

Classification of Comments Related to Virtual Reality Technology Using Topic Modeling

Fariba Karimi 

Master of Information Technology Management e-business, Faculty of Social Sciences and Economics, Alzahra University, Tehran, Iran

Ameneh Khadivar *

Associate Professor, Department of Management, Faculty of Social Sciences and Economics, Alzahra University, Tehran, Iran

Fatemeh Abbasi 

Assistant Professor, Faculty of Management and Accounting, Shahid Beheshti University, Tehran, Iran

Abstract

In recent years, the rapid growth of virtual space has made people devote more of their time in virtual space, especially to social networks, which can be attributed to the remarkable features of virtual space; including increasing the speed of information exchange, easy and free access to information and variety of knowledge topics. In this regard, the opinions recorded by users in virtual networks have grown day by day and have become very important, and extracting the opinions and feelings of users' opinions for more informed decision-making is of great help to businesses, on the other hand, virtual reality technology in the past few decades It has undergone technical changes and improved immersion and the feeling of remote presence; This technology is used in various fields such as education, tourism, health, sports, entertainment, architecture and construction, etc. The increasing progress of virtual reality technology has caused many businesses to operate in this field, but due to changes Continuous market and the need for timely information, companies should use differentiation and growth strategies, in this regard, they need to ask users' opinions and in line with that, try to grow and improve their business, considering that Users' comments are textual,

* Corresponding Author: a.khadivar@alzahra.ac.ir

How to Cite: Karimi, F., Khadivar, A., Abbasi, F. (2024). Classification of Comments Related to Virtual Reality Technology Using Topic Modeling, *Journal of Business Intelligence Management Studies*, 12(47), 1-43.

and reading and summarizing them is time-consuming and difficult. Based on this, the aim of the current research was to categorize comments related to virtual reality technology using machine learning methods and a dictionary-based approach. Therefore, about one million tweets in the field of virtual reality technology were collected by the web crawler, and after data preprocessing, 480,432 samples remained in the data, then Dirichlet's hidden allocation topic modeling was implemented on the data. This modeling separated different topics by examining the distribution of words in tweets; The tweets whose distribution of words were similar were placed into a topic and the number of topics with the highest coherence score was selected, the number of topics 9 had higher coherence and the data were grouped into 9 topics, so once again the Dirichlet hidden allocation modeling was set to 9. The topic was done, with this the tweets were grouped into 9 different topics. To evaluate the model, considering that we had a probability distribution, the confusion criterion was used, the value of which was -9.44, and the coherence score was used for the degree of semantic similarity between words and the distinction between subjects, and the result was 0.47. The lower the confusion criterion and the higher the coherence score, the more efficient the model is. With the help of keyword weights obtained by Dirichlet hidden allocation modeling and examining at least 5 different tweets from each topic, 9 topics related to virtual reality technology were identified: "New Technology", "Creation and Make", "Technological Business", "Education", "Virtual Games", "Progress", "Gadget", "Metaverse", and "Indiegame", the topics were analyzed with the help of several graphs. We found that the number of neutral comments on topics such as "New Technology" and "Metaverse" is more than positive and negative comments, which indicates the lack of sufficient information or the lack of use of these technologies, and it is necessary for businesses in this field, to try more in this regard, in the same way, if we observe the graph of "Virtual Games" and "Technological Business", we can see that it changes almost with the same ratio in different years, in the sense that this The two graphs are related, in fact, businesses should keep in mind that the factors affecting these two issues are the same, but users pay more attention to the issue of "Virtual Games", as a result, if the creators of "Technological Business" Focus specifically on "Virtual Games", they will grow more due to the more attention of users, also the creators of games should consider that "Virtual Games" are a topic of more attention than "Indiegame". Is. In the subjects of "Education" and "Gadget", users lost their attention to these subjects in the field of virtual reality over time, in fact they showed their

attention to other subjects, so it is better for businesses that operate in this field to take measures To advertise and attract users or change their user area if there is no growth.

1. Introduction

Constant changes in the market and the need for timely information force companies to use differentiation and growth strategies appropriate to the needs of customers. (Sánchez, Folgado-Fernández, & Sánchez, 2022). Companies can check and analyze their customers' opinions through microblogging sites (Facebook, Twitter, etc.) and finally improve the desired products or services (Ahmad, Aftab, Bashir, & Hameed, 2018). Today, users express their opinions and feelings and review products in online social networks. Therefore, user comments and the analysis of these comments have become a valuable resource for businesses (Kim et al., 2015; Loureiro et al., 2019).

Virtual reality and augmented reality have undergone technical developments in the past few decades and have improved immersion and the feeling of remote presence. Several examples of applications of such techniques can be found in stores, the tourism industry, hotels, restaurants, etc. (Loureiro, Guerreiro, & Ali, 2020). Due to the constant changes in the market and the need for timely information, companies should use differentiation and growth strategies, nowadays, due to the rapid evolution of the Internet, instead of collecting their opinions through time-consuming and expensive methods such as questionnaires and interviews, etc., they express in the context of social networks, which is very useful for businesses in their development, and they can measure the feelings of customers towards products and services, and understand the needs of users, and finally make appropriate and appropriate decisions in the direction of adopt growth, but in order to use the produced content correctly, text mining and sentiment analysis techniques should be used, which has not been researched in Iran so far. Analysis of users' opinions and feelings about virtual reality technology can help businesses that operate in the field of metaverse, virtual game production, virtual education, virtual tourism, etc., to make better decisions and plans.

2. Literature Review

Social media generates a large amount of real-time social signals that can provide new insights into human behavior and emotions. People around the world are constantly engaged with social media. (Al-Samarraie, Sarsam, & Alzahrani, 2023).

On the other hand, the amount of data is increasing day by day. Almost all institutions, organizations and business industries store their data electronically. A huge amount of text is circulating on the Internet in the form of digital libraries, repositories, and other textual information such as blogs, social media networks, and emails (Sagayam, Srinivasan, & Roshni, 2012).

Topic modeling is one of the most powerful techniques in text mining for data mining, discovering hidden data and finding relationships between data and textual documents (Jelodar et al., 2017).

The technological advances of the last century have confronted societies with new realities that have indisputably improved daily life, making it more convenient and interesting. In recent decades, technology using virtual reality and wearable devices have had a significant impact in the fields of education, tourism, health, sports, entertainment, architecture and construction, etc. (Kosti et al., 2023).

Virtual reality is a technology that allows a user to interact with a computer-simulated environment, whether that environment is a simulation of the real world or an imaginary one. With virtual reality, we can experience the most frightening and overwhelming situations with safe play and a learning perspective (Mandal, 2013). Most people are curious about the possibilities and future of new technologies, considering the various applications it is supposed to offer such as virtual meetings, learning environments and many others, however, there are also concerns about potential negative effects. because real world signals can be transmitted in the virtual world. In this regard, people express their feelings in different social networks (Bhattacharyya et al., 2023).

3. Methodology

According to the main goal of the research, which is to classify comments related to virtual reality technology using machine learning methods and a dictionary-based approach, therefore, about one million

tweets in the field of virtual reality technology were collected by the web crawler and After data preprocessing, 480,432 samples remained in the data, then Dirichlet hidden allocation thematic modeling was implemented on the data. By examining the distribution of words in tweets, this modeling tries to separate different topics by detecting the distribution of words; The tweets whose distribution of words are similar were put into a topic, and the number of topics with the highest score was selected, the number of topics 9 has higher coherence, and the data was grouped into 9 topics, so once again, Dirichlet hidden allocation modeling was applied 9 topics were done, whereby the tweets were grouped into 9 different topics. Considering that we have a probability distribution, the confusion criterion was used to evaluate the model. The lower the confusion criterion and the higher the coherence score, the more efficient the model is. With the help of keyword weights obtained by Dirichlet hidden allocation modeling and examining at least 5 different tweets from each topic, 9 topics related to virtual reality technology were identified: "New Technologies", "Creation and Make", "Technological Business", "Education", "Virtual Games", "Progress", "Gadget", "Metaverse" and "Indiegame" were named.

4. Discussion and Conclusion




In this research, by examining topics in different years, we observed that the topic of "Progress" was the most popular topic among users from 2017 to the end of 2021, in early 2022, this topic gave way to "Metaverse", currently "Metaverse" is one of the most popular topics being discussed by users. Businesses in the field of virtual reality should strive for the attractiveness of "Metaverse" and attract users. Likewise, if we observe the "Virtual Games" and "Technological Business" graphs, we can see that they change with almost the same ratio in different years, meaning that these graphs are related to each other, in fact, business and keep in mind that the factors affecting these two issues are the same, but in the case of "Virtual Games" it has more effects, and if "Technological Businesses" specifically focus on virtual games, they will grow more due to the greater attention of users. had Similarly, "Indiegame" which have had a series of changes but in recent years have had a declining trend and then no change, now the creators of these games should check, and in general "Virtual

Games" are a more interesting topic than "Indiegame". In the subjects of "Education" and "Gadget" it has been decreasing since the beginning of 2017, which shows that users lost their attention to these subjects in the field of virtual reality over time, in fact to other topics showed their attention, so it is better for businesses that are active in this field to take measures to advertise and attract users, or change their user field if there is no growth.

Keywords: Data Mining, Text Mining, Virtual Reality Technology, Topic Modeling, Latent Dirichlet Allocation.



دسته‌بندی نظرات مرتبط با فناوری واقعیت مجازی با استفاده از مدل سازی موضوعی

- فریبا کریمی  کارشناسی ارشد مدیریت فناوری اطلاعات گرایش کسب و کار الکترونیک، دانشکده علوم اجتماعی و اقتصادی، دانشگاه الزهرا(س)، تهران، ایران
- * آمنه خدیور  دانشیار، گروه مدیریت، دانشکده علوم اجتماعی و اقتصادی، دانشگاه الزهرا(س)، تهران، ایران
- فاطمه عباسی  استادیار، گروه مدیریت صنعتی و فناوری اطلاعات، دانشکده مدیریت و حسابداری، دانشگاه شهید بهشتی، تهران، ایران

چکیده

در سال‌های اخیر رشد سریع فضای مجازی، باعث شده افراد بیشتر زمان خود را در فضای مجازی به‌ویژه فعالیت در شبکه‌های اجتماعی اختصاص دهند، در این راستا نظرات ثبت شده توسط کاربران در شبکه‌های مجازی، رشد روزافزونی داشته و اهمیت بسیاری پیدا کرده است. استخراج عقاید و احساسات نظرات کاربران جهت تصمیم‌گیری آگاهانه‌تر کمک زیادی به کسب و کارها می‌کند، بر این اساس، هدف پژوهش حاضر تحلیل و بررسی نظرات کاربران توئیتر درباره‌ی فناوری واقعیت مجازی با بهره‌گیری از روش‌های یادگیری ماشین و رویکرد مبتنی بر واژه‌نامه بوده است. از این رو حدود یک میلیون توئیتر در زمینه فناوری واقعیت مجازی توسط خزشگر وب جمع‌آوری گردید و بعد از پیش‌پردازش، مدل‌سازی موضوعی تخصیص پنهان دیریکله روی داده‌ها اجرا شد. این مدل‌سازی با استفاده از تشخیص توزیع کلمات، سعی می‌کند موضوعات مختلف را از هم جدا کند؛ تعداد موضوعات ۹ دارای انسجام بالاتری بود (۰٫۴۷) و داده‌ها در ۹ موضوع مختلف دسته‌بندی شدند. برای ارزیابی مدل نیز با توجه به اینکه یک توزیع احتمالی داشتیم، از معیار سرگشتگی استفاده شد که مقدار آن ۹/۴۴- به دست آمد و که هرچه قدر معیار سرگشتگی پایین‌تر و امتیاز انسجام بالاتر نشان از کارایی مدل دارد. نتایج نشان می‌دهد تعداد نظرات خنثی در موضوعاتی مثل فناوری‌های جدید و متاورس بیشتر از نظرات مثبت و منفی است که نشان از عدم اطلاعات کافی و یا عدم استفاده از این فناوری‌ها می‌باشد و لازم است کسب و کارها در این خصوص بیشتر تلاش نمایند.

کلیدواژه‌ها: داده کاوی، متن کاوی، فناوری واقعیت مجازی، مدل‌سازی موضوعی تخصیص پنهان دیریکله.

مقاله حاضر برگرفته از پایان‌نامه ارشد رشته مدیریت فناوری اطلاعات دانشگاه الزهرا است.

* نویسنده مسئول: a.khadivar@alzahra.ac.ir

مقدمه

تغییرات مداوم بازار و نیاز به اطلاعات به موقع، شرکت‌ها را مجبور می‌کند تا از استراتژی‌های تمایز و رشد مناسب متناسب با نیازهای مشتریان استفاده کنند. (Sánchez, Folgado-Fernández, & Sánchez, 2022). شرکت‌ها می‌توانند از طریق سایت‌های میکرو بلاگینگ^۱ (فیس‌بوک، توئیتر و غیره) نظرات مشتریان خود را بررسی و تجزیه و تحلیل کنند و در نهایت محصولات یا خدمات مورد نظر را بهبود ببخشند (Ahmad, Aftab, Bashir, & Hameed, 2018). با ظهور رسانه‌های اجتماعی مردم ایده‌ها، تجربیات، نظرات و مقاصد خود را در توئیتر، فیس‌بوک و دیگر پلتفرم‌های رسانه اجتماعی به اشتراک می‌گذارند. تحلیل و بررسی این نظرات در رسانه‌های اجتماعی اطلاعات سودمندی را در رابطه با بازار، مشتریان و رقبا به سازمان‌ها می‌دهد (نوروزی و همکاران، ۱۴۰۲).

امروزه کاربران نظرات و احساسات خود را بیان می‌کنند و محصولات را در شبکه‌های اجتماعی آنلاین بررسی می‌کنند؛ بنابراین، نظرات کاربر و تجزیه و تحلیل این نظرات به یک منبع ارزشمند برای کسب و کارها تبدیل شده است (Kim et al., 2015; Loureiro et al., 2019). در سال‌های اخیر استفاده از خدمات آنلاین برای انتخاب و رزرو هتل شاهد رشد پررونقی بوده است. نظرات مشتریان جایگزین بازاریابی دهان‌به‌دهان شده‌اند (Abbasi et al., 2019). با توجه به آنکه نظرات کاربران به صورت متنی است و خواندن و جمع‌بندی آن‌ها زمان‌بر و مشکل است، خودکارسازی استخراج عقاید و احساسات نظرات کاربران یکی از راهکارهای پیشنهادی برای سایت‌های فروش آنلاین جهت ارائه خدمات کارا تر به مشتریان جهت تصمیم‌گیری آگاهانه‌تر است (عباسی و همکاران، ۱۳۹۶).

تکنیک‌های داده کاوی و متن کاوی به عنوان ابزاری قدرتمند برای تجزیه و تحلیل این

۱. Microblogging: نوعی وبلاگ نویسی است که به کاربران امکان نوشتن متن‌های کوتاه و منتشر کردن آن را

داده‌ها در صنایع مختلف استفاده می‌شوند (Liao et al., 2012). حوزه‌های کاربردی مانند موتورهای جستجو، سیستم مدیریت ارتباط با مشتری، فیلتر ایمیل‌ها و تجزیه و تحلیل رسانه‌های اجتماعی از متن کاوی برای استخراج ویژگی، احساسات، پیش‌بینی و تحلیل روند استفاده می‌کنند (He, 2013).

واقعیت مجازی^۱ و واقعیت افزوده^۲ در چند دهه گذشته دستخوش تحولات فنی شده‌اند و باعث بهبود در غوطه‌وری و احساس حضور از راه دور شده‌اند. چندین نمونه از کاربردهای چنین تکنیک‌هایی را می‌توان در فروشگاه‌ها، صنعت گردشگری، هتل‌ها، رستوران‌ها و غیره یافت (Loureiro et al., 2020). در این راستا باید به واقعیت ترکیبی^۳ نیز اشاره کرد که یکی از فناوری‌های نوظهور جهان است که از توسعه و ترکیب دو فناوری واقعیت مجازی و واقعیت افزوده شکل گرفته، واقعیت ترکیبی با حداکثر تعامل کاربر در دنیای واقعی در مقایسه با سایر فناوری‌های مشابه سروکار دارد و پتانسیل ایجاد تحول عظیم در آینده را دارد (Rokhsaritalemi et al., 2020). تجربه فروشگاه واقعیت مجازی باعث بیشتر برانگیخته شدن احساسات مثبت مشتریان می‌شود و جذابیت فروشگاه را افزایش می‌دهد (Jin et al., 2021).

طبق بررسی‌های انجام‌شده، واقعیت مجازی در زمینه‌های مختلفی کاربرد دارد، همین‌طور پیشرفت روزافزون فناوری واقعیت مجازی، باعث شده کسب و کارها زیادی در این زمینه فعالیت کنند، به دلیل تغییرات مداوم بازار و نیاز به اطلاعات به‌موقع، شرکت‌ها باید از استراتژی‌های تمایز و رشد استفاده کنند، امروزه با توجه به تکامل سریع اینترنت افراد به جای اینکه نظراتشان را از طریق روش‌های زمان‌بر و پرهزینه مثل پرسش‌نامه و مصاحبه و غیره بیان کنند، در بستری از شبکه‌های اجتماعی بیان می‌کنند که برای کسب و کارها در جهت پیشرفتشان بسیار مفید است و می‌تواند احساسات مشتریان را نسبت به محصولات و خدمات بسنجند و نیازهای کاربران را درک کنند و در نهایت تصمیمات

-
1. Virtual Reality (VR)
 2. Augmented Reality (AR)
 3. Mixed Reality (MR)

درخور و مناسبی در جهت رشد اتخاذ کنند، ولی برای اینکه محتوای تولیدشده را بتوان به درستی به کار برد، باید از تکنیک‌های متن‌کاوی و تجزیه و تحلیل احساسات استفاده کرد که تاکنون پژوهشی در این زمینه در ایران صورت نگرفته است. تحلیل نظرات و احساسات کاربران درباره فناوری واقعیت مجازی می‌تواند به کسب و کارهایی که در زمینه متاورس^۱، تولید بازی‌های مجازی، آموزش مجازی، گردشگری مجازی و غیره فعالیت می‌کنند، برای تصمیم‌گیری و برنامه‌ریزی بهتر کمک نماید. با توجه به اهمیت موضوع و خلأ شناسایی‌شده در این پژوهش از طریق دسترسی به تئیت‌های کاربران در شبکه اجتماعی توییتر و استفاده از روش‌های یادگیری ماشین، تجزیه و تحلیل احساسات و مدل‌سازی موضوعی می‌خواهیم به دسته‌بندی نظرات مرتبط با فناوری واقعیت مجازی بپردازیم.

پیشینه پژوهش

پیشینه نظری پژوهش

رسانه‌های اجتماعی حجم زیادی از سیگنال‌های اجتماعی را در لحظه تولید می‌کنند که می‌تواند بینش جدیدی در مورد رفتار و احساسات انسان ارائه دهد، مردم در سراسر جهان به طور مداوم با رسانه‌های اجتماعی درگیر هستند. تعامل مردم با رسانه‌های اجتماعی همیشه به‌عنوان منبعی در نظر گرفته شده است که می‌تواند به سیاست‌گذاران، کسب و کارها و سازمان‌ها کمک کند تا بفهمند یک فناوری یا خدمات چقدر می‌تواند یکپارچه یا مورد استفاده قرار گیرد (Al-Samarraie et al., 2023).

از طرفی حجم داده‌ها روزبه‌روز در حال افزایش است. تقریباً همه مؤسسات، سازمان‌ها و صنایع تجاری داده‌های خود را به‌صورت الکترونیکی ذخیره می‌کنند. حجم عظیمی از متن به شکل کتابخانه‌های دیجیتال، مخازن و سایر اطلاعات متنی مانند وبلاگ‌ها، شبکه‌های رسانه‌های اجتماعی و ایمیل‌ها در اینترنت در جریان است (Sagayam et al., 2012). تعیین الگوها و روندهای مناسب برای استخراج دانش ارزشمند از این حجم

عظیم داده، کار چالش‌برانگیزی است (Padhy et al., 2012). ابزارهای داده‌کاوی سنتی قادر به مدیریت داده‌های متنی نیستند زیرا استخراج اطلاعات به زمان و تلاش نیاز دارد (Talib et al., 2016). داده‌کاوی رویکرد پردازش اطلاعات است که از فناوری پایگاه داده استفاده می‌کند. فناوری پایگاه داده یک علم نرم‌افزاری است که در مورد مدیریت و اعمال پایگاه‌های داده تحقیق می‌کند (Yang et al., 2020). داده‌کاوی برای کمک به سازمان‌ها، جهت تمرکز بر مهم‌ترین اطلاعات در ابزارهای داده ذخیره‌شده مورد استفاده قرار می‌گیرد. تکنیک‌های داده‌کاوی می‌توانند پاسخگوی سؤالاتی باشند که به کسب‌وکار مربوط می‌شوند و حل آن‌ها معمولاً زمان‌بر است (Osman, 2019).

مدل‌سازی موضوعی یکی از قوی‌ترین تکنیک‌ها در متن‌کاوی برای داده‌کاوی، کشف داده‌های پنهان و یافتن روابط بین داده‌ها و اسناد متنی است (Jelodar et al., 2017). مدل‌سازی موضوعی یک یادگیری بدون ناظر است؛ در حقیقت مدل‌سازی موضوع روش‌هایی را برای شناسایی کلمات کلیدی هم‌زمان برای خلاصه کردن مجموعه‌ای از اطلاعات متنی ارائه می‌دهد. تعدادی الگوریتم برای حل این مشکل وجود دارد، الگوریتم‌های یادگیری ماشین بدون ناظر برای بررسی موضوعات توئیت‌ها شامل: تحلیل معنایی پنهان^۱، فاکتورگیری غیر منفی ماتریس^۲ و تخصیص پنهان دیریکله^۳ و غیره می‌باشند (Islam, 2019).

پیشرفت‌های فناوری قرن گذشته، جوامع را با واقعیت‌های جدیدی روبه‌رو کرده است که زندگی روزمره را به طور غیرقابل بحثی بهبود بخشیده و آن را راحت‌تر و جالب‌تر کرده است. در دهه‌های اخیر، فناوری با استفاده از واقعیت مجازی و دستگاه‌های پوشیدنی تأثیر قابل توجهی در حوزه‌های آموزش، گردشگری، بهداشت، ورزش، سرگرمی، معماری و ساخت‌وساز و غیره داشته‌اند (Kosti et al., 2023).

واقعیت مجازی فناوری است که به کاربر اجازه می‌دهد با یک محیط شبیه‌سازی شده

-
1. Latent Semantic Analysis (LSA)
 2. Non-negative Matrix Factorization (NMF)
 3. Latent Dirichlet Allocation (LDA)

توسط کامپیوتر تعامل داشته باشد، خواه آن محیط شبیه‌سازی دنیای واقعی یا خیالی باشد. با واقعیت مجازی، می‌توانیم ترسناک‌ترین و طاقت‌فرساترین موقعیت‌ها را با بازی ایمن و با دیدگاه یادگیری تجربه کنیم (Mandal, 2013). کریگ و همکارانش به‌درستی واقعیت مجازی را به‌عنوان رسانه‌ای متشکل از شبیه‌سازی‌های کامپیوتری تعاملی که احساس غوطه‌ور شدن را می‌دهد، تعریف می‌کند (Craig et al., 2009). درحالی‌که ژانگ و همکارانش آن را به‌عنوان رابط انسان و ماشین پیشرفته بیان کرده‌اند که فناوری‌هایی مانند گرافیک رایانه‌ای، پردازش تصویر، تشخیص الگو، هوش مصنوعی و غیره را برای شبیه‌سازی و تعامل رایانه‌ای ترکیب می‌کند (Zhang et al., 2018). بسیاری از نام‌های دیگر به جای واقعیت مجازی استفاده می‌شوند که شامل محیط مجازی، واقعیت مصنوعی، دنیای مجازی، دنیای مصنوعی و فضای سایبری^۱ است (Bamodu & Ye, 2013). بسیاری از مفاهیم در حوزه علم نیاز به درک روابط سه‌بعدی دارند. به‌عنوان مثال، در زمینه پزشکی، درک آناتومی انسان از منظر سه‌بعدی نقش مهمی در طول جراحی ایفا می‌کند. بررسی‌های اخیر نشان می‌دهد که فقدان آموزش سه‌بعدی، یادگیری یک مفهوم را برای دانش‌آموزان بسیار چالش‌برانگیز می‌سازد که به نوبه خود، بر دستاوردهای آن‌ها تأثیر منفی می‌گذارد (Boulter & Gilbert, 2000; Harle & Towns, 2011). اکثر مردم در مورد احتمالات و آینده فناوری‌های جدید، با توجه به برنامه‌های کاربردی مختلفی که قرار است ارائه دهد مانند جلسات مجازی، محیط‌های یادگیری و بسیاری موارد دیگر کنجکاو هستند، با این حال، نگرانی‌هایی در مورد اثرات منفی بالقوه نیز وجود دارد زیرا سیگنال‌های دنیای واقعی می‌توانند در دنیای مجازی منتقل شوند. در این راستا، مردم احساسات خود را در شبکه‌های اجتماعی مختلف بیان می‌کنند (Bhattacharyya et al., 2023).

چرخه هایپ گارتنر برای نشان دادن نموداری از مراحل حیات یک فناوری نوظهور استفاده می‌شود؛ در این چرخه، مراحل اولیه نوآوری با افزایش سریع انتظارات از نوآوری به سمت اوج انتظارات می‌رود و سپس کاهش سریع از طریق ناامیدی مشخص می‌شود،

سپس با درک سودمندی فناوری منجر به پذیرش اولیه و رشد سریع بازار که همان سطح سوددهی است می‌رسد (Shi & Herniman, 2023).

پیشینه تجربی پژوهش

ما در مورد نحوه درک کاربران از فناوری‌های جدید، به‌ویژه در مورد خوش‌بینی یا بدبینی آن‌ها، دانش کافی نداریم، در پژوهشی که توسط العداله و همکاران انجام گرفت، برای رفع این شکاف، بیش از ۳۰۰۰۰۰۰ توثیت را در مورد متاورس تجزیه و تحلیل شد، از جمله بیش از ۱۰۰۰ توثیت که کلمات کلیدی امنیت و حریم خصوصی را ذکر کرده بودند. هدف این مطالعه درک چگونگی درک مردم از امنیت و حریم خصوصی متاورس با استفاده از تحلیل احساسات در توثیت‌های جمع‌آوری شده بود. نتایج نشان می‌دهد که تقریباً ۵۹٪ از کاربران دیدگاه مثبتی نسبت به امنیت خود در متاورس داشتند، در حالی که حدود ۶۶٪ دیدگاه مثبتی نسبت به حریم خصوصی خود داشتند (Al-Kfairy et al., 2023).

پژوهشی دیگر توسط کائور و همکاران انجام گرفت که در آن تجزیه و تحلیل احساسات توثیت برای تشخیص افسردگی با استفاده از الگوریتم‌های یادگیری ماشین صورت گرفت، توثیت‌ها ابتدا جمع‌آوری شدند تا مدل ساخته شود و سپس از توثیت‌های پیش‌پردازش شده برای تجزیه و تحلیل احساسات استفاده شد. سپس مجموعه داده به دست آمده به نسبت‌های ۸۰ به ۲۰ تقسیم شدند تا از روش‌های یادگیری ماشین برای شناسایی تعداد توثیت‌هایی با مضمون افسردگی و غیرافسردگی استفاده شود. سپس عملکرد الگوریتم‌های یادگیری ماشین با استفاده از ماتریس‌های ارزیابی مختلف مقایسه شد در آخر نتایج نشان داد که الگوریتم جنگل تصادفی به بالاترین دقت پیش‌بینی ۹۰ درصد دست می‌یابد (Kaur et al., 2023).

در پژوهشی توسط عالی خانی و همکاران، درک کاربران شبکه اجتماعی توثیت به‌عنوان یک بستر میکرو بلاگینگ در مورد نسل پنجم ارتباطات سیار با استفاده از روش‌های یادگیری ماشینی و تحلیل احساسات مورد بررسی قرار گرفت، از این‌رو،

مجموعه‌ای از بیش از ۴۰۰۰۰ توییت در این زمینه از طریق رابط کاربری توئیتر گردآوری شد و مدل شبکه عصبی بازگشتی با دقت ۷۹ درصد ایجاد شد. در نهایت مدل‌سازی موضوع به روش آنالیز تشخیصی خطی برای تعمیق بیشتر انجام شد. نتایج نشان می‌دهد که اگرچه نارضایتی از کیفیت ارائه‌شده، هزینه و پوشش نسل پنجم ارتباطات سیار، نگرانی‌های بهداشتی، شایعات مربوط به کووید ۱۹ با نسل پنجم وجود دارد، اما همچنان افراد بیشتری نسبت به آینده این فناوری در زمینه‌های مختلف مانند اینترنت اشیا و هوش مصنوعی خوش‌بین هستند (عالی‌خانی و همکاران، ۱۴۰۱).

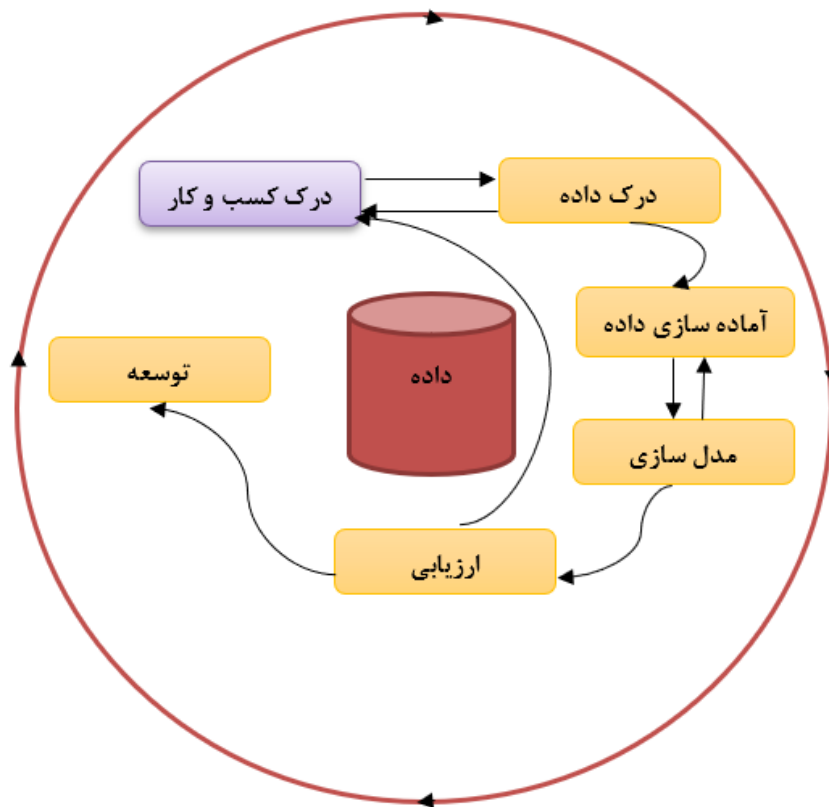
محمدی و ناظمی در پژوهشی ضمن بررسی نظرات کاربران در شبکه اجتماعی توئیتر در مورد ویژگی‌های مختلف دو محصول رقیب تلفن همراه-شرکتهای اپل و سامسونگ- در بازار، احساسات آنان را بر اساس جنسیت مصرف‌کنندگان این دو محصول مورد بررسی قرار داده‌اند. این بررسی با استفاده از روش مبتنی بر رابطه در مرحله استخراج ویژگی و رویکرد مبتنی بر لغتنامه احساسی در مرحله تعیین قطبیت نظرات انجام شد، نتایج این پژوهش، حاکی از آن بود که محبوبیت ویژگی‌های مختلف محصول بین کاربران مرد و زن متفاوت بوده و بر اساس این نتایج، صاحبان کسب‌وکار می‌توانند اقدام به تولید محصولاتی با تمرکز بر جنسیت افراد کرده که در نهایت به افزایش سوددهی کسب‌وکار و رضایتمندی مشتریان منجر می‌شود (محمدی و ناظمی، ۱۴۰۰).

بصیری و همکاران در پژوهشی به تحلیل احساسات موجود در نظرات مردم ایران در شبکه اجتماعی توئیتر در طول بحران کرونا پرداخته‌اند. برای این منظور یک مدل شبکه عصبی عمیق ارائه می‌شود. با توجه به اینکه داده‌های برچسب‌گذاری شده از توییت‌های مرتبط با کرونا در دسترس نیست، مدل پیشنهادی ابتدا روی مجموعه داده Sentiment140 دانشگاه استنفورد شامل یک میلیون و ششصد هزار توییت آموزش داده شد، سپس برای طبقه‌بندی احساسات موجود در توییت‌های جمع‌آوری شده مرتبط با کرونا در ایران استفاده می‌شود. نتایج نشان می‌دهد درصد توییت‌ها دارای احساسات منفی نسبت به توییت‌های مثبت به شکل معنی‌داری بیشتر است (بصیری و همکاران، ۱۴۰۰).

روش پژوهش

این پژوهش از لحاظ هدف در زمره پژوهش‌های کاربردی قرار می‌گیرد. جامعه مورد تحلیل این پژوهش فضای شبکه اجتماعی توئیتر است و نمونه مورد مطالعه نظرات کاربران در خصوص تکنولوژی واقعیت مجازی از سال ۲۰۱۷ تا سال ۲۰۲۲ می‌باشد که حدود ۱۲۹۰۵۴۰ سطر داده دریافت شد. در این پژوهش زبان برنامه‌نویسی پایتون بوده، در محیط ژوپیتر برنامه‌نویسی انجام شده است. در این پژوهش مطابق روش‌شناسی کریسپ گام برداشته شد؛ که مطابق مراحل زیر می‌باشد (Chapman et al, 2000):

شکل ۱. چرخه متدولوژی کریسپ



یافته‌ها

در این بخش به بررسی مراحل انجام شده می‌پردازیم.

۱. جمع‌آوری داده‌ها

اولین گام در این بخش استخراج داده‌ها و جمع‌آوری نظرات کاربران توئیتر درباره‌ی واقعیت مجازی با استفاده از برنامه‌نویسی پایتون می‌باشد، در ابتدا از طریق دسترسی به api توئیتر و با استفاده از پکیج Snsrape نظرات کاربران جمع‌آوری شد و تمام توئیتهای شامل #virtualreality، توسط خزنده وب^۱ جمع‌آوری شده‌اند، در مجموع ۱۲۹۰۵۴۰ سطر دیتا جمع‌آوری و ذخیره شد این مجموعه داده شامل ستون‌های زیر می‌باشد:

- تاریخ و زمان انتشار توئیت
 - شناسه توئیت (به هر حساب کاربری در توئیتر عدد منحصر به فردی تعلق می‌گیرد)
 - متن توئیت منتشر شده
 - نام کاربری
- با فراخوانی دستور df.head() می‌توانیم یک نمای اولیه از داده‌های مورد نظر را به صورت زیر مشاهده کنیم:

جدول ۱. پنج سطر اول داده‌های جمع‌آوری شده

نام کاربری	متن توئیت	شناسه توئیت	تاریخ زمان
۲۲۶sigP	#virtualreality #vr Virtual Trolls...	۸۷۹۸۵۱۴۷۰۹۸۷۹۵۶۲۲۵	۲۰۱۷،۰۶،۲۷ ۲۳:۵۸:۳۲
bjbig	#UnityEngineer #Game Programmer #virtualReality...	۸۷۹۸۵۰۷۲۷۰۶۰۹۸۳۸۰۸	۲۰۱۷،۰۶،۲۷ ۲۳:۵۵:۳۴
Official VRology	7 things to do with your DJI Goggles, if you d...	۸۷۹۸۵۰۱۰۸۴۹۷۶۵۳۷۶۰	۲۰۱۷،۰۶،۲۷ ۲۳:۵۳:۰۷
VRDomain	Another Super Awesome VR Domain you...	۸۷۹۸۴۹۳۷۴۶۹۱۵۹۰۱۴۴	۲۰۱۷،۰۶،۲۷ ۲۳:۵۰:۱۲
myleejoseph	Virtual Reality for #Libraries on a Shoestring...	۸۷۹۸۴۹۱۷۰۶۵۵۲۶۴۷۶۸	۲۰۱۷،۰۶،۲۷ ۲۳:۴۹:۲۳

1. Web Crawler

۲. آماده‌سازی داده‌ها

توئیت‌های جمع‌آوری‌شده مربوط به سال‌های ۲۰۱۷ تا ۲۰۲۲ هستند که شامل ۱۲۹۰۵۴۰ سطر دیتا است که با حذف توئیت‌های تکراری تعداد سطر دیتا به ۱۲۲۱۸۹۱ کاهش می‌یابد.

داده‌های متنی قبل از اعمال هر الگوریتم یادگیری ماشین، نیاز به آماده‌سازی دارد، آماده‌سازی داده‌ها با پیش‌پردازش انجام می‌شود. حال در مرحله‌ی پیش‌پردازش در اولین گام باید ستون تاریخ زمان را به دو ستون مجزا زمان و تاریخ تبدیل کنیم، زیرا روی تاریخ‌ها تحلیل و بررسی انجام دهیم.

در این گام به دلیل حجم زیاد داده و با توجه به اینکه کدهای پایتون را به صورت آنلاین روی گوگل کولب^۱ می‌نویسیم و اجرا می‌کنیم، با توجه به ارتباط اینترنتی اجرای کدها خیلی طولانی می‌شود و گاهی مشکلاتی از قبیل قطعی اینترنت و ریست شدن گوگل کولب وجود دارد به همین دلیل ما حجم داده‌ها را کم کردیم. برای کاهش حجم و زمان محاسبات با کمک دستور `df.sample(frac=0.5)`، ۵۰٪ از توئیت‌ها را به عنوان نمونه در نظر می‌گیریم و بر روی آن کار می‌کنیم. با این کار تعداد توئیت‌ها به ۶۱۰۹۴۶ کاهش می‌یابد.

حال به بررسی طول توئیت‌ها با استفاده از تابع `Len()` می‌پردازیم. طول توئیت‌ها از ۱۶ کاراکتر تا ۹۸۸ کاراکتر است. همچنین با کمک تابع `Split()` تعداد کلمات توئیت‌ها را به دست می‌آوریم که از یک کلمه تا ۲۱۲ کلمه را شامل می‌شوند، باید در نظر داشت توئیت‌هایی که طولشان خیلی کوتاه باشد معمولاً محتوای مفیدی ندارند یا برعکس توئیت‌هایی که خیلی طولانی باشند در بیشتر مواقع توئیت‌های تبلیغاتی یا غیر مرتبط هستند. برای اینکه دریابیم توئیت‌ها از چه تعداد کمتر و یا از چه تعداد بیشتر حذف کنیم تا توئیت‌های مفید را تحلیل کنیم، توئیت‌ها را بر اساس تعداد کلمات بررسی کردیم، اعداد مختلفی را امتحان کردیم و مشاهده کردیم توئیت‌هایی که بین ۱۰ تا ۸۰ کلمه هستند

توئیت‌های مفیدتری هستند، اغلب توئیت‌هایی با تعداد کلمات کمتر از ۱۰ شامل اطلاعات مفیدی برای مدل‌سازی موضوعی نیستند. همچنین توئیت‌هایی با تعداد کلمات بیشتر از ۸۰ معمولاً هرزنامه هستند؛ بنابراین این دسته توئیت‌ها از داده‌ها حذف می‌شوند تا توئیت‌های قابل استنادتری برای مدل‌سازی موضوعی داشته باشیم، با انجام این مرحله ۵۳۷۴۵۱ نمونه خواهیم داشت، (لازم به ذکر است این مرحله را برای کم کردن تعداد داده‌ها نیز انجام دادیم).

زمانی که تمام توئیت‌های شامل #virtualreality جمع‌آوری می‌کنیم، ممکن است توئیت‌هایی با زبان‌های دیگر را نیز جمع‌آوری کند. در این گام توسط کتابخانه Langdetect، توئیت‌هایی که زبانی غیر از انگلیسی داشتند را از داده‌ها حذف کردیم که در نتیجه ۵۰۰۷۷۵ نمونه با زبان انگلیسی باقی ماند.

در گام پاک‌سازی داده‌ها به حذف آدرس‌های اینترنتی، حذف کاراکترهای اضافه، حذف فاصله‌های اضافه، حذف ایست‌واژه‌ها، بن‌واژه‌سازی کلمات و غیره پرداختیم. بعد از پاک‌سازی داده‌ها برای اینکه بتوانیم مدل‌سازی انجام دهیم، باید داده‌ها را از حال متن به صورت یک لیست استخراج کنیم. تعداد کلمات موجود در هر توئیت مورد بررسی قرار گرفته است. توئیت‌ها از ۰ کلمه تا ۷۰ کلمه را شامل می‌شوند که در ویژگی تعداد کلمات بعد از پاک‌سازی ذخیره شده است.

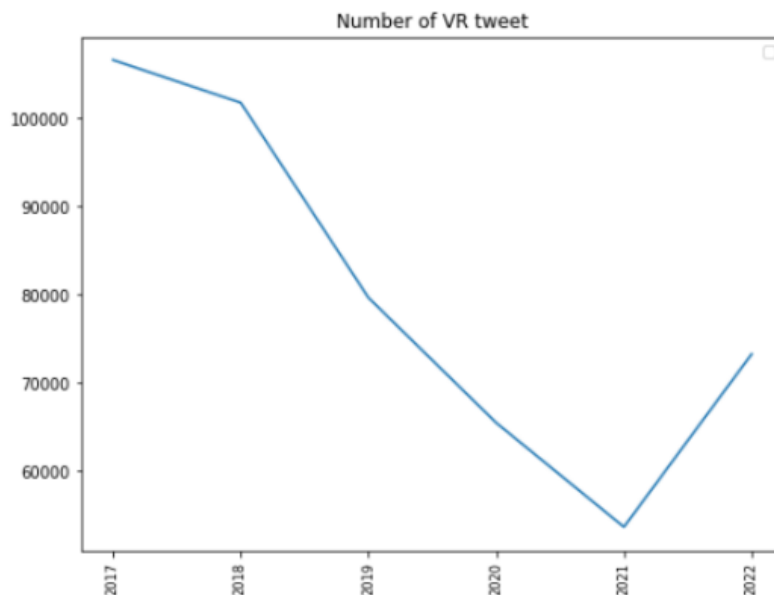
مجدداً با بررسی طول توئیت‌ها، توئیت‌هایی با کلمات مفید کمتر از ۵ را بررسی می‌کنیم (به صورت رندوم تعدادی از توئیت‌ها را مورد بررسی قرار گرفت) و دریافتیم توئیت‌هایی با کمتر از این تعداد شامل اطلاعات مفیدی برای مدل‌سازی موضوعی نیستند و به تشخیص موضوع توئیت به ما کمکی نمی‌کند. همین‌طور برای اینکه حجم داده‌هایمان نیز کمتر شود این فیلتر را اعمال می‌کنیم، پس توئیت‌هایی با کلمات مفید کمتر از ۵ را از داده‌ها حذف می‌کنیم تا با حداقل ۵ کلمه بتوانیم موضوع توئیت را تشخیص دهیم. با این کار ۴۸۰۴۳۲ نمونه در داده‌ها باقی می‌ماند.

جدول ۲. داده‌های جمع‌آوری شده بعد از پاک‌سازی داده‌ها

شناسه توثیت	متن	نام کاربری	تاریخ	زمان	طول توثیت ها	تعداد کلمات	زبان	متن پاک‌سازی شده	تعداد کلمات بعد از پاک‌سازی
۱۴۸۶۲۸۷ ۹۰۹۷۲۸۷ ۸۸۴۸۱	#VR Immersive Tech announces ...	PDH_ Meta verse	۲۰۲۲,۰۱,۲۶	۱۰:۴۰:۲۶	۲۶۳	۲۹	انگلیسی	['Immersive', , 'tech', , 'announce', , ...	۱۸
۱۵۷۰۸۰۷ ۷۱۲۴۵۱۲ ۶۸۶۰۸	One of the best drawings of all time...	Bmano 931	۲۰۲۲,۰۹,۱۶	۱۶:۱۱:۵۷	۱۵۰	۱۲	انگلیسی	['one', 'good', , 'drawing', , 'time', ...	۸
۸۹۴۶۸۳۳ ۱۲۵۵۱۳۳ ۷۹۸۴	Just had my first #Virtual Reality expreience...	G_ Cero Medo	۲۰۱۷,۰۸,۰۷	۲۲:۱۴:۵۸	۱۰۰	۱۴	انگلیسی	['first', , 'experience', , 'woooooow', , ...	۸
۸۸۱۷۸۰۸ ۰۶۵۲۳۵۱ ۸۹۷۷	What #VR and #AR could mean for your #brand...	Kt- verran	۲۰۱۷,۰۷,۰۳	۰۷:۴۵:۰۱	۱۵۸	۱۷	انگلیسی	['ar', 'could', , 'mean', , 'brand', , 'amp', ..	۷
۸۲۴۶۵۹۳ ۰۶۹۰۹۹۹ ۵۰۱۰	#virtual reality UploadVR: The fast-paced ...	Kageon Plays	۲۰۱۷,۰۱,۲۶	۱۶:۴۴:۳۴	۱۳۹	۱۷	انگلیسی	['uploadvr', , 'fast', , 'pace', , 'moto', ..	۱۰

پس از پاک‌سازی داده‌ها، تعداد توثیت‌های شامل واقعیت مجازی را در سال‌های ۲۰۱۷ تا ۲۰۲۲ را رسم می‌کنیم تا با بررسی روند توثیت‌ها در این سال‌ها بتوانیم به نتایجی برسیم:

شکل ۲. تعداد توئیتهای شامل واقعیت مجازی در سالهای ۲۰۱۷ تا ۲۰۲۲



این نمودار به دست آمده را می توان با کمک چرخه هایپ گارتنر تفسیر نمود، این چرخه همان طور که در ادبیات پژوهش گفته شد، یک نمای کلی از چگونگی رشد و تکامل تکنولوژی، در یک بازه زمانی مشخص ارائه می دهد، با توجه به اینکه ما از سال ۲۰۱۷ این فناوری را مورد بررسی قرار دادیم، سال های ظهور فناوری پیشرفت نمودار را نمی بینیم، بدین صورت که در سال های اولیه، فناوری واقعیت مجازی با شیب زیادی در حال پیشرفت بوده و اشتیاق ها نسبت به این فناوری در سال ۲۰۱۷ به اوج خود رسیده است پس از اینکه اوج انتظارات برآورده شد، از سال ۲۰۱۷ به بعد نمودار سیر نزولی طی کرده، دلایلی مختلفی وجود دارد ممکن است، این فناوری انتظارات کاربران را برآورده نکرده و رسانه ها و مطبوعات فناوری را کنار گذاشتند و یا سازندگان فناوری را فراموش کردند.

در سال ۲۰۲۱ به بعد مشاهده می کنیم که سازندگان تمرکز ویژه ای روی فناوری واقعیت مجازی داشتند و با موفقیت روبه رو شده اند در نتیجه مردم به سودمندی فناوری و فواید آن پی بردند. در سال ۲۰۲۲ مشاهده می کنیم که این فناوری با شیب نسبتاً زیادی در

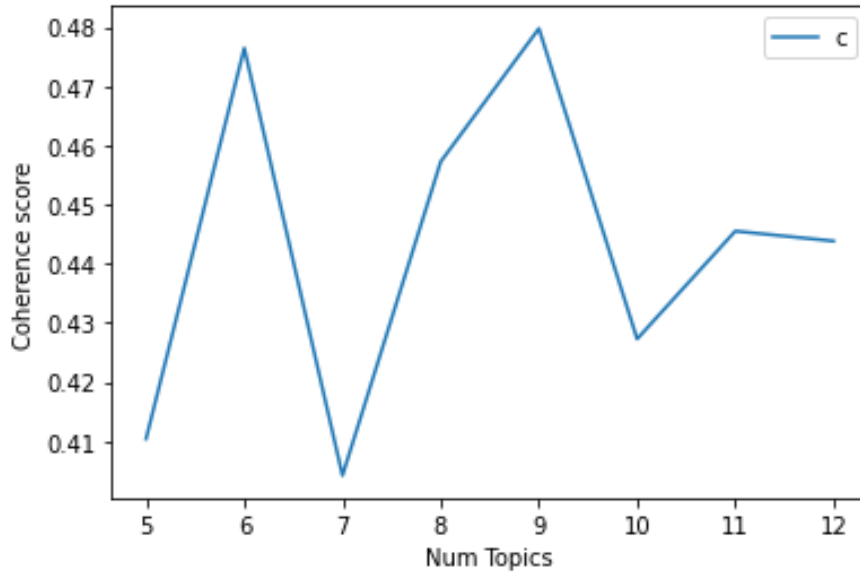
حال افزایش می‌باشد به این معنا که مزایای اصلی و کاربرد عملی این فناوری مورد استقبال گرفته و در حال سودهی می‌باشد.

در مرحله بعد برای مدل‌سازی به تشخیص کلمات تکراری در توئیت‌ها و تشخیص تعداد تکرار کلمه در هر توئیت از هر نیاز داریم، پس از طریق کتابخانه ast آن‌ها را ایجاد می‌کنیم.

۳. مدل‌سازی و ارزیابی مدل

در این مرحله توسط کتابخانه Gensim به مدل‌سازی موضوعی تخصیص پنهان دیریکله می‌پردازیم. این مدل‌سازی برای پردازش زبان طبیعی، متن‌کاوی و تحلیل رسانه‌های اجتماعی، بازیابی اطلاعات به کار گرفته شده است. مدل‌سازی موضوعی، بر اساس تجزیه و تحلیل رسانه‌های اجتماعی؛ درک واکنش‌ها و مکالمات بین افراد در جوامع آنلاین را تسهیل می‌کند، همچنین الگوهای مفید و قابل‌درک را از تعاملات آن‌ها استخراج می‌کند (Jelodar et al., 2019). برای ارزیابی بحث مدل‌سازی موضوعی یک تابعی به نام انسجام وجود دارد که نشان می‌دهد توئیت‌هایی که داخل یک موضوع شناسایی شده چقدر کلماتشان به هم نزدیک می‌باشد، به معنای دیگر درجه تشابه معنایی بین کلمات و تمایز بین موضوعات را می‌سنجد، هرچقدر مقدار انسجام ما بالاتر باشد مدل‌سازی موضوعی ما دقیق‌تر خواهد بود و با توجه به اینکه تعداد موضوعاتی که در توئیت‌های ما وجود دارد را نمی‌دانیم و باید آن‌ها به دست آوریم، از کمترین تعداد موضوعات شروع به امتحان کردن می‌کنیم و مدل‌سازی توسط مدل تخصیص پنهان دیریکله انجام می‌دهیم، در آخر تعداد موضوعی که بیشترین انسجام را به ما بدهد را انتخاب می‌کنیم. در نتیجه این بررسی، نمودار زیر به دست آمد.

شکل ۳. نمودار بررسی مقدار انسجام موضوع



که در آن مقدار دقیق انسجام هر موضوع به صورت زیر می باشد:

جدول ۳. مقدار انسجام ۹ موضوع به دست آمده

مقدار انسجام	شماره موضوع
۰/۴۱۰۴	۵
۰/۴۷۶۴	۶
۰/۴۰۴۲	۷
۰/۴۵۷۲	۸
۰/۴۷۹۷	۹
۰/۴۲۷۳	۱۰
۰/۴۴۵۵	۱۱
۰/۴۴۳۸	۱۲

همان طور که مشاهده می شود، تعداد موضوعات ۹ دارای انسجام بالاتری می باشد، پس بار دیگر مدل سازی تخصیص پنهان دیریکله را روی ۹ موضوع انجام دهیم، با این کار

توئیت‌ها را در ۹ تا موضوع مختلف دسته‌بندی می‌کند. برای ارزیابی مدل نیز با توجه به اینکه یک توزیع احتمالی داریم، از معیار سرگشتگی استفاده کردیم و برای درجه تشابه معنایی بین کلمات و تمایز بین موضوعات نیز از امتیاز انسجام استفاده کردیم که نتیجه ارزیابی مدل‌سازی به صورت زیر است:

جدول ۴. مقدار معیار سرگشتگی و امتیاز انسجام به دست آمده با استفاده از کد ارزیابی

معیار سرگشتگی	امتیاز انسجام
-۹/۴۴۱۷۰۹۴۱۸۹۹۶۶۵۶	۰/۴۷۹۷۴۰۲۵۰۳۸۵۵۰۸۱

هرچقدر معیار سرگشتگی پایین‌تر و امتیاز انسجام بالاتر مدل ما دقت بالاتری دارد.

۴. نام‌گذاری موضوعات

کاری که مدل‌سازی موضوعی انجام می‌دهد این است که با استفاده از بررسی توزیع کلمات در توئیت‌ها، سعی می‌کند با استفاده از تشخیص توزیع کلمات، موضوع‌های مختلف را از هم جدا کند، توئیت‌هایی که توزیع کلماتشان به هم شبیه است داخل یک موضوع قرار می‌گیرند. در ابتدا تعداد موضوعات مختلف را توسط مدل‌سازی تخصیص پنهان دیریکله بررسی کردیم، دیدیم با ۹ موضوع نتیجه بهتری می‌گیریم، حال توسط مدل‌سازی تخصیص پنهان دیریکله توزیع کلمات در ۹ دسته انجام می‌شود، مدل‌سازی تخصیص پنهان دیریکله وزن کلمات کلیدی را به ما برمی‌گرداند، حال با در نظر داشتن کلمات کلیدی و بررسی حداقل ۵ توئیت مختلف از هر موضوع باید بررسی کنیم که چه عنوان مناسب با کلمات کلیدی و توئیت‌ها می‌توانیم انتخاب کنیم. وزن کلمات کلیدی به صورت زیر می‌باشد:

شکل ۴. نمایشی از وزن کلمات کلیدی ۹ موضوع به دست آمده

```
[
(0,
'0.017*experience' + 0.014*video' + 0.012*new' + 0.011*world' + '
'0.010*look' + 0.010*make' + 0.009*like' + 0.008*take' + 0.008*use' + '
'0.008*real' + 0.007*see' + 0.006*immersive' + 0.006*one' + '
'0.006*first' + 0.006*feel'),
(1,
'0.014*amp' + 0.012*pm' + 0.010*friend' + 0.009*we' + 0.009*game' + '
'0.009*secondlife' + 0.008*join' + 0.008*race' + 0.007*news' + '
'0.007*live' + 0.007*life' + 0.007*shape' + 0.007*share' + 0.006*link' + '
'+ 0.006*rt'),
(2,
'0.095*ar' + 0.037*technology' + 0.033*ai' + 0.024*xr' + 0.023*tech' + '
'0.020*mixedreality' + 0.019*iot' + 0.017*future' + 0.016*innovation' + '
'0.015*rt' + 0.014*mr' + 0.014*augment' + 0.012*business' + '
'0.011*startup' + 0.011*marketing'),
(3,]
'0.030*education' + 0.029*edtech' + 0.020*free' + 0.017*make' + '
'0.015*download' + 0.013*arvrinedu' + 0.013*art' + 0.013*click' + '
'0.011*ar' + 0.011*museum' + 0.011*trend' + 0.011*history' + '
'0.010*ebook' + 0.009*top' + 0.008*giveaway'),
(4,
'0.044*game' + 0.038*oculus' + 0.017*quest' + 0.012*new' + 0.012*play' + '
'+ 0.012*oculusquest' + 0.011*gaming' + 0.011*vrgame' + '
'0.011*oculusrift' + 0.010*go' + 0.008*video' + 0.008*check' + '
'0.008*youtube' + 0.008*uploadvr' + 0.008*rift'),
(5,
'0.020*use' + 0.016*amp' + 0.014*experience' + 0.012*help' + 0.012*new' + '
'+ 0.012*learn' + 0.010*training' + 0.010*we' + 0.010*technology' + '
'0.008*great' + 0.007*way' + 0.007*see' + 0.007*work' + 0.006*student' + '
'+ 0.006*immersive'),
(6,
'0.031*headset' + 0.022*google' + 0.014*ar' + 0.011*enter' + '
'0.010*samsung' + 0.009*price' + 0.008*video' + 0.008*announce' + '
'0.008*apple' + 0.008*wearable' + 0.008*mixedreality' + 0.007*win' + '
'0.007*vision' + 0.007*mobile' + 0.006*gear'),
(7,
'0.053*metaverse' + 0.034*gt' + 0.016*nft' + 0.014*game' + 0.012*meta' + '
'+ 0.011*facebook' + 0.010*web' + 0.010*new' + 0.010*ps' + '
'0.009*playstation' + 0.008*headset' + 0.008*blockchain' + 0.007*market' + '
'+ 0.007*launch' + 0.007*news'),
(8,
'0.075*htcvive' + 0.050*vive' + 0.036*gamedev' + 0.031*indiedev' + '
'0.029*unrealengine' + 0.029*steam' + 0.028*pc' + 0.026*indiegamae' + '
'0.026*gamer' + 0.024*oculus' + 0.022*sale' + 0.019*available' + '
'0.018*indiegamedev' + 0.018*ue' + 0.017*pcgame')]
```

یکی دیگر از خروجی‌هایی که می‌توانیم بر اساس آن کلمات کلیدی را بررسی کنیم، استفاده از تحلیل مؤلفه اساسی^۱ می‌باشد با توجه به اینکه مجموعه موضوعات ما دارای متغیرهای زیادی (کلمات کلیدی) می‌باشد، برای اینکه درک بهتری از کلمات کلیدی و

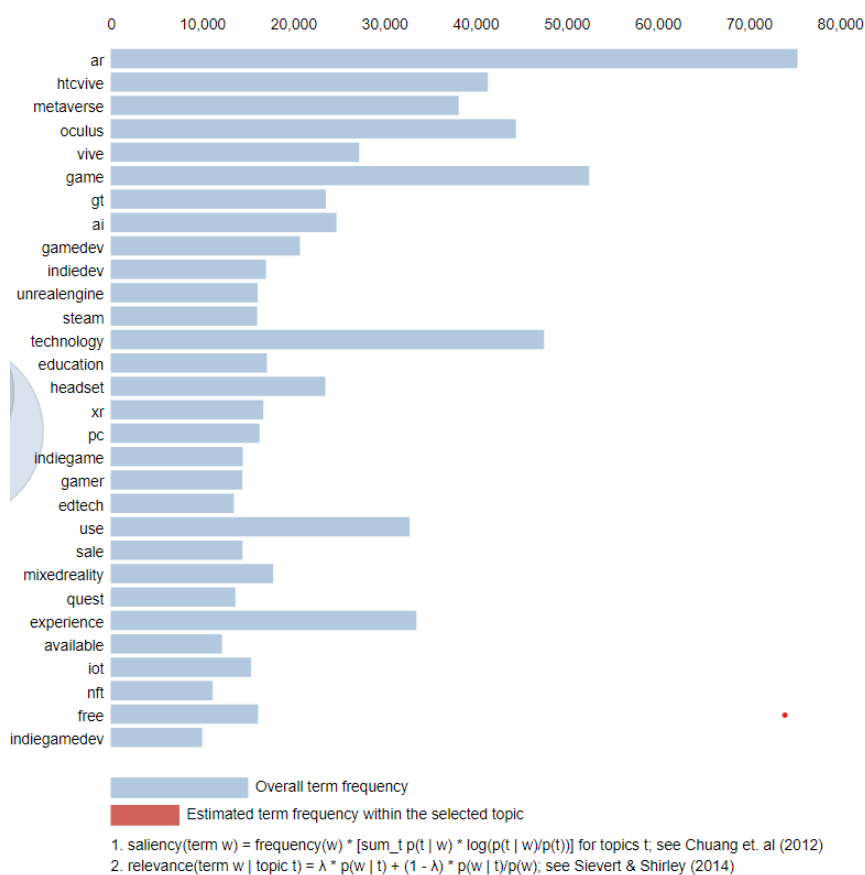
1. Principal Component Analysis (PCA)

فراوانی آن‌ها داشته باشیم؛ می‌توانیم از تحلیل مؤلفه‌های اساسی جهت تفسیرپذیری و درک بهتر با حفظ حداکثر مقدار اطلاعات استفاده کنیم، در اینجا موضوعات در فضای دوبعدی رسم شده‌اند و کلمات کلیدی بر اساس فراوانی‌شان نمایش داده می‌شوند. از کتابخانه PyLDAvis برای مصورسازی خروجی استفاده کردیم، در زیر نمایش داده شده است:

شکل ۵. نمای کلی مدل‌سازی موضوع



شکل ۶. نمای کلی مدل‌سازی موضوع (کلمات کلیدی)

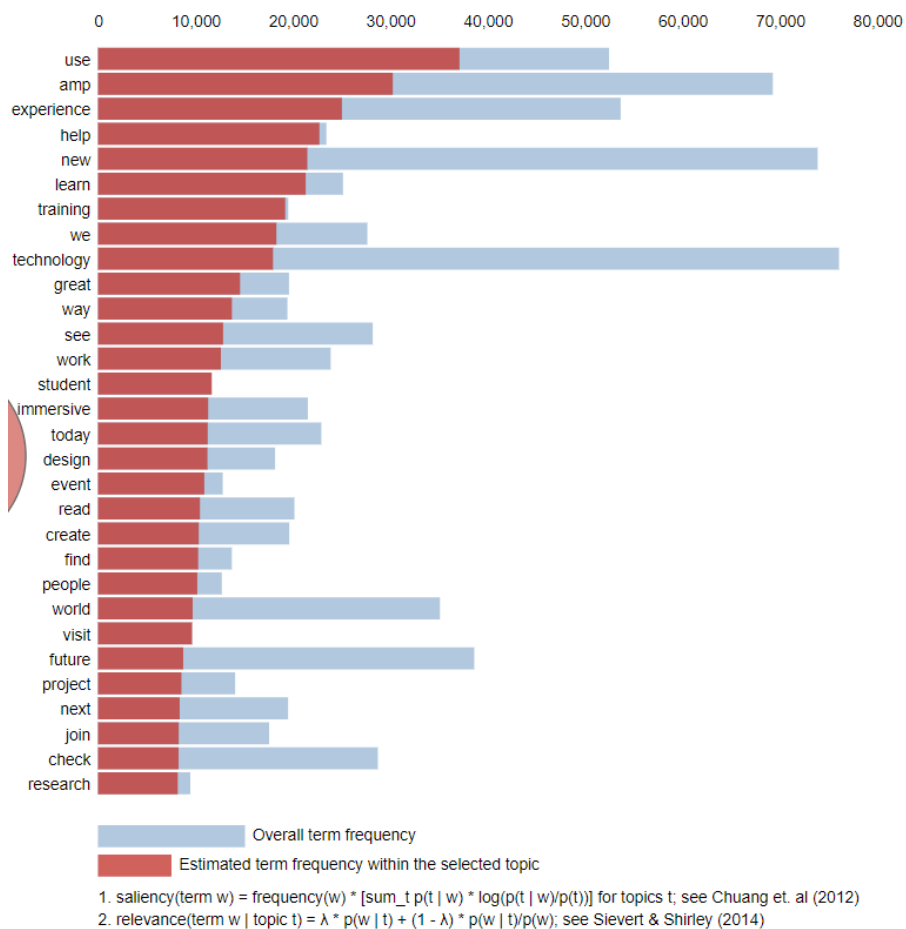


در شکل ۵، هر کدام از دایره‌ها نمایانگر یک موضوع است با عددی مشخص شده است و شکل ۶، کلمات کلیدی با بیشترین تکرار. اندازه هر دایره نشان از تکرار بیشتر و نزدیکی دایره‌ها به هم، نشان‌دهنده نزدیکی موضوعات به هم می‌باشد.

با انتخاب هر موضوع، کلمات کلیدی آن موضوع نمایش داده می‌شوند که رنگ قرمز تعداد تکرار در این موضوع و رنگ آبی تعداد تکرار در کل متون را نشان می‌دهد، از تحلیل این نمودار می‌توانیم برای نام‌گذاری موضوعات استفاده کنیم به این شکل کلمات کلیدی که زیاد تکرار شده‌اند با احتمال زیاد می‌توانند در تشخیص موضوع به ما کمک

کنند، در زیر به مصورسازی خروجی در فضای دوبعدی به ۹ موضوع به دست آمده پرداختیم:

شکل ۷. موضوع پرتکرار اول



اگر موضوع پرتکرار اول را بررسی کنیم، یکسری کلمات می‌بینیم که بر اساس میزان فراوانی‌شان با رنگ قرمز نشان داده شده است: «استفاده کردن»، «صفحات موبایلی

شتاب یافته^۱، «تجربه»، «کمک کردن»، «جدید»، «یادگیری»، «فناوری» و غیره مثلاً کلمه‌ی «استفاده کردن» را در نظر می‌گیریم.

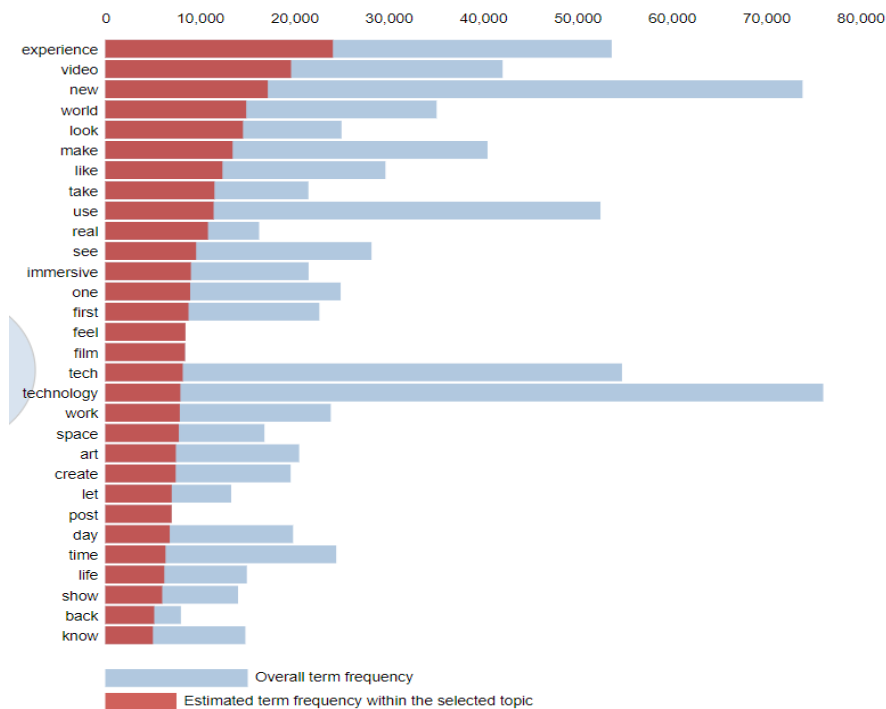
نمونه توئیت:

Topic 5.0: Virtual reality may help improve the quality of life for people with dementia

topic 5.0: #VR can improve quality of life for people with dementia

که با بررسی این کلمات و توئیت‌ها می‌توانیم تحلیل کنیم که موضوع این توئیت در راستای بهبود بخشیدن و ارتقا در زمینه فناوری‌های مختلف توسط واقعیت مجازی صحبت می‌کند که طبق نظر پژوهشگر این موضوع «پیشرفت» نام‌گذاری شده است.

شکل ۸. موضوع دوم



1. saliency(term w) = frequency(w) * [sum_t p(t | w) * log(p(t | w)/p(t))] for topics t; see Chuang et. al (2012)
 2. relevance(term w | topic t) = λ * p(w | t) + (1 - λ) * p(w | t)/p(w); see Sievert & Shirley (2014)

1. Accelerated Mobile Pages (AMP)

همین‌طور اگر موضوع ۲ را بررسی کنیم کلمات «تجربه»، «ویدئو»، «جدید»، «جهان»، «ساختن»، «واقعی»، «فیلم»، «دیدن»، «غوطه‌ور» و غیره، از کلمات پر استفاده در این موضوع هستند.

نمونه توئیت:

Topic 0.0: Disney files patent for 3D hologram immersive ride experience
#VR #VirtualReality
Topic 0.0 Creators can Touch Their CAD Designs With exiii Inc's EXOS Haptic System

با کمک این کلمات و بررسی تعدادی توئیت می‌توان نتیجه گرفت توئیت‌های دسته‌بندی‌شده در این موضوع حول کسب تجربه‌های جدید از فناوری‌های نوظهور بوده است که می‌توانیم این موضوع را «فناوری جدید» نام‌گذاری کنیم. اگر موضوع ۳ را بررسی کنیم درمی‌یابیم با توجه به کلمات «بازی»، «اوکلوس»، «جستجو»، «جدید»، «بازی»، «بازی‌های واقعیت مجازی»، «عینک واقعیت مجازی اوکلوس» «یوتیوب» و غیره از کلمات پرتکرار این موضوع هستند.

نمونه توئیت:

Topic 4.0: Great to see youtubers enjoy your new VIVE Game by MalboMX
Topic 4.0: Operation Warcade VR Review – Nostalgic Gaming at its Best

که با توجه به کلمات به‌دست‌آمده و بررسی توئیت‌های دسته‌بندی‌شده می‌توان نتیجه گرفت که این موضوع درباره‌ی بازی‌های جدید و مجازی می‌باشد که می‌توان این موضوع را «بازی‌های مجازی» نام‌گذاری کرد.

اگر موضوع ۴ را بررسی کنیم کلمات «متاورس»، «توکن غیرقابل تعویض^۱»، «بازی»، «متا»، «فیس‌بوک»، «وب»، «بلاک‌چین»، «پلی‌استیشن»، «دنیای مجازی»، «کریپتو»، «خریدن» و غیره وجود دارد.

1. Non-Fungible Token (NFT)

نمونه توثیت:

Topic 7.0: The world's first metaverse ATM was just launched

Topic 7.0: The metaverse & web 3.0 revolution is the future of blockchain or cryptocurrency.

که با بررسی کلمات و توثیت‌های دسته‌بندی‌شده در این موضوع متوجه می‌شویم این موضوع در رابطه با رمز ارز و تعامل در دنیای مجازی است که می‌توان آن را «متاورس» نام‌گذاری کرد.

اگر موضوع ۵ را بررسی کنیم کلمات «واقعیت افزوده»، «اینترنت اشیا»، «هوش مصنوعی»، «واقعیت گسترده^۱»، «فناوری»، «خلاقیات»، «آینده»، «واقعیت ترکیبی^۲»، «دیجیتال»، «استارت‌آپ» و غیره وجود دارد.

نمونه توثیت:

Topic 2.0: Round Sunglasses #AR Snapchat lens is now available on Catchar

Topic 2.0: Crane Craze #AR App by @magic leap is now available on Catchar

با توجه به این کلمات و توثیت‌های دسته‌بندی‌شده درمی‌یابیم این موضوع در رابطه با کسب‌وکارهایی است در راستای فناوری‌های نوظهور فعالیت می‌کنند، به همین دلیل این موضوع را «کسب‌وکار مبتنی بر فناوری» نام‌گذاری کردیم.

موضوع ۶ کلمات «بازی»، «هدست واقعیت مجازی HTCvive»، «توسعه بازی‌ها»، «بازی‌های مستقل»، «Unreal Engine^۳»، «کامپیوتر شخصی»، «گیمر»، «اوکلوس»، «خریدن»، «هلی‌کوپتر» و غیره را شامل می‌شود.

نمونه توثیت:

Topic 8.0: And so it begins...who's ready for #UniteBerlin?!

Topic 8.0: Helicopter VR | Get it now.

-
1. Extended Reality
 2. Mixed Reality

۳. یک موتور بازی گرافیک رایانه‌ای سه‌بعدی است.

با توجه به این کلمات و توئیت‌های دسته‌بندی‌شده درمی‌یابیم این موضوع در رابطه با بازی‌های کامپیوتری است که نوآوری داشتند و توسعه پیدا کردند، طبق نظر پژوهشگر این موضوع را «بازی‌های مستقل» نام‌گذاری شده است.

در موضوع ۷ کلمات «آموزش»، «فناوری‌های آموزشی»، «رایگان»، «دانلود»، «کلیک»، «واقعیت افزوده»، «تمایل»، «کتاب الکترونیک» و غیره به چشم می‌خورد.

نمونه توئیت:

Topic 3.0: Virtual reality #eBook Will Make You The Authority

Topic 3.0: Designing for virtual reality and the impact on education

که با توجه به این کلمات و بررسی تعدادی توئیت درمی‌یابیم توئیت‌های دسته‌بندی‌شده در این موضوع در رابطه با فناوری جدید در راستای آموزش بوده است که این موضوع را «آموزش» نام‌گذاری کردیم.

در موضوع ۸ «صفحات موبایلی شتاب‌یافته»، «دوست»، «پیوند»، «بازی» «بازی زندگی دوم»، «پیوستن»، «شکل دادن» و غیره به چشم می‌خورد که برای تشخیص دقیق‌تر و بهتر باید توئیت‌های این موضوع را نیز بررسی کنیم:

نمونه توئیت:

Topic 1.0: doing what we do best! Today we are filming a #virtualreality aka #360degree tour.

Topic 1.0 getting ready to burn a COVID at 20:20 o'clock (8:20 pm SLT PST) Come join us.

با بررسی کلمات کلیدی درمی‌یابیم این موضوع در ارتباط با دنیایی است که در آن افراد به هم می‌پیوندند به شکل دادن و ساختن اطراف خود می‌پردازند، این موضوع را می‌توان «خلق کردن و ساختن» نام‌گذاری کرد.

در موضوع ۹ کلمات «هدست»، «گوگل»، «واقعیت افزوده»، «گوشی‌های هوشمند»، «موبایل»، «سامسونگ»، «اعلام کردن»، «اپل»، «قیمت»، «ویدئو» و غیره به چشم می‌خورد.

نمونه توثیت:

Topic 6.0: APPLE Mac mini (2018) goes official with 8th-gen Intel Core processor

Topic 6.0: Will this device be successful? Are we ready to adopt the technology?

که با بررسی این کلمات و تعدادی توثیت، درمی‌یابیم با توجه به این که این موضوع در رابطه با وسایل الکترونیکی برای کاربردهای خاصی طراحی شده‌اند و به آسان‌تر شدن زندگی کمک می‌کنند و همین‌طور کسب‌وکارهایی که این وسایل را تولید می‌کنند، می‌باشد پس این موضوع را می‌توان «گجت» نام‌گذاری کرد.

با توجه به وزن کلمات مشخص شده برای هر موضوع و بررسی تعدادی از توثیت‌ها، عنوان پیشنهادی برای موضوع‌های مختلف به صورت زیر مشخص شده است:

جدول ۵. نام‌گذاری ۹ موضوع مرتبط با فناوری واقعیت مجازی

شماره موضوع	نام‌گذاری شده
۰	فناوری جدید
۱	خلق و ساخت
۲	کسب‌وکار مبتنی بر فناوری
۳	آموزش
۴	بازی‌های مجازی
۵	پیشرفت
۶	گجت
۷	متاورس
۸	بازی‌های مستقل

حال باید با استفاده از مدل‌سازی موضوعی تخصیص پنهان دیریکله که ساختیم تک‌تک توثیت‌ها را بررسی کنیم و ببینیم جز کدام موضوع هستند با استفاده از این مدل، توزیع توثیت‌ها در موضوع‌های مختلف به صورت زیر است:

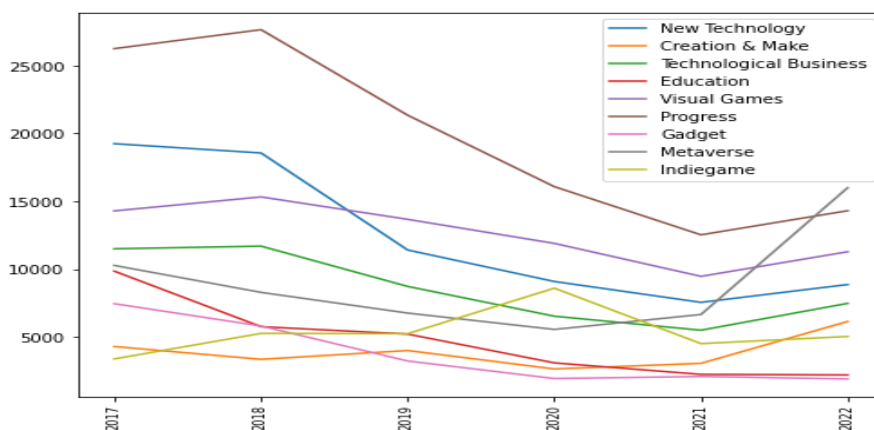
جدول ۶. تعداد توئیت‌ها در موضوع‌های مختلف

شماره موضوع	نام‌گذاری شده	تعداد
۰	فناوری جدید	۷۴۷۴۹
۱	خلق و ساخت	۲۳۵۴۸
۲	کسب‌وکار مبتنی بر فناوری	۵۱۴۶۷
۳	آموزش	۲۸۴۰۶
۴	بازی‌های مجازی	۷۵۹۵۹
۵	پیشرفت	۱۱۸۱۶۴
۶	گجت	۲۲۴۷۷
۷	متاورس	۵۳۵۷۲
۸	بازی‌های مستقل	۳۲۰۹۰

۵. تحلیل مدل‌سازی موضوعی

حال خروجی‌هایی که از مدل‌سازی موضوعی به دست آوردیم با کمک چند نمودار تحلیل می‌کنیم، ابتدا مدل تخصیص پنهان دیریکله و داده‌هایمان را فراخوانی می‌کنیم، سپس یکسری ویژگی‌های جدید سال، تاریخ و ماه ایجاد کردیم تا بر اساس موضوعات در سال‌های مختلف و تعداد توئیت‌ها در این سال‌ها تا بتوانیم تجسم بهتری از این مدل‌سازی داشته باشیم. با بررسی روند موضوع‌های مختلف در طول زمان، به شکل زیر می‌رسیم:

شکل ۹. بررسی روند موضوعات مرتبط با فناوری واقعیت مجازی در بازه‌ی ۲۰۱۷ تا ۲۰۲۲



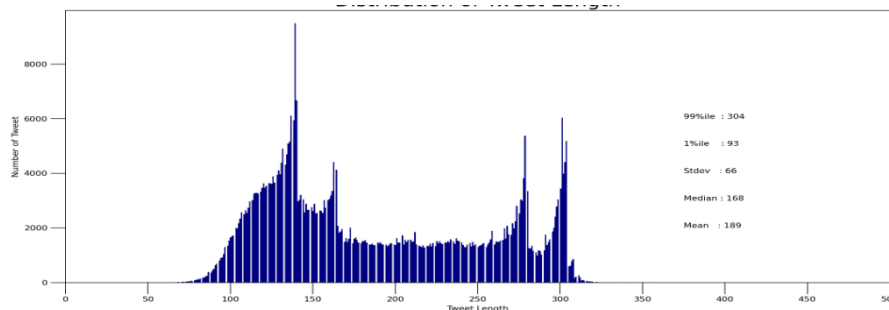
با توجه به شکل ۹، «پیشرفت» داغ‌ترین موضوع در طول زمان بوده به جز اواخر سال ۲۰۲۱ که «متاورس» از آن پیشی گرفته است که نشان‌دهنده اهمیت و توجه زیاد کاربران به این موضوع در سال‌های اخیر بوده است.

همین‌طور اگر نمودار «بازی‌های مجازی» و «کسب‌وکار مبتنی بر فناوری» را مشاهده کنیم می‌بینیم که تقریباً با یک نسبت در سال‌های مختلف تغییر می‌کند به این معنا که این نمودار باهم ارتباط دارند، درواقع باید کسب‌وکارها در نظر داشته باشند که عوامل تأثیرگذار روی این دو موضوع یکی است ولی موضوع «بازی‌های مجازی» طرفداران بیشتری دارد و اگر کسب‌وکار مبتنی بر فناوری به‌طور خاص روی بازی‌های مجازی تمرکز کنند به دلیل توجه بیشتر کاربران، رشد بیشتری خواهند داشت.

همین‌طور «بازی‌های مستقل» که یکسری تغییرات داشته ولی در سال‌های اخیر یک روند کاهشی و سپس بدون تغییر داشته است، حال پدیدآوردگان این بازی‌های باید بررسی کنند و در کل «بازی‌های مجازی» موضوع موردتوجه بیشتری نسبت به «بازی‌های مستقل» بوده است.

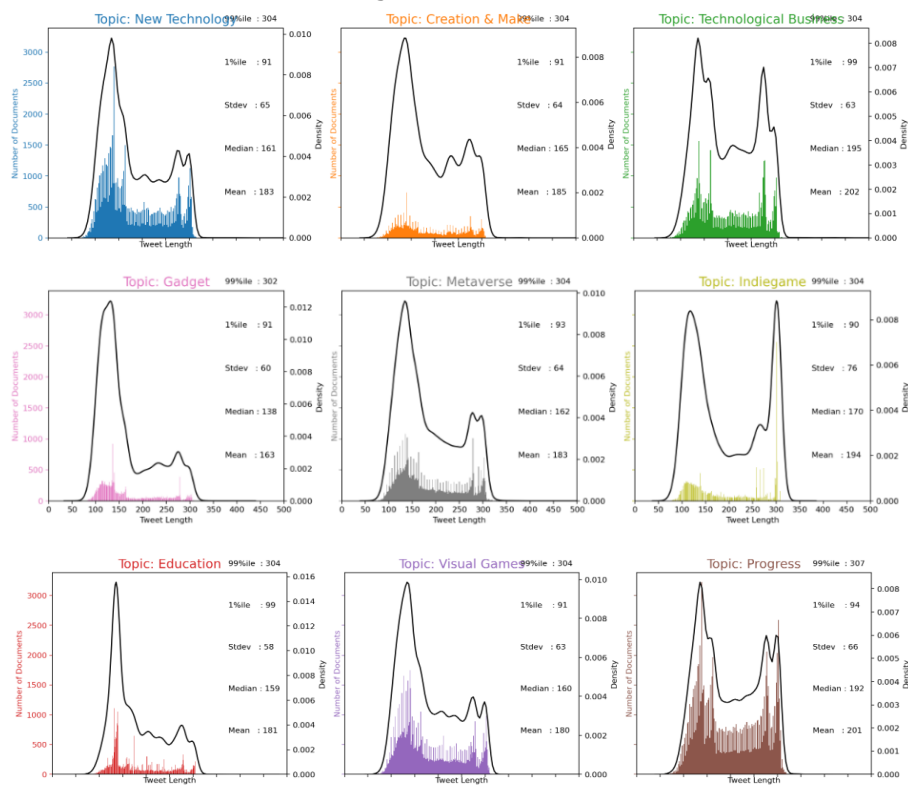
در موضوعات «آموزش» و «گجت» از ابتدای سال ۲۰۱۷ دارای سیر نزولی بوده است که نشان می‌دهد کاربران در طول زمان توجه خود به این موضوعات در حوزه واقعیت مجازی را از دست دادند، درواقع به موضوعات دیگری توجه خود را نشان دادند. حال در این بخش طول توئیت‌ها را بر اساس تعداد کلمات موردبررسی قرار دهیم، به‌صورت میانگین هر توئیت ۱۸۹ کاراکتر طول دارند.

شکل ۱۰. نمودار تعداد کل کلمات توئیت‌ها بر اساس طول توئیت‌ها



اگر طول توئیتهای را بر روی موضوعهای مختلف بررسی کنیم، به نمودار زیر می‌رسیم، همان‌طور که می‌بینیم نمودار «پیشرفت» با ۱۱۸۱۶۴ توئیته (جدول ۶) بیشترین تعداد توئیته را دارا می‌باشد و موضوع «خلق کردن و ساختن» کمترین تعداد توئیته را دارا است.

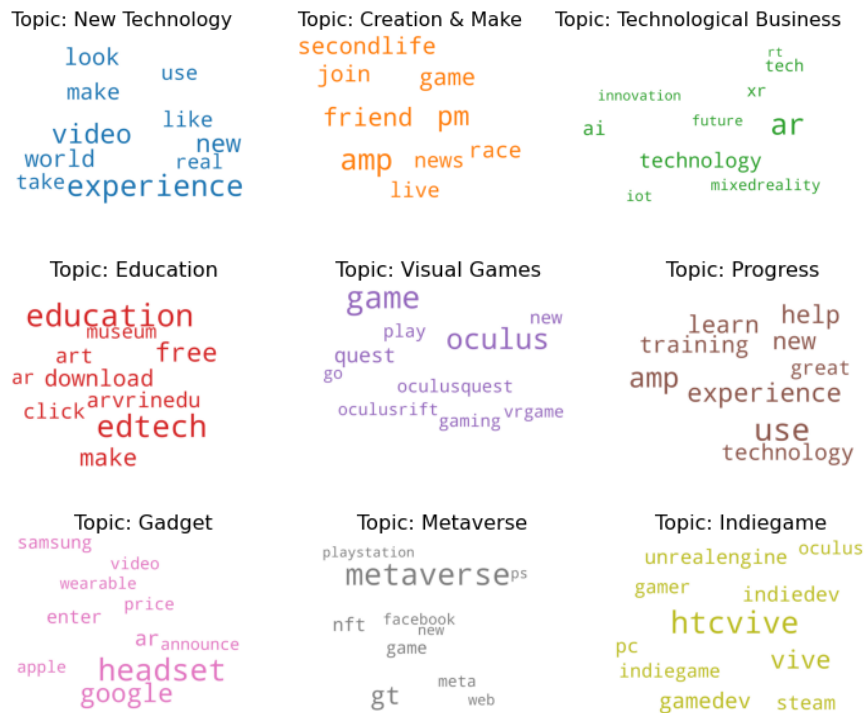
شکل ۱۱. برآورد چگالی ۹ موضوع مرتبط به دست‌آمده



خروجی دیگری که می‌توانیم رسم کنیم، نمایش کلمات کلیدی در قالب ابر کلمات^۱ برای موضوعهای مختلف می‌باشد، هرچقدر کلمات کلیدی ما پرتکرارتر باشد، کلمه بزرگ‌تر نشان داده می‌شود، به‌عنوان مثال در موضوع «فناوری‌های جدید» همان‌طور که قبلاً بررسی کردیم، بیشترین کلمه تکرار شده «تجربه» بوده است که مشاهده می‌کنیم این کلمه بزرگ‌تر از همه نشان داده شده است.

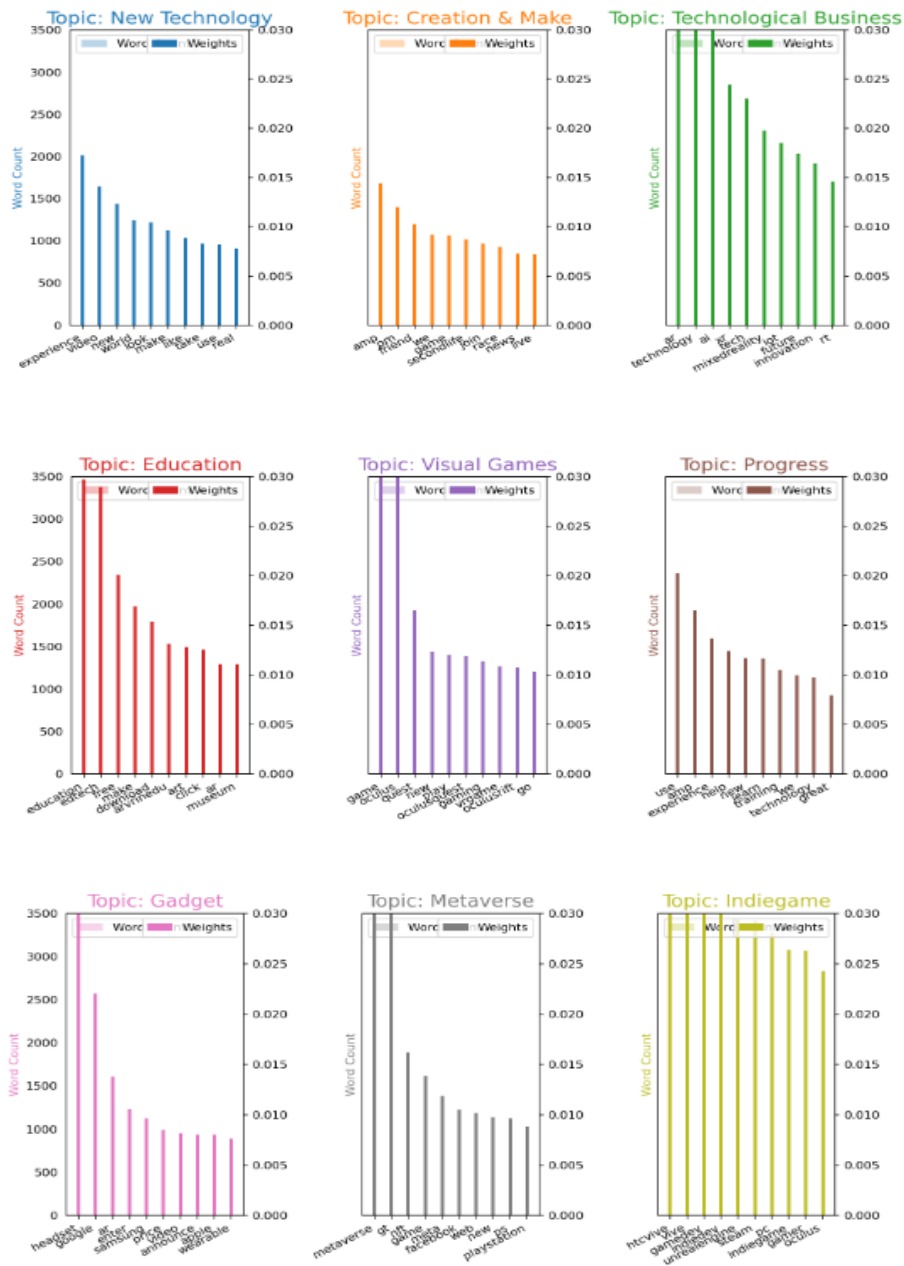
1. Word Clouds

شکل ۱۲. نمایی از ابرکلمات ۹ موضوع مرتبط به دست آمده



در تحلیلی دیگر تعداد تکرار کلمات کلیدی را در هر کدام از موضوعات بررسی کردیم (شکل 13) که این بررسی کلمات کلیدی در هر موضوع نمودارهای زیر را ایجاد می‌کند که در این نمودار هم کلمات «تجربه»، «صفحات موبایلی شتاب یافته»، «کسب و کار مبتنی بر فناوری»، «واقعیت افزوده»، «بازی»، «آموزش»، «استفاده کردن»، «هدست»، «متاورس» و «هدست واقعیت مجازی HTCvive» بیشترین کلمات پرتکرار بوده است.

شکل ۱۳. نمودار کلمات کلیدی هر موضوع به دست آمده



بحث و نتیجه‌گیری

در این پژوهش ما از الگوریتم‌های یادگیری ماشین و رویکرد مبتنی بر واژه جهت مدل‌سازی و تحلیل احساسات استفاده کردیم، در مجموع توانستیم ۹ موضوع مرتبط با فناوری واقعیت مجازی را استخراج کنیم، با توجه به کلمات پرتکرار و همین‌طور در نظر گرفتن اصل توثیق به نام‌گذاری موضوعات پرداختیم. با بررسی موضوعات در سال‌های مختلف مشاهده کردیم که موضوع «پیشرفت» از سال ۲۰۱۷ تا اواخر سال ۲۰۲۱ پرتعدادترین موضوع کاربران بوده است، در اوایل سال ۲۰۲۲ این موضوع جای خود را به «متاورس» داد، در حال حاضر «متاورس» از پرتعدادترین موضوعات در حال بحث کاربران می‌باشد. کسب‌وکارها در حوزه واقعیت مجازی باید در راستای جذابیت «متاورس» و جذب کاربران تلاش کنند. همین‌طور اگر نمودار «بازی‌های مجازی» و «کسب‌وکار مبتنی بر فناوری» را مشاهده کنیم می‌بینیم که تقریباً با یک نسبت در سال‌های مختلف تغییر می‌کند به این معنا که این نمودار باهم ارتباط دارند، در واقع باید کسب‌وکارها در نظر داشته باشند که عوامل تأثیرگذار روی این دو موضوع یکی است ولی در موضوع «بازی‌های مجازی» تأثیرات بیشتری دارد و اگر کسب‌وکار مبتنی بر فناوری به‌طور خاص روی بازی‌های مجازی تمرکز کنند به دلیل توجه بیشتر کاربران، رشد بیشتری خواهند داشت. همین‌طور «بازی‌های مستقل» که یکسری تغییرات داشته ولی در سال‌های اخیر یک روند کاهشی و سپس بدون تغییر داشته است، حال پدیدآوردگان این بازی‌های باید بررسی کنند و در کل «بازی‌های مجازی» موضوع موردتوجه بیشتری نسبت به «بازی‌های مستقل» بوده است. در موضوعات «آموزش» و «گجت» از ابتدای سال ۲۰۱۷ دارای سیر نزولی بوده است که نشان می‌دهد کاربران در طول زمان توجه خود به این موضوعات در حوزه واقعیت مجازی را از دست دادند، در واقع به موضوعات دیگری توجه خود را نشان دادند پس بهتر است کسب‌وکارهایی که در این حوزه فعالیت دارند تدابیری جهت تبلیغ و جذب کاربران نمایند و یا در صورت عدم رشد حوزه کاربری خود را تغییر دهند.

در مسیر انجام این مطالعه، پژوهشگران از روش یادگیری بدون ناظر استفاده کردند. با توجه به اینکه اعتبار سنجی متقاطع^۱ یک روش محبوب برای انتخاب مدل است؛ با تخمین خطای پیش‌بینی هر مدل و سپس انتخاب مدل با بهترین عملکرد کار می‌کند، با این حال، در زمینه‌های بدون نظارت، هیچ مفهوم روشی در مورد اینکه دقیقاً «خطای پیش‌بینی» چیست، وجود ندارد؛ بنابراین، استفاده از اعتبارسنجی متقابل برای انتخاب مدل در زمینه‌های بدون نظارت یا اکتشافی دشوار است (Perry, 2009).

در ادامه، به محققان آتی پیشنهاد می‌گردد، با توجه به اینکه کاربران توئیتر اغلب در پست‌های خود اغلب از ایموجی استفاده می‌کنند و ایموجی می‌تواند منبع خوبی برای تحلیل احساسات باشد، پیشنهاد می‌شود در پژوهش‌های آتی کاراکترهای ایموجی‌ها در فرآیند تحلیل احساسات مورد ارزیابی قرار گیرند، همین‌طور می‌توان تحلیل احساسات درباره فناوری واقعیت مجازی روی سایر شبکه‌های اجتماعی مانند لینک دین^۲ و یوتیوب^۳ انجام شود همین‌طور می‌توان از سایر روش‌های متن‌کاوی LSA و PLSA جهت تحلیل متون استفاده کرد.

تعارض منافع

تعارض منافع وجود ندارد.

ORCID

Fariba Karimi
Ameneh khadivar
Fateme Abbasi



<https://orcid.org/0009-0007-2669-3422>
<https://orcid.org/0000-0003-4620-1409>
<https://orcid.org/0000-0002-6918-9679>

-
1. Cross Validation
 2. LinkedIn
 3. YouTube

منابع

- نوروزی، م.، خدیور، آ.، عباسی، ف. (۱۴۰۲). مدل‌سازی و پیش‌بینی قصد خرید تلفن همراه کاربران توئیتر بر مبنای تحلیل احساسات، ۸(۱)، ۹۱-۱۱۲.
- عباسی، ف.، سهرابی، ب.، مانیان، ا.، خدیور، آ. (۱۳۹۶). ارائه مدلی جهت دسته‌بندی احساسات خریداران کتاب با استفاده از رویکرد ترکیبی، ۶(۲۱)، ۶۵-۹۲.
- عالی خانی، ی.، خدیور، آ.، عباسی، ف. (۱۴۰۱). ارزیابی ادراک عمومی از نسل پنجم ارتباطات سیار از طریق تحلیل احساسات کاربران شبکه اجتماعی توئیتر. پژوهش‌های نوین در تصمیم‌گیری، ۷(۲)، ۱۱۱-۱۳۵.
- محمدی و ناظمی. (۱۴۰۰). تجزیه و تحلیل احساسات در سطح ویژگی محصول و مبتنی بر جنسیت کاربران. مطالعات مدیریت کسب و کار هوشمند، ۱۰(۳۷): ۲۶۷-۲۹۶.
doi:10.2204/IMS.2021.52110.17235
- بصیری؛ حبیبی و نعمتی. (۱۴۰۰). تحلیل احساسات توئیتهای مرتبط با کرونا در ایران با استفاده از شبکه عصبی عمیق. مطالعات مدیریت کسب و کار هوشمند، ۱۰(۳۷): ۱۰۹-۱۳۴.
doi:10.22054/ims.2021.54705.1799

References

- Shi, Y., & Herniman, J. (2023). The role of expectation in innovation evolution: Exploring hype cycles. *Technovation*, 119, 102459. doi:https://doi.org/10.1016/j.technovation.2022.102459
- Al-Samarraie, H., Sarsam, S. M., & Alzahrani, A. I. (2023). Haptic technology in society: A sentiment analysis of public engagement. *Computers in Human Behavior*, 147, 107862. doi: https://doi.org/10.1016/j.chb.2023.107862
- Kosti, M. V., Georgakopoulou, N., Diplaris, S., Pistola, T., Chatzistavros, K., Xefteris, V.-R.,... Kompatsiaris, I. (2023). Assessing Virtual Reality Spaces for Elders Using Image-Based Sentiment Analysis and Stress Level Detection. *Sensors*, 23(8), 4130. doi: https://doi.org/10.3390/s23084130
- Bhattacharyya, M., Roy, A., Midya, S., Mitra, A., Ghosh, A., & Roy, S. (2023). *An Emoticon-Based Sentiment Aggregation on Metaverse Related Tweets*. Paper presented at the The International Conference on Artificial Intelligence and Computer Vision.
- Al-Kfairy, M., Al-Adaileh, A., Tubishat, M., Alfandi, O., BinAmro, M., & Alomari, A. (2023). *A Sentiment Analysis Approach for Identifying Users' Security and Privacy Perception of Metaverse In Twitter*.

- Paper presented at the 2023 International Conference on Smart Applications, Communications and Networking (SmartNets). doi: 10.1109/SmartNets58706.2023.10215677
- Kaur, R., Kaur, S., Mufassir Yassen.,S. (2023). *Twitter sentiment analysis for depression detection using machine learning algorithms*.
- Sánchez, P. R. P., Folgado-Fernández, J. A., & Sánchez, M. A. R. (2022). Virtual Reality Technology: Analysis based on text and opinion mining. *Mathematical Biosciences and Engineering*, 19(8), 7856-7885. doi:<https://doi.org/10.3934/mbe.2022367>
- Kim, Y. B., Lee, S. H., Kang, S. J., Choi, M. J., Lee, J., & Kim, C. H. (2015). Virtual world currency value fluctuation prediction system based on user sentiment analysis. *PloS one*, 10(8), e0132944
- Jin, B., Kim, G., Moore, M., & Rothenberg, L. (2021). Consumer store experience through virtual reality: its effect on emotional states and perceived store attractiveness. *Fashion and Textiles*, 8, 1-21. doi:<https://doi.org/10.1186/s40691-021-00256-7>
- Abbasi, F., Khadivar, A., & Yazdinejad, M. (2019). A grouping hotel recommender system based on deep learning and sentiment analysis. *Journal of Information Technology Management*, 11(2).
- Rokhsaritalemi, S., Sadeghi-Niaraki, A., & Choi, S.-M. (2020). A review on mixed reality: Current trends, challenges and prospects. *Applied Sciences*, 10(2), 636. doi:<https://doi.org/10.3390/app10020636>
- Yang, J., Li, Y., Liu, Q., Li, L., Feng, A., Wang, T.,... Lyu, J. (2020). Brief introduction of medical database and data mining technology in big data era. *Journal of Evidence-based Medicine*, 13(1), 57-69. doi:<https://doi.org/10.1111/jebm.12373>
- Osman, A. S. (2019). Data mining techniques. Retrieved from <http://ojs.mediu.edu.my/index.php/IJDSR/article/view/1841>
- Talib, R., Hanif, M. K., Ayesha, S., & Fatima, F. (2016). Text mining: techniques, applications and issues. *International Journal of Advanced Computer Science and Applications*, 7(11). Retrieved from <https://www.proquest.com/openview/86b5831a74364ad4b36255cc0f697c52/1?pq-origsite=gscholar&cbl=5444811>
- He, W. (2013). Examining students' online interaction in a live video streaming environment using data mining and text mining. *Computers in Human Behavior*, 29, 90-102. doi:<https://doi.org/10.1016/j.chb.2012.07.020>
- Mandal, S. (2013). Brief introduction of virtual reality & its challenges. *International Journal of Scientific & Engineering Research*, 4(4), 304-309.
- Craig, A. B., Sherman, W. R., & Will, J. D. (2009). *Developing virtual reality applications: Foundations of effective design*: Morgan Kaufmann.

- Zhang, L., Wang, S., & Liu, B. (2018). Deep learning for sentiment analysis: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(4), e1253. doi:<https://doi.org/10.1002/widm.1253>
- Bamodu, O., & Ye, X. M. (2013). *Virtual reality and virtual reality system components*. Paper presented at the Advanced materials research.
- Boulter, C. J., & Gilbert, J. K. (2000). Challenges and opportunities of developing models in science education. *Developing models in science education*, 343-362. doi:https://doi.org/10.1007/978-94-010-0876-1_18
- Sagayam, R., Srinivasan, S., & Roshni, S. (2012). A survey of text mining: Retrieval, extraction and indexing techniques. *International Journal of Computational Engineering Research*, 2(5), 1443-1446. Retrieved from https://fayllar.org/pars_docs/refs/650/649516/649516.pdf
- Padhy, N., Mishra, D., & Panigrahi, R. (2012). The survey of data mining applications and feature scope. *arXiv preprint arXiv:1211.5723*. doi:<https://doi.org/10.5121/ijcseit.2012.2303>
- Liao, S.-H., Chu, P.-H., & Hsiao, P.-Y. (2012). Data mining techniques and applications—A decade review from 2000 to 2011. *Expert systems with applications*, 39(12), 11303-11311. doi:<https://doi.org/10.1016/j.eswa.2012.02.063>
- Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., & Zhao, L. (2017). Latent Dirichlet Allocation (LDA) and Topic modeling: models, applications, a survey. *arXiv preprint arXiv:1711.04305*. doi:<https://doi.org/10.48550/arXiv.1711.04305>
- Islam, T. (2019). Yoga-veganism: Correlation mining of twitter health data. *arXiv preprint arXiv:1906.07668*. doi:<https://doi.org/10.48550/arXiv.1906.07668>
- Jelodar, H., Wang, Y., Yuan, C., Xia, F., Jiang, X., Li, Y., & Zhao, L. (2019). Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey. *Multimedia Tools and Applications*, 78(11), 15169-15211. doi:<https://doi.org/10.48550/arXiv.1711.04305>
- Ahmad, M., Aftab, S., Bashir, M. S., & Hameed, N. (2018). Sentiment analysis using SVM: a systematic literature review. *International Journal of Advanced Computer Science and Applications*, 9(2). doi:<https://doi.org/10.14569/IJACSA.2018.090226>
- Jang, J. Y., Hur, H. J., & Choo, H. J. (2019). How to evoke consumer approach intention toward VR stores? Sequential mediation through telepresence and experiential value. *Fashion and Textiles*, 6(1), 1-16. doi:<https://doi.org/10.1186/s40691-018-0166-9>
- Loureiro, S. M. C., Guerreiro, J., & Ali, F. (2020). 20 years of research on virtual reality and augmented reality in tourism context: A text-mining approach. *Tourism management*, 77, 104028.

doi:<https://doi.org/10.1016/j.tourman.2019.104028>

Loureiro, S. M. C., Guerreiro, J., Eloy, S., Langaro, D., & Panchapakesan, P. (2019). Understanding the use of Virtual Reality in Marketing: A text mining-based review. *Journal of Business Research*, 100, 514-530. doi:<https://doi.org/10.1016/j.jbusres.2018.10.055>

P.O.Perry. (2009). Cross-Validation for Unsupervised Learning doi: <https://doi.org/10.48550/arXiv.0909.3052>

References [in Persian]

Noroozi, M., khadivar, a., & Abbasi, F. (2023). Modeling and predicting mobile phone purchase intention of Twitter users based on sentiment analysis. *Modern Research in Decision Making*, 8(1), 91-112. Retrieved from https://journal.saim.ir/article_706378_376f06e3364192c35dc5e6d70453a5b1.pdf

Abbasi, F., Sohrabi, B., Manian, A., & Khadivar, A. (2017). A Model to Classify Book Buyers' Sentiments Using Ensemble Approach. *Business Intelligence Management Studies*, 6(21), 65-92. doi:10.22054/ims.2018.8512

Alikhani, Y., khadivar, a., & abbasi, f. (2022). Assessing the public perception of the fifth generation of mobile communication (5G) by sentiment analysis of Twitter users. *Modern Research in Decision Making*, 7(2), 111-135. Retrieved from https://journal.saim.ir/article_253208_99b09f7ff204d88f5915f532bfb16287.pdf

Mohammadi, S., & Nazemi, E. (2021). Sentiment Analysis at the Product Feature Level and Based on Users Gender. *Business Intelligence Management Studies*, 10(37), 267-296. doi:10.22054/ims.2021.52110.1723

Basiri, M. E., Habibi, S., & Nemati, S. (2021). Sentiment Analysis of Corona-Related Tweets in Iran Using Deep Neural Network. *Business Intelligence Management Studies*, 10(37), 109-134. doi:10.22054/ims.2021.54705.1799

استناد به این مقاله: کریمی، فریبا، خدیور، آمنه، کریمی، فاطمه. (۱۴۰۳). دسته‌بندی نظرات مرتبط با فناوری واقعیت مجازی با استفاده از مدل‌سازی موضوعی، مطالعات مدیریت کسب و کار هوشمند، ۱۲(۴۷)، ۱-۴۳.

DOI: 10.22054/ims.2023.74147.2342



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

