتدا متن ها را پاک سازی کرده و به صورت بردار کلمات نمایش می دهیم.
۲. مدل سازی: سپس با استفاده از الگوریتم LDA، مدل موضوعی را بر روی داده های پیش پردازش شده اعمال می کنیم. این الگوریتم به ما اجازه می دهد تا موضوعات مختلف موجود در مجموعه داده های متنی را شناسایی کنیم.
۳. تخصیص موضوعات: در این مرحله، هر سند به یک یا چند موضوع اختصاص داده می شود. این اختصاص بر اساس توزیع کلمات در هر سند و توزیع کلمات در هر موضوع است.
۴. تفسیر نتایج: در نهایت، می توان با تحلیل توزیع موضوعات در اسناد و توزیع کلمات در موضوعات، به تفسیر نتایج پرداخت.

یکی از چالش های اصلی در استفاده از الگوریتم LDA تعیین تعداد بهینه موضوعات یا K است. برای این منظور می توان از روش های مختلفی مانند روش Elbow یا Silhouette استفاده کرد. با استفاده از الگوریتم LDA (Latent Dirichlet Allocation)، مدل موضوعی بر روی داده های پیش پردازش شده اعمال گردید. این الگوریتم امکان

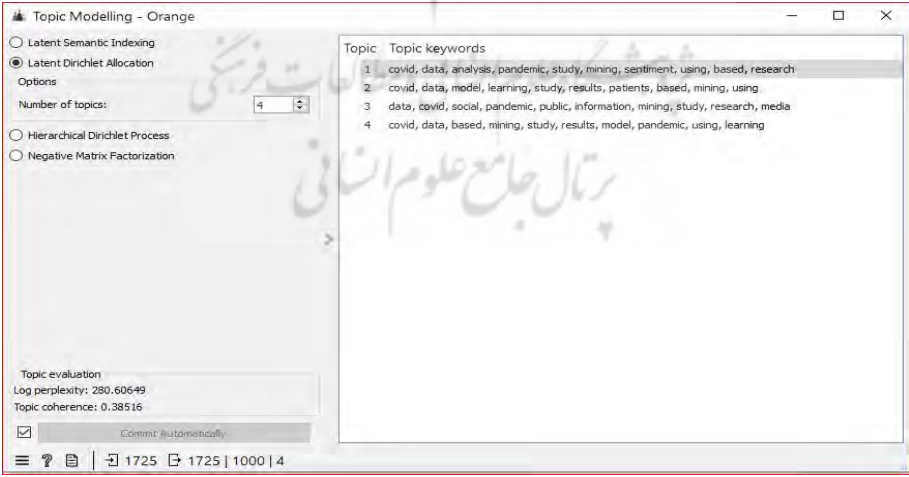
شناسایی موضوعات مختلف موجود در مجموعه داده‌های متنی را فراهم نمود. سپس با استفاده از ابزار concordance میزان تکرار پذیری روش‌های مختلف داده‌کاوی در اسناد استخراج گردید.

در ادامه تحقیق از نرم‌افزار VOS viewer برای ساخت و مصورسازی مفاهیم مربوط به کشورها و نویسندگان و ارتباطات این مفاهیم استفاده شد. تحلیل‌های هم‌واژگانی (هم‌رخدادی)، هم‌استنادی، هم‌نویسندگی مطالعات کتاب‌سنجی و علم‌سنجی از قابلیت‌های خوب این نرم‌افزار است. به‌وسیله این امکانات دوری و نزدیکی رشته‌های مختلف از هم، نقاط داغ پژوهشی، نقاط بکر و دست‌نخورده، همکاری‌های علمی محققان و حتی اثرات پژوهشی را می‌توان مطالعه و بررسی کرد.

یافته‌ها

در نتیجه تحلیل موضوعی یا Topic Modeling با کمک الگوریتم LDA یا Latent Dirichlet Allocation، چهار مجموعه کلیدواژه موضوعی به شرح شکل ۲ نمایش داده شده است:

شکل ۲. خروجی مدل‌سازی موضوعی تحقیقات داده‌کاوی کووید-۱۹



The screenshot shows the 'Topic Modelling - Orange' window. On the left, the 'Options' section is visible, with 'Latent Dirichlet Allocation' selected and 'Number of topics' set to 4. The 'Topic evaluation' section shows 'Log perplexity: 280.60649' and 'Topic coherence: 0.38516'. The 'Topic keywords' section on the right lists four topics with their corresponding keywords:

Topic	Topic keywords
1	covid, data, analysis, pandemic, study, mining, sentiment, using, based, research
2	covid, data, model, learning, study, results, patients, based, mining, using
3	data, covid, social, pandemic, public, information, mining, study, research, media
4	covid, data, based, mining, study, results, model, pandemic, using, learning

گروه ۱: شامل کلمات (کووید، داده، تحلیل، همه‌گیری، مطالعه، کاوش، احساس، استفاده، مبتنی، تحقیق) بر روی تجزیه و تحلیل داده‌های مربوط به بحران کووید-۱۹ تمرکز دارد. احتمالاً از تکنیک‌های تحلیل احساسات و استخراج داده برای درک واکنش‌های عمومی یا تأثیر بحران استفاده می‌کند.

گروه ۲: شامل کلمات (کووید، داده، مدل، یادگیری، مطالعه، نتایج، بیماران، مبتنی، کاوش، استفاده) به نظر می‌رسد که بر روی مدل‌های یادگیری ماشینی که برای مطالعه کووید-۱۹ استفاده می‌شوند، متمرکز است. شاید برای پیش‌بینی نتایج یا تجزیه و تحلیل داده‌های بیماران استفاده شده باشد.

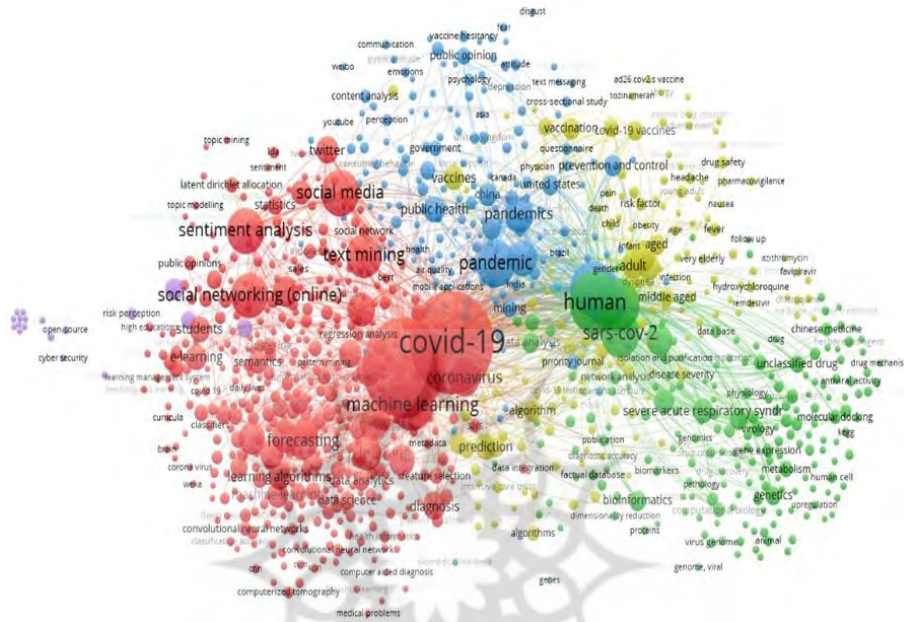
گروه ۳: شامل کلمات (داده، کووید، اجتماعی، همه‌گیری، عمومی، اطلاعات، کاوش، مطالعه، تحقیق، رسانه) در مورد اطلاعات عمومی و جوانب اجتماعی بحران کووید-۱۹ است که احتمالاً شامل مطالعاتی در مورد چگونگی انتشار اطلاعات از طریق رسانه‌ها و واکنش‌های عمومی است.

گروه ۴: شامل کلمات (کووید، داده، مبتنی، کاوش، مطالعات، نتیجه، مدل، همه‌گیری، استفاده، یادگیری) نیز شامل استخراج داده و مدل‌سازی است اما به نظر می‌رسد بیشتر بر روی نتایج به دست آمده از این روش‌ها در مطالعه جنبه‌های مختلف بحران تمرکز دارد. به نظر می‌رسد که این گروه‌ها به موضوعات مختلف تحقیقات COVID-19 مرتبط هستند و هر یک از آن‌ها بر روی یک جنبه خاص از این تحقیقات تمرکز دارند. این می‌تواند نشان‌دهنده تنوع و گستردگی تحقیقات در این زمینه باشد.

در بخش دوم تحلیل نقشه هم‌رخدادی کلمات کلیدی در تحقیقات داده‌کاوی کووید-۱۹، ساختار مفهومی ارتباط میان مفاهیم یا کلمات را در یک مجموعه از نشریات نمایش می‌دهد. کلماتی که در کنار هم در یک مقاله می‌آیند به یک شبکه مرتبط خواهد بود، این شبکه تحت شبکه هم‌رخدادی واژه شناخته می‌شود. این ساختار برای درک سرفصل‌های که توسط یک حوزه تحقیق پوشش داده می‌شود، استفاده می‌گردد. در شکل شماره ۳ نقشه هم‌رخدادی کلمات کلیدی در تحقیقات داده‌کاوی کووید-۱۹ که توسط

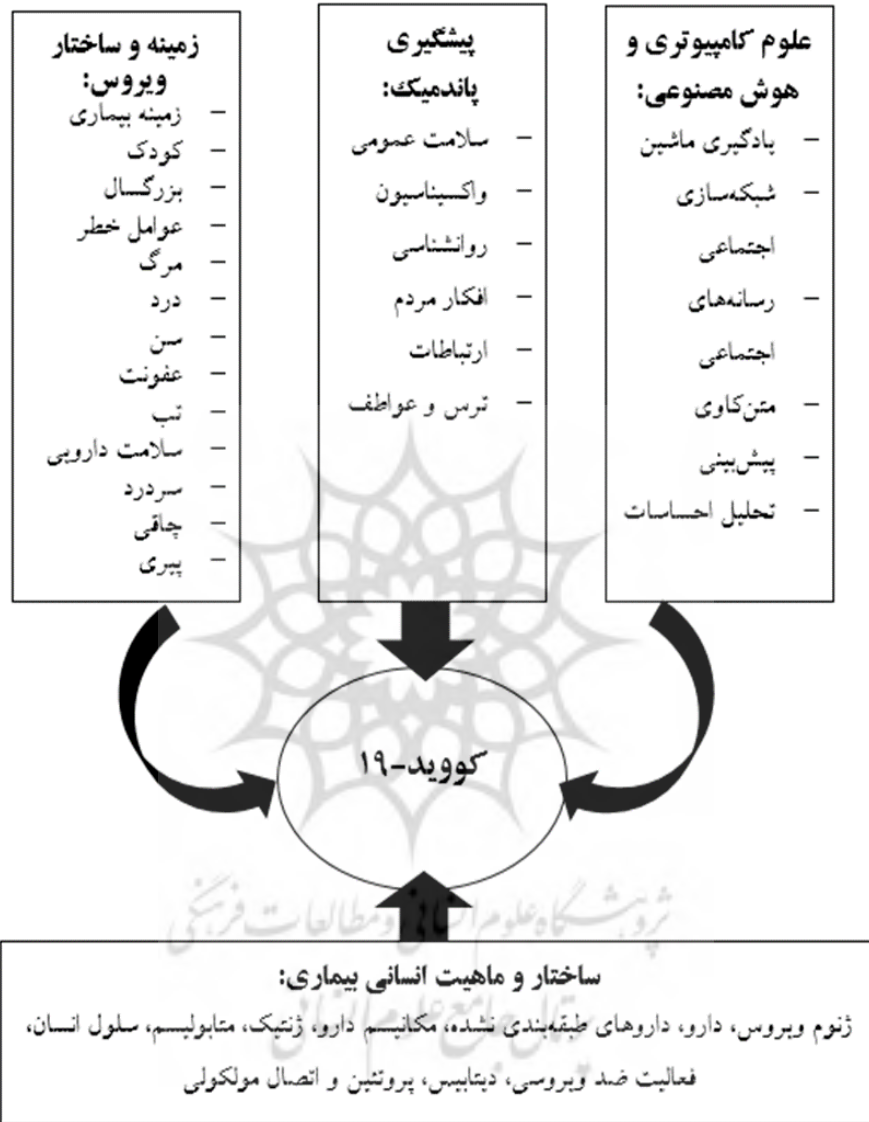
نرم افزار VosViewer استخراج گردیده نشان داده شده است.

شکل ۳. نقشه هم رخدادی کلمات کلیدی در تحقیقات داده کاوی کووید-۱۹

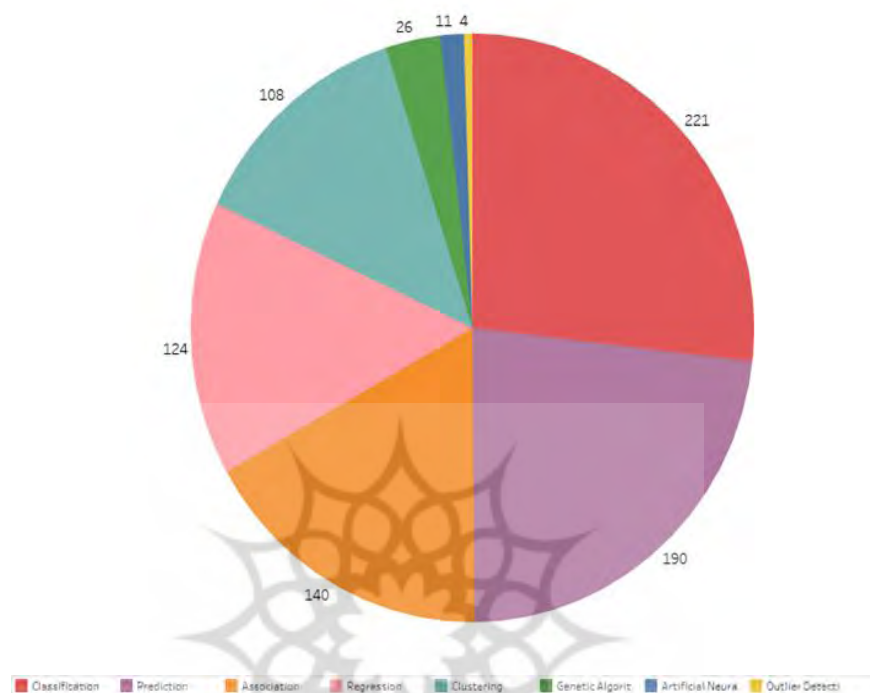


با توجه به شناسایی اجزاء و عوامل مؤثر در مطالعات داده کاوی کووید-۱۹، می توان بر اساس نتایج آن، شکل ۴ نقشه داده کاوی کووید-۱۹ بر اساس مدل سازی موضوعی را برای پیاده سازی مطالعات آتی ارائه داد. اجزای شناسایی شده شامل چهار دسته کلی ساختار و ماهیت انسانی بیماری، زمینه و ساختار بیماری، پیشگیری پاندمیک و هوش مصنوعی قرار دارد. هر کدام از این موارد می تواند به عنوان عوامل مرتبط با کووید -۱۹ شناسایی شوند و لذا پیاده سازی همه اجزاء به صورت موازی می تواند در نقشه دانش مطالعات داده کاوی کووید-۱۹ مؤثر باشد.

شکل ۴. نقشه داده کاوی کووید-۱۹ بر اساس مدل سازی موضوعی



در بررسی میزان تکرارپذیری متدهای داده کاوی در نتیجه فرآیند متن کاوی و استفاده از ابزار Concordance در نرم افزار اورنج، شکل شماره ۵ نشان دهنده پرکاربردترین نوع الگوریتم های داده کاوی به همراه میزان تکرارپذیری در تحقیقات کووید-۱۹ می باشد.



شکل ۵. نمودار میزان تکرارپذیری متدهای داده‌کاوی در تحقیقات کووید-۱۹

بحث و نتیجه‌گیری

برای اینکه اجزا و روابط میان مفاهیم داده‌کاوی در حوزه‌های مختلف مطالعات کووید-۱۹ مورد شناسایی قرار گیرد، با استفاده از روش متن‌کاوی و بیلیومتریک اقدام به شناسایی آن‌ها گردید. در نتیجه فرآیند متن‌کاوی و اعمال الگوریتم‌های مختلف می‌توان گفت داده‌کاوی از پرکاربردترین روش‌های تحقیق در ۴ حوزه ذیل در تحقیقات کووید-۱۹ می‌باشد:

۱. تکنیک‌های تحلیل احساسات و استخراج داده برای درک واکنش‌های عمومی یا تأثیر

بحران

۲. پیش‌بینی نتایج یا تجزیه و تحلیل داده‌های بیماران

۳. چگونگی انتشار اطلاعات از طریق رسانه‌ها و واکنش‌های عمومی

۴. مطالعه جنبه‌های مختلف بحران

در نتیجه بررسی نقشه هم‌رخدادی کلمات کلیدی، مدل‌سازی موضوعی، به ۴ زیر دسته ذیل تقسیم شد:

دسته اول با عبارات و مفاهیمی مانند یادگیری ماشین، شبکه‌سازی اجتماعی، رسانه‌های اجتماعی، متن کاوی، پیش‌بینی و تحلیل احساسات دیده می‌شود، کلمه مرکزی کووید-۱۹ است. همان‌طور که از کلمات و مفاهیم این دسته مشخص است، این دسته در حوزه رسانه‌های اجتماعی، یادگیری ماشین و به عبارتی حول مفاهیم مربوط به هوش مصنوعی است. از این رو در تبیین این دسته می‌توان گفت که از آنجا که کووید-۱۹ در عصر انقلاب صنعتی چهارم و پنجم به وجود آمد و از آنجا که در این عصر هر روز در زمینه پیشرفت و توسعه تکنولوژی تحولاتی در سطح دنیا رخ می‌دهد، لذا قسمتی از تحقیقات و مطالعات مربوط به داده کاوی در زمینه کووید-۱۹ به حوزه علوم کامپیوتری و هوش مصنوعی ارتباط دارد. در این حوزه پژوهش‌های مختلفی از طریق داده کاوی جهت پیش‌بینی کووید-۱۹ در کشورهای مختلف انجام شده است. به همین دلیل منطقی به نظر می‌رسد که بخشی از مطالعات را در حوزه علوم کامپیوتری دسته‌بندی کرد؛ بنابراین یکی از اجزای مهم تحقیقات مربوط به داده کاوی در زمینه مطالعات مربوط به کووید-۱۹، علوم کامپیوتر و هوش مصنوعی است.

دسته دیگری که در زمینه مطالعات مربوط به داده کاوی در زمینه کووید-۱۹، مورد شناسایی قرار گرفت، مربوط به مطالب و مفاهیمی از جمله سلامت عمومی، واکسیناسیون، روانشناسی، افکار مردم، ارتباطات، ترس و عواطف دیده می‌شود. مجموعه این خوشه را می‌توان با عنوان پاندمیک دانست. این دسته از پژوهش‌ها به ارتقای سلامت عمومی جامعه اختصاص دارد و پژوهش‌هایی که نحوه واکسیناسیون، استفاده از اصول روانشناسی برای کاهش ترس مردم از بیماری پاندمیک، همچنین نحوه ارتباطات مردم جهت پیشگیری از این ویروس پاندمیک و بالا بردن اطلاعات عمومی مردم جهت مقابله با خطرات این

بیماری پرداخته است. از این رو، ارتقای آگاهی مردم می‌تواند باعث ارتقای سلامتی آنان شود.

دسته سوم از مطالعات با مفاهیم و کلماتی مانند؛ ژنوم ویروس، دارو، داروهای طبقه‌بندی نشده، مکانیسم دارو، ژنتیک، متابولیسم، سلول انسان، فعالیت ضدویروسی، دیتابیس، پروتئین و اتصال مولکولی و غیره تشکیل شده است. این بخش از پژوهش‌ها به ماهیت بیماری، سلول‌های انسانی اثرگذار در این بیماری، داروهای اثرگذار در این بیماری، ژنتیک و روابط مولکولی در این بیماری اختصاص دارد؛ بنابراین این دسته از متغیرها را می‌توان به عنوان متغیرهای وابسته تلقی کرد که سایر متغیرها بر آن اثرگذار هستند. از این رو می‌توان این دسته از متغیرها را با عنوان «ساختار و ماهیت انسانی بیماری» تلقی کرد.

دسته چهارم از مطالعات مربوط به داده‌های کووید-۱۹، متغیرها و مفاهیمی مانند؛ بیماری، کودک، بزرگسال، عوامل خطر، مرگ، درد، سن، عفونت، تب، الگوریتم، سلامت دارویی، سردرد، چاقی، نوزاد، پیری و داروهای از جمله آزیترامین، فابپیرایور، رم‌دیسویور می‌باشد. مجموعه این عوامل را می‌توان با بزرگسالی در ارتباط دانست. لذا ماهیت این متغیرها را می‌توان به عنوان زمینه‌ساز جذب ویروس کووید-۱۹ دانست. از این رو، عواملی مانند تب، سردرد، چاقی، پیری به عنوان عوامل جذب کووید-۱۹ دانست. همچنین داروهای پر استفاده در این زمینه مانند آزیترامین، فابپیرایور و رم‌دیسویور می‌باشد.

پرکاربردترین متد داده‌های کووید-۱۹ با بیشترین میزان تکرارپذیری الگوریتم Classification با تعداد ۲۲۱ تکرار می‌باشد.

پیشنهاد کاربردی: با توجه به نقشه همکاری کشورهای پیشنهاد می‌شود کشورهای پیشرو در تحقیقات برای مشارکت و بررسی تحقیقات مرتبط با کشور ایران مورد استفاده قرار گیرند.

پیشنهاد پژوهشی: با توجه به تحقیق پیش رو عنوان «به کارگیری مدل‌های داده‌های کووید-۱۹» جهت استخراج دانش مرتبط با پیش‌بینی کووید-۱۹ برای تحقیقات آتی پیشنهاد می‌گردد.

سپاسگزاری

مقاله حاضر مستخرج از پایان نامه کارشناسی ارشد بود و در اینجا فرصتی است که از داوران محترم پایان نامه مربوطه و همچنین داور گرامی مجله مطالعات کسب و کار هوشمند که با بیان نظرات ارزشمند خود به هر چه بهتر شدن این پژوهش کمک شایانی نمودند، تشکر و قدردانی گردد.

تعارض منافع

تعارض منافع ندارم.

ORCID

Roya Abadeh

Alireza Hassanzadeh

Shaban Elahi



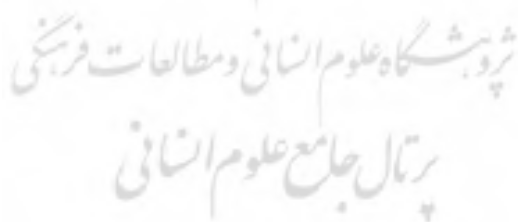
<http://orcid.org/0009-0005-7978-0648>



<http://orcid.org/0000-0003-3015-3019>



<http://orcid.org/0000-0002-1598-4431>



منابع

- مقدس، حمید؛ حسینی، اعظم السادات؛ اسدی، فرخنده؛ جهانبخش، مریم. (۱۳۹۱). داده کاوی و کاربرد آن در سلامت. *مدیریت اطلاعات سلامت*، ۹، ۲، ۲۹۷-۳۰۴.
- سهیلی، فرامرز؛ شعبانی، علی؛ خاصه، علی اکبر. (۱۳۹۴). ساختار فکری دانش در حوزه رفتار اطلاعاتی: مطالعه هم واژگانی. *تعامل انسان و اطلاعات*، جلد ۲، شماره ۴، ۲۱-۳۶.
- شفیعی نیکابادی، محسن؛ اسمعیل زاده، رویا؛ آبروش، مینا. (۱۴۰۲). طراحی مدلی پویا مبتنی بر متن کاوی و روش های نرم جهت تعیین مهمترین عوامل کلیدی مدل های کسب و کار الکترونیک. *مطالعات مدیریت کسب و کار هوشمند*، سال دوازدهم، شماره ۴۶، زمستان، ص ۴۱ تا ۸۰. doi: 10.22054/ims.2023.69067.2199
- محمدزاده، زینب؛ مسرت، الهام؛ کریمی نژاد اصل، رضا. (۱۴۰۰). کاربرد مدل ها، رویکردها و ابزارهای فناوری اطلاعات در مدیریت کووید-۱۹: مرور سریع. *تصویر سلامت*، ۱۲(۱): ۷۷-۹۵. doi: https://doi.org/10.34172/doh.2021.09.45

References

1. Anderson, B. S. (2021). Using text mining to glean insights from COVID-19 literature. *Journal of Information Science*. <https://doi.org/10.1177/01655515211001661>
2. Antons, D., Grünwald, E., Cichy, P., & Salge, T. O. (2020). The application of text mining methods in innovation research: current state, evolution patterns, and development priorities. *R and D Management*, 50(3), 329–351. <https://doi.org/10.1111/radm.12408>.
3. Chalortham, N., Leesawat, P., Buranarach, M. and Supnithi, T., 2008, May. Ontology development for pharmaceutical tablet production expert system. In (2008) 5th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (Vol. 1, pp. 205-208). IEEE. DOI: 10.1109/ECTICON.2008.4600408
4. Chen, Y.L., Analytis, J.G., Chu, J.H., Liu, Z.K., Mo, S.K., Qi, X.L., Zhang, H.J., Lu, D.H., Dai, X., Fang, Z. and Zhang, S.C., (2009). Experimental realization of a three-dimensional topological insulator, Bi₂Te₃. *Science*, 325(5937), pp.178-181. DOI: 10.1126/science.1173034.
5. Dastani, M., & Danesh, F. (2021). Iranian COVID-19 Publications in

- LitCovid: Text Mining and Topic Modeling. *Scientific Programming*, 2021, 3315695. <https://doi.org/10.1155/2021/3315695>
6. Davis, R., Shrobe, H. and Szolovits, P., 1993. What is a knowledge representation?. *AI magazine*, 14(1), pp.17-17 <https://doi.org/10.1609/aimag.v14i1.1029>.
 7. Elsayed, A.E., El-Beltagy, S.R., Rafea, M. and Hegazy, O., (2007). Applying data mining for ontology building. Proc. of ISSR. https://www.academia.edu/14323140/Applying_data_mining_for_ontology_building
 8. Eppler, M.J., (2004). Making knowledge visible through knowledge maps: concepts, elements, cases. Handbook on knowledge management 1: *Knowledge matters*, pp.189-205. https://doi.org/10.1007/978-3-540-24746-3_10
 9. Eppler, M.J., 2006. A comparison between concept maps, mind maps, conceptual diagrams, and visual metaphors as complementary tools for knowledge construction and sharing. *Information visualization*, 5(3), pp.202-210. DOI:10.1057/palgrave.ivs.9500131.
 10. Gandhi, P. and Pruthi, J., (2020). Data visualization techniques: traditional data to big data. *Data Visualization: Trends and Challenges Toward Multidisciplinary Perception*, pp.53-74. https://doi.org/10.1007/978-981-15-2282-6_4.
 11. Gorskis, H. and Chizhov, Y., (2012). Ontology building using data mining techniques. *Information technology and management science*, 15(1), pp.183-188. DOI:10.2478/v10313-012-0024-5.
 12. Gorskis, H. and Chizhov, Y., (2012). Ontology building using data mining techniques. *Information technology and management science*, 15(1), pp.183-188. DOI:10.2478/v10313-012-0024-5.
 13. Ho, M. H.-C., & Liu, J. S. (2021). The swift knowledge development path of COVID-19 research: the first 150 days. *Scientometrics*, 126(3), 2391–2399. <https://doi.org/10.1007/s11192-020-03835-5>
 14. Jiang, C., Ngo, V., Chapman, R., Yu, Y., Liu, H., Jiang, G. and Zong, N., (2022). Deep Denoising of Raw Biomedical Knowledge Graph From COVID-19 Literature, LitCovid, and Pubtator: Framework Development and Validation. *Journal of medical Internet research*, 24(7), p.e38584. DOI: 10.2196/385841.
 15. Kim, S., Suh, E. and Hwang, H., (2003). Building the knowledge map: an industrial case study. *Journal of knowledge management*, 7(2), pp.34-45. DOI: 10.1108/136732703104772702.
 16. Levesque, H.J., (1986). Knowledge representation and reasoning. *Annual review of computer science*, 1(1), pp.255-287. DOI:

10.1146/annurev.cs.01.060186.0013513.

17. Mansingh, G., Osei-Bryson, K.M. and Reichgelt, H., (2009). Building ontology-based knowledge maps to assist knowledge process outsourcing decisions. *Knowledge Management Research & Practice*, 7(1), pp.37-51. DOI: 10.1057/kmrp.2008.374
18. Mohammadzadeh, Z., Maserat, E., & Kariminezhad, R. (2021). Application of Information Technology Models, Approaches and Tools in Covid-19 Management: Rapid Review. *Depiction of Health*, 12(1), 77–95. <https://doi.org/10.34172/doh.2021.09>
19. Ngo, Q.Q., Dennig, F.L., Keim, D.A. and Sedlmair, M., (2022). Machine learning meets visualization—Experiences and lessons learned. *it-Information Technology*, 64(4-5), pp.169-180. DOI: 10.1515/itit-2022-00341.
20. Panov, P., Soldatova, L. N., & Džeroski, S. (2009). Towards an ontology of data mining investigations. Lecture Notes in Computer Science (Including Subseries *Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics*), 5808 LNAI(October), 257–271. DOI: 10.1007/978-3-642-04747-3_212.
21. Peim, N., (2001). The history of the present: towards a contemporary phenomenology of the school. *History of Education*, 30(2), pp.177-190. DOI: 10.1080/004676000100124543.
- A. L., Zhang, Y., Huang, Y., & Wu, M. (2020). Tracking and Mining the COVID-19 Research Literature. *Frontiers in Research Metrics and Analytics*, 5(November 2020), 1–18. <https://doi.org/10.3389/frma.2020.594060>
22. Safdari, R., Rezayi, S., Saedi, S., Tanhapour, M., & Gholamzadeh, M. (2021). Using data mining techniques to fight and control epidemics: A scoping review. *Health and Technology*, 11(4), 759–771. <https://doi.org/10.1007/s12553-021-00553-7>

References (In Persian)

1. Moghaddassi H, Hoseini A, Asadi F, Jahanbakhsh M. Application of Data Mining. *Health Information Management* 2012; 9(2): 304. (In Persian)
2. Mohammadzadeh Z, Moserrat A, Kariminjad Asl R. Application of models, approaches and information technology tools in COVID-19 management: A quick review. *Health Image* 2022; 30(9): 15-66. SID. <https://sid.ir/paper/385727/fa> (In Persian)
3. Shafiei Nikabadi, Mohsen, Esmaeilzadeh, Roya, & Abfroosh, Mina. (2023). Designing a dynamic model based on text mining and soft

methods to determine the most important key factors of e-business models. *Studies in Smart Business Management*, 12(46), . doi: 10.22054/ims.2023.69067.2199 (In Persian)

4. Soheili Faramarz, Shabani Ali, Khasse Ali Akbar. The structure of thought knowledge in the field of information behavior: A study of homonymy. *Human and Information Interaction*. 1394; 2 (4). URL: <http://hii.khu.ac.ir/article-1-2446-fa.html>. (In Persian)



استناد به این مقاله: آباده، رؤیا، حسن‌زاده، علیرضا، الهی، شعبان. (۱۴۰۳). استخراج و مصورسازی نقشه دانش

داده‌کاوی در تحقیقات کووید-۱۹، مطالعات مدیریت کسب و کار هوشمند، ۱۳(۴۹)، ۲۵۵-۲۸۱. DOI:

10.22054/ims.2024.77278.2432



Journal of Business Intelligence Management Studies is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License..



پروہشگاہ علوم انسانی و مطالعات فرہنگی
پرتال جامع علوم انسانی