



<http://doi.org/10.22133/TLJ.2024.422177.1124>

A Systematic Review on the Use of Big Data in Tourism

Jafar Ahangaran^{1*} , Abbas Sadeghnia² 

¹ Assistant Prof. Faculty of Tourism, University of Science and Culture, Tehran, Iran.

² Ph.D. Student in Faculty of Tourism, University of Science and Culture, Tehran, Iran.

Article Info

Abstract

Original Article

Received:
24-10-2023

Accepted:
16-12-2023

Keywords:
Big Data
Social Media
Machine Learning
Systematic Mapping
Review

As technology advances and the use of social networks and the Internet grows, there's a significant increase in the production of data. The concept of "big data" is becoming more prominent in various scientific studies. However, understanding how big data is used in tourism is a challenge addressed in this study. The research investigates the application of big data in tourism studies published from 2016 to 2022, aiming to identify research gaps and potential areas for future development. To provide an overview of published works in tourism, we used the systematic mapping study method. Articles were assessed based on various factors, including research topic, conceptual and theoretical descriptions, data sources, data types and sizes, data collection methods, data analysis techniques, and reporting and visualization methods. The findings indicate a growth in research within hospitality and tourism management using analytical techniques for big data. Despite this, the research in this field is scattered and limited in scope and methodology, revealing several gaps. Additionally, the study suggests that the full potential of big data remains untapped in tourism, offering numerous research opportunities for the industry. A notable challenge highlighted is the researchers' incomplete knowledge of data-driven analysis and decision-making tools in this field.

***Corresponding author**

e-mail: a.sadeqnia@gmail.com

How to Cite:

Ahangaran, J., & Sadeghnia, A. (2024). A Systematic Review on the Use of Big Data in Tourism. *Tourism and Leisure Time Journal*, 9(18), 65-86.

Published by University of Science and Culture <https://www.usc.ac.ir>
Online ISSN: 2783-3836

1. Introduction

The speed of technological changes and their profound effects on various industries, coupled with unpredictable events such as the recent pandemic—having a significant impact on various sectors—requires these industries to cope with uncertainty (Piyush Sharma, 2020). Among them, dynamic, fluctuating, and time-sensitive industries like the tourism sector must respond promptly and more efficiently to emerging challenges and threats in order to survive (Vikrant Kaushal, 2021). Hence, the roles of technology and digitization are vital in helping industries, countries, and organizations overcome rapid challenges (Soluk, 2021).

Digitization and the adoption of new technologies have been increasing dramatically in recent years (Ritter, 2020). This includes the use of methods that change the business model to create more effective and efficient processes toward technology—arguably the most crucial change in exploitation. One significant aspect is the analysis of consumer behavior, market dynamics, supply and delivery processes, and customer communication using data obtained from the digitization process (Youssef, 2021). Meanwhile, big data stands out as one of the most well-known paradigms of information analysis, attracting both researchers and experts (José Álvarez García, 2023).

Methodologically, big-data-based approaches allow researchers to overcome the problems associated with traditional sampling methods because big data practically enables working with the entire population under investigation. This capability makes it possible to address any question related to the opinions, views, ideas, and behavior of all individuals. Simultaneously, it proves to be a powerful tool for addressing new research questions, developing innovative research designs useful in advancing knowledge, and ultimately providing support for policy and managerial decision-making (Gerard, 2016).

The most crucial factor is that the tourism industry grows based on information. Thus, tourism research requires a large amount of strong, up-to-date, timely, and relevant information to support and assist decision-making processes. Big data can provide up-to-date and highly informed inferences about human behavior and activities that boost the tourism industry (Feifei Xu, 2020). Through the traveler, a large number of data sources can inform decisions made at different stages, such as before, during, and after the trip. One popular data source is social media, used by tourists and tourism businesses to communicate, find or provide relevant information or advice, and obtain information about important news and crises (Park, 2019). Therefore, considering the importance of using big data analysis in tourism studies and stating the issue that no monitoring and analysis of the state of research conducted in this field with this theme has been done so far, the upcoming research on the analysis of the use of big data is focused on tourism research by conducting a systematic mapping review.

2. Literature review

Since no systematic mapping research has been conducted to investigate the status and process of research on the use of big data in tourism studies, the concern of this research is to use the opportunities created by big data and the need to use big data and analyze and make decisions on its basis in the tourism industry, by reviewing the literature and explaining the different dimensions of the needs of this industry, this concern and issue will become clearer, therefore, in this section, we will review the researches conducted in recent years that emphasized the necessity of using big data in the field of tourism and each offered a solution for a different need.

Describing the need to create a personal customer experience in the tourism industry, Chen et al. state that personal experiences lead to increased customer satisfaction and loyalty. Big data allows the tourism industry to analyze customer preferences, behaviors and feedback to provide recommendations (Chen, 2017). By describing the need for demand forecasting and price optimization in tourism, Jiang et al. believe that accurate demand forecasting helps tourism businesses to optimize pricing and resource allocation. Big data analytics enables the analysis of historical data, seasonal patterns and market trends for better decision making (Xiang, 2017). Referring to the need for operational efficiency and resource management in the tourism industry, Lee et al. have stated that big data helps simplify operations by providing insights into resource use, crowd management, and operational efficiency. This leads to cost savings and improved service quality (Li, 2018). Gretzel et al. also stated the need for risk management and crisis response in this industry, stating that big data helps to identify potential risks and predict crises in the tourism sector. Analyzing data from various sources helps to develop effective risk management strategies (Gretzel, 2015). Stating the need to improve marketing and customer engagement in the tourism industry, Neuhofer et al. believe that big data enables targeted marketing campaigns based on customer insights and improves the effectiveness of promotional efforts. It also increases customer engagement through data-driven

strategies (Neuhofer, 2015). Pointing to the need for tourism destination management and planning, Wang has stated that big data supports destination management by providing insight into tourist movements, preferences and impact on local infrastructure. This contributes to sustainable destination planning (Wang, 2016). Cigala also points to the need for customer feedback and reputation management in the tourism industry, noting that big data analytics enable monitoring of online comments and social media sentiment, helping businesses manage their reputation. and resolve customer concerns quickly (Sigala, 2017). With the mentioned sources, it can be concluded that the upcoming issue is the concern of knowing the amount of attention that is given to the use of big data in scientific researches in the field of tourism, and with the clarification of the necessity of using big data in tourism researches, it is necessary to address the theoretical definitions of this research.

3. Research Methodology

The purpose of this research is to review the studies that have been done in recent years in the field of tourism using big data. According to the purpose of the research, which includes answers to questions focused on the classification and organization of research conducted in the field of tourism with a focus on big data, and in which the identification, classification and analysis of existing literature related to a specific topic will be done, Petersen et al. (2008) recommend the systematic mapping review method (SMS). Systematic mapping study is a special form of literature review that complements the systematic review method (Banaeianjahromi, 2016). While a systematic review examines the research question in detail (Petersen, 2008), a systematic mapping study is a means of categorizing and summarizing the available information about the research question in an unbiased way (Wendler, 2012). In other words, the study of systematic mapping provides an overview of a specific research field with the aim of examining issues related to the identification, analysis, and organization of primary objectives, methods, and content (Kitchenham et al., 2011). In this method, various consecutive activities by following a series of independent tasks lead to reaching the final goal in this system. In this method, we often collect and review scientific articles from a specific subject area to answer some predetermined questions. The strategy behind this method is to find and evaluate all applicable articles to address specific problems that exist around a specific topic (P. Di Francesco, 2017). Therefore, in this research, we reached six different stages to reach the desired result: 1- Determining the purpose of the research 2- Determining the research questions 3- Research strategy 4- Selection criteria 5- Selecting suitable researches 6- Analyzing the results.

4. Results

By examining the results obtained from the research questions, it was found that, firstly, the use of big data analysis in scientific articles has grown significantly, and secondly, in most of the researches conducted around the topic of tourism using big data analysis, the most source of big data collection is the data produced. by users (UGC) with the approach of statistical analysis and then analysis by artificial intelligence and machine learning. This means that social networks, which are responsible for the dissemination of user-generated data, can play a more prominent role in scientific research, and the traditional method of collecting questionnaires will give way to examining the real opinions of users on social networks about a specific topic. Therefore, in response to the first question of this research, which source of big data is often used by researchers in tourism studies, it should be said that the research findings show the interest of researchers in using big data produced by users in social networks, because access to them Easier and less challenging to analyze them by artificial intelligence. In response to the second question of this research, which approaches and tools are used to analyze the big data extracted, the findings of the research indicate the dominance of the statistical analysis approach on the content generated by the user. Because according to the answer to the first question, there is often big data produced by users in the form of text and comments in social networks, and the collection of these opinions can gradually take the place of scientific research questionnaires. In response to the third question of this research, which is the type and classification of big data that support this tool and methods of analysis, it should be said that the findings of the research show that the textual analysis of the data collected with the purpose of predictive analysis has been used the most in the selected articles. Therefore, one of the most used and challenging aspects in the management of tourist destinations, i.e. the management of the high volume of input in the peak seasons, will be somewhat predictable. On the other hand, the analysis of keywords confirms that these articles are often used in order to predict the tourism demand of a destination and arrive at a prediction model, and in the second step, to examine the conditions of sustainable

development of tourist destinations by using big data analysis. In addition, the results show that most of the research in which the researcher has the desire to use big data has been used more than quantitative data analysis and less qualitative analysis has been used.

5. Conclusion

The purpose of this article was to investigate the amount of use of big data in scientific research in the field of tourism published from 2016 to 2022. Based on the purpose of this research, questions were designed and an attempt was made to answer these three questions in order to achieve the goal. The results obtained from the analysis of the findings of this research show that most of the researchers who have conducted research in the field of tourism using big data analysis are more interested in using the big data produced by users in the form of textual data and comments, as well as the purpose of their use. It has been used to forecast and predict the trend or volume of input or seasonal demand in the tourism industry. Perhaps the main reason for this is the ease of analyzing the data generated by users in the form of text in social networks. However, big data has different types and forms that can be used depending on the researcher's needs, but perhaps the researchers' lack of knowledge about other big data analysis methods makes them closer to the easiest method, which is qualitative big data analysis.

Limitations: Although the use of big data in academic research has made significant progress, there are still challenges regarding data quality, data cost, and user privacy concerns. In addition, there are several issues regarding the reliability of user-generated data, such as faking reviews about a tourist attraction to reach more users. Also, due to privacy concerns, mobile phone roaming and user transaction data have not been widely used in tourism research, but despite their benefits, they are costly and difficult to access. Because usually mobile phone operator networks, as well as business solution providers, do not want to share private information of their users. To overcome these challenges, mutual cooperation between university, government and industries becomes necessary. This collaboration not only ensures data availability and data cost reduction for tourism research using big data, but also addresses practical issues in return. Another way that comes to mind to overcome this challenge is to use social media data, as it allows people to provide their views on events, products, tools, and other topics, and this data is easier to access and should be explored in future research. Let's take advantage of this huge amount of data.

Suggestions for future research: Considering the diverse and extensive dimensions and forms of big data as well as various methods and tools that are added to their diversity day by day, the need to take advantage of this diversity in scientific research is felt, but the gap that is considered in this research It was found that researchers are not sufficiently informed about this tool and the variety of data types that are produced in large volumes in the field of information technology. One of the most important parts after collecting big data is analyzing this big data and then making data-driven decisions based on the results of these analyses. Therefore, it is suggested that researchers should get to know these methods before conducting research and choosing tools and methods, and after having sufficient mastery in the field of big data analysis, they should choose the most optimal method based on their needs and make the most of the available potential. To gain in the use of big data in the field of tourism.



<http://doi.org/10.22133/TLJ.2024.422177.1124>

گردشگری و اوقات فراغت

مروری نظام مند بر کاربرد کلان داده‌ها در صنعت گردشگری

جعفر آهنگران^{۱*}، عباس صادق‌نیا^۲

^۱ استادیار، دانشکده گردشگری، دانشگاه علم و فرهنگ، تهران، ایران.

^۲ دانشجوی دکتری، دانشکده گردشگری، دانشگاه علم و فرهنگ، تهران، ایران.

اطلاعات مقاله

مقاله پژوهشی

تاریخ دریافت:

۱۴۰۲/۰۸/۰۲

تاریخ پذیرش:

۱۴۰۲/۰۹/۲۵

واژگان کلیدی:

کلان داده‌ها
مطالعه نگاشت نظام مند
شبکه‌های اجتماعی
یادگیری ماشین

با پیشرفت فناوری و افزایش استفاده از شبکه‌های اجتماعی و اینترنت، حجم تولید داده‌ها افزایش چشمگیری داشته و بهره‌گیری از مفهوم «کلان داده‌ها» نیز به تبع آن، به‌طور فزاینده‌ای در پژوهش‌های علمی مختلف به چشم می‌خورد؛ اما سازمان‌دهی و تحلیل ادبیات موجود با محوریت بهره‌گیری از کلان داده‌ها در حوزه گردشگری برای شناسایی وضعیت تحقیقات صورت گرفته در این عرصه، خود مسئله‌ای است که هدف این مطالعه است؛ از این رو این پژوهش به بررسی میزان بهره‌گیری از کلان داده در تحقیقات گردشگری منتشر شده از سال ۲۰۱۶ تا ۲۰۲۲، به شناسایی شکاف‌های پژوهشی و فرصت‌های پیشرفت‌های آتی پرداخته است. از آنجاکه این پژوهش در پی ارائه تصویری کلی از کارهای انجام شده در مقالات منتشر شده در حوزه گردشگری است، از روش مطالعه نگاشت نظام مند استفاده شده است. مقاله‌ها براساس این ویژگی‌ها بررسی شدند: موضوع تحقیق، توصیف مفهومی و نظری، منابع داده، نوع داده و اندازه، روش‌های جمع‌آوری داده‌ها، روش‌های تجزیه و تحلیل داده‌ها و گزارش و تجسم داده‌ها. یافته‌های این مطالعه حاکی از افزایش تحقیقات انجام شده در حوزه مدیریت مهمان‌نوازی و گردشگری با استفاده از روش‌های تحلیلی برای کلان داده‌هاست. با این حال، این حوزه پژوهشی از نظر دامنه نسبتاً پراکنده و در روش‌شناسی محدود است و شکاف‌های متعددی را نشان می‌دهد. نتیجه دیگر آن‌که پژوهش‌های گردآوری شده نشان می‌دهند هنوز از پتانسیل کلان داده‌ها به‌طور کامل استفاده نشده و فرصت‌های زیادی پیش‌روی صنعت گردشگری از بعد پژوهشی وجود دارد. همچنین آشنانودن کامل پژوهشگران این حوزه با ابزارهای تحلیل و تصمیم‌گیری داده‌محور چالشی است که در این پژوهش نمایان شده است.

*نویسنده مسئول

رایانامه: a.sadeqnia@gmail.com

نحوه استناددهی:

آهنگران، جعفر، و صادق‌نیا، عباس (۱۴۰۳). مروری نظام مند بر کاربرد کلان داده‌ها در صنعت گردشگری. گردشگری و اوقات فراغت، ۹(۱۸)، ۶۵-۸۶.

ناشر: دانشگاه علم و فرهنگ <https://www.usc.ac.ir>

شاپای الکترونیکی: ۲۷۸۳-۳۸۳۶

مقدمه

گونگون داشت، این صنایع را ملزم کرد تا با عدم قطعیت کنار آیند (Sharma et al., 2020). از این میان، صنایع پویا و پرنوسان و همچنین حساس به زمان مانند صنعت گردشگری باید به سرعت و درعین حال با کارآمدی بیشتری به چالش‌ها و تهدیدهای درحال ظهور پاسخ دهند تا بتواند به بقای خود ادامه دهد (Kaushal & Srivastava, 2021)؛ از این رو نقش فناوری و دیجیتالی شدن در کمک به صنایع، کشورها و سازمان‌ها برای غلبه بر چالش‌های سریع بسیار حیاتی است (Soluk et al., 2021). دیجیتالی شدن و بهره‌گیری از فناوری‌های جدید در سال‌های اخیر، به‌طور چشم‌گیری رو به افزایش است (Ritter et al., 2020) و شامل بهره‌گیری از روش‌هایی می‌شود که مدل کسب‌وکار را در راستای خلق فرایندهای مؤثرتر و کارا تر به سمت فناوری تغییر می‌دهد. شاید مهم‌ترین تغییر بهره‌جویی از تحلیل‌هایی باشد که بر رفتار مصرف‌کننده، رفتار بازار، فرایند تأمین و تحویل و ارتباط با مشتری از طریق داده‌های به‌دست‌آمده از فرایند دیجیتالی شدن انجام می‌شود (Youssef et al., 2021). در این میان، کلان‌داده‌ها یکی از شناخته‌شده‌ترین پارادایم‌های تحلیل اطلاعات است که هم محققان و هم متخصصان را به خود جذب کرده است (Garcia et al., 2023).

از نظر روش‌شناسی، رویکردهای مبتنی بر کلان‌داده به محققان اجازه می‌دهد بر مشکلات روش‌های نمونه‌گیری سنتی غلبه کنند؛ زیرا کلان‌داده عملاً اجازه می‌دهد با کل جمعیت تحت بررسی کار کنند. ظاهراً این امکان را فراهم می‌کند تا به هر سؤالی که مربوط به نظرات، دیدگاه‌ها، ایده‌ها و رفتار تمامی افراد باشد پاسخ دهد. درعین حال، به نظر می‌رسد که ابزار قدرتمندی برای پرداختن به سؤالات پژوهشی جدید، توسعه طرح‌های پژوهشی نوآورانه مفید در پیشرفت دانش، و درنهایت حمایت از سیاست و تصمیم‌مدیریتی باشد (Gerard et al., 2016). نکته بسیار مهم این است که صنعت گردشگری براساس اطلاعات رشد می‌کند. به این ترتیب، تحقیق درباره گردشگری نیازمند حجم زیادی از اطلاعات قوی، به‌روز، به‌موقع و مرتبط برای پشتیبانی و کمک به فرایندهای تصمیم‌گیری است. کلان‌داده‌ها می‌توانند استنباط‌های به‌روز و بسیار آگاهانه درباره رفتار و فعالیت‌های انسانی، که صنعت گردشگری را تقویت می‌کنند، ارائه دهند (Xu et al., 2020). از طریق گردشگر، تعداد زیادی از منابع داده می‌توانند تصمیم‌های اتخاذشده در مراحل مختلف، مانند قبل، حین و بعد از سفر را به اطلاع برسانند. یکی از نمونه‌های یک منبع داده محبوب، رسانه‌های اجتماعی هستند که گردشگران و همچنین کسب‌وکارهای گردشگری برای برقراری ارتباط، یافتن یا ارائه اطلاعات مرتبط یا هر توصیه‌ای و همچنین کسب اطلاعات درباره اخبار و بحران‌های مهم استفاده می‌کنند (Park, 2019)؛ از این رو با در نظر گرفتن اهمیت بهره‌گیری از تحلیل کلان‌داده‌ها در مطالعات حوزه گردشگری و بیان این مسئله که تاکنون هیچ نظارت و تحلیلی بر وضعیت پژوهش‌های انجام‌شده در این حوزه با این مضمون انجام نشده است، تحقیق پیش‌رو بر تجزیه و تحلیل میزان استفاده از داده‌های بزرگ در تحقیقات گردشگری با انجام یک مرور نگاشت نظام‌مند متمرکز است.

مبانی نظری و مرور پیشینه

از آنجاکه تاکنون پژوهشی با روش نگاشت نظام‌مند برای بررسی وضعیت و روند تحقیقات صورت‌گرفته بر روی استفاده از کلان‌داده‌ها در مطالعات گردشگری انجام نشده است، دغدغه این پژوهش استفاده از فرصت‌هایی که کلان‌داده‌ها خلق کردند، لزوم بهره‌گیری از کلان‌داده‌ها و تحلیل و تصمیم‌گیری بر مبنای آن در صنعت گردشگری است، با بررسی ادبیات و تشریح ابعاد مختلف نیازهای این صنعت این دغدغه و مسئله شفاف‌تر خواهد شد؛ از این رو در این بخش به بررسی پژوهش‌های انجام‌شده در سال‌های اخیر می‌پردازیم که بر ضرورت بهره‌گیری از کلان‌داده در حوزه گردشگری تأکید داشتند و هر یک راه‌حلی برای نیازهای مختلف ارائه کردند.

چن^۱ و همکاران (2017) با تشریح نیاز به خلق تجربه شخصی مشتری در صنعت گردشگری بیان می‌دارند که تجربیات شخصی به افزایش رضایت و وفاداری مشتری منجر می‌شوند. کلان‌داده‌ها به صنعت گردشگری اجازه می‌دهند ترجیحات، رفتارها و بازخوردهای مشتری را برای ارائه پیشنهادهای تجزیه‌وتحلیل‌کننده. ژیانگ و همکاران با تشریح نیاز به پیش‌بینی تقاضا و بهینه‌سازی قیمت در گردشگری، معتقد است پیش‌بینی دقیق تقاضا به کسب‌وکارهای گردشگری کمک می‌کند تا قیمت‌گذاری و تخصیص منابع را بهینه کنند. تجزیه‌وتحلیل داده‌های بزرگ، تجزیه‌وتحلیل داده‌های تاریخی، الگوهای فصلی و روند بازار را برای تصمیم‌گیری بهتر امکان‌پذیر می‌کند (Xiang et al., 2017). لی و همکاران با اشاره به نیاز به بهره‌وری عملیاتی و مدیریت منابع در صنعت گردشگری، بیان داشته‌اند که کلان‌داده با ارائه بینش‌هایی درباره استفاده از منابع، مدیریت جمعیت و کارایی عملیاتی به ساده‌سازی عملیات کمک می‌کند. این امر به صرفه‌جویی در هزینه و بهبود کیفیت خدمات منجر می‌شود (Li et al., 2018). گرتزل و همکاران نیز با بیان نیاز به مدیریت ریسک و واکنش به بحران در این صنعت، بیان داشته است که کلان‌داده‌ها به شناسایی ریسک‌های بالقوه و پیش‌بینی بحران‌ها در بخش گردشگری کمک می‌کند. تجزیه‌وتحلیل داده‌ها از منابع مختلف به توسعه راهبردهای مدیریت ریسک مؤثر کمک می‌کند (Gretzel et al., 2015). نیوهوفر و همکاران با بیان نیاز به بهبود بازاریابی و تعامل با مشتری در صنعت گردشگری، معتقدند که کلان‌داده کمپین‌های بازاریابی هدفمند را براساس بینش مشتری امکان‌پذیر می‌کند و اثربخشی تلاش‌های تبلیغاتی را بهبود می‌بخشد. همچنین تعامل مشتری را از طریق راهبردهای داده‌محور افزایش می‌دهد (Neuhof et al., 2015). وانگ با اشاره به نیاز مدیریت و برنامه‌ریزی مقصد گردشگری بیان داشته که کلان‌داده با ارائه بینشی درباره حرکات توریستی، ترجیحات و تأثیر در زیرساخت‌های محلی، از مدیریت مقصد پشتیبانی می‌کند. این به برنامه‌ریزی مقصد پایدار کمک می‌کند (Wang et al., 2016). سیگالا و همکاران نیز با اشاره به نیاز به بازخورد مشتری و مدیریت شهرت در صنعت گردشگری، به این نکته اشاره دارد که تجزیه‌وتحلیل داده‌های بزرگ نظارت بر نظرات آنلاین و احساسات رسانه‌های اجتماعی را امکان‌پذیر می‌کند و به کسب‌وکارها کمک می‌کند شهرت خود را مدیریت کنند و نگرانی‌های مشتریان را به سرعت برطرف سازند (Sigala et al., 2017). با منابع ذکرشده می‌توان نتیجه گرفت مسئله پیش‌رو دغدغه اطلاع از میزان توجهی است که بر استفاده از کلان‌داده در پژوهش‌های علمی حوزه گردشگری صورت می‌گیرد و با روشن شدن لزوم بهره‌گیری از کلان‌داده‌ها در پژوهش‌های گردشگری، لازم است به تعاریف نظری این پژوهش بپردازیم.

کلان‌داده‌ها

برخی از محققان کلان‌داده را نوعی نوش‌دارو توصیف کرده‌اند که می‌تواند بینش مفیدی را درباره بسیاری از جنبه‌های زندگی افراد، سازمان‌ها و بازارها ارائه دهد. برخی دیگر از محققان آن را پدیده‌ای فرهنگی، فناوری و علمی ترسیم کرده‌اند که متکی بر تعامل فناوری، تحلیل و حتی روش‌شناسی است (Mayer-Schonberger, 2014)؛ البته باید توجه داشت کلان‌داده‌ها بدون توسعه فناوری وجود نخواهند داشت. در طول سه دهه گذشته، تعداد دستگاه‌هایی که امکان اتصال به اینترنت را فراهم می‌کنند به طور تصاعدی افزایش یافته است. این دستگاه‌ها باعث ازدیاد داده‌ها شده‌اند که اغلب با محتوای تولیدشده به دست کاربر در شبکه‌های اجتماعی آنلاین مرتبط هستند؛ اما به اختصار باید گفت کلان‌داده‌ها به مجموعه‌های داده‌های بزرگ و پیچیده‌ای اشاره دارد که به ابزارهای مدیریت و تجزیه‌وتحلیل داده‌های پیچیده و مقرون‌به‌صرفه برای استخراج بینش و تصمیم‌گیری نیاز دارند که اغلب از نظر نوع در سه دسته داده‌های ساختاریافته، نیمه‌ساختاریافته و بدون ساختار جای می‌گیرند (Turikpenova & Abitova, 2023). داده‌های ساختاریافته داده‌هایی هستند که قابلیت اندازه‌گیری، تقسیم‌بندی و مرتب‌سازی دارند؛ مانند داده‌های موجود در یک پایگاه داده. داده‌های بدون ساختار شامل داده‌هایی مانند متن و عکس و صدا و تصویری هستند که قابلیت سازمان‌دهی و دسته‌بندی ندارند و برای تحلیل آن‌ها نیاز به ابزارهای خاص و ویژه است. داده‌های نیمه‌ساختاریافته نیز ترکیبی از این دو نوع داده هستند و بهترین مثال برای آن، داده‌های تگ‌بندی‌شده برای اشتراک‌گذاری داده‌ها در وبسایت‌ها مانند داده‌های XML است (Amudhavel et al., 2015).

1. Chen

در سال ۲۰۰۱، داگلاس لین^۱ برای اولین بار کلان داده را برحسب حجم، سرعت و تنوع تعریف کرد که این تعریف در آن زمان به منطقی ترین و محبوب ترین تعریف داده های بزرگ (تعریف 3V) تبدیل شد. پس از چندی، مارک بیر^۲ در تیم تحقیق بر روی توسعه کلان داده ها به V اضافه کرد و در نتیجه در سال ۲۰۱۲ آن دو تعریف کلان داده ها را با اضافه کردن دو V دیگر بسط داد: ارزش و صحت. این دو خصیصه جدید برای برطرف کردن نیازها و اهداف کسب و کارها ضروری بود؛ برای مثال بدون وجود صحت و ارزش در کلان داده ها، هر داده جعلی و بی معنی ای می تواند در دسته کلان داده ها جای گیرد و تصمیم گیری های مبتنی بر داده یک کسب و کار را دچار انحراف کند. تا به امروز این پنج خصیصه (5V) جامع ترین تعریف برای کلان داده ها هستند (UNWTO, 2021).

جدول ۱: پنج ویژگی کلان داده ها (UNWTO, 2021)

حجم (Volume)	تنوع (Variety)	سرعت (Velocity)	صحت (Veracity)	ارزش (Value)
بیانگر اندازه و حجم میزان داده ای است که از منبع مدنظر جمع آوری شده است	به معنای نوع داده در یکی از اشکال ساختار بندی شده، نیمه ساختار یافته و یا ساختار بندی نشده است.	بیانگر میزان سرعتی است که در آن کلان داده های مدنظر تولید و گردآوری می شوند.	به معنای میزان درجه ای از اعتبار است که می توان به آن مقدار به کلان داده های مدنظر اعتماد کرد.	به ارزش داده های گردآوری شده از منبع مدنظر اشاره دارد.

یادگیری ماشین

یادگیری ماشین^۳ زیرمجموعه ای از هوش مصنوعی^۴ است که به منزله مجموعه ای از الگوریتم های مختلف برای آموزش رایانه ها به منظور یافتن الگوهایی در داده ها با هدف پیش بینی و پیشگویی آینده یا بررسی کیفیت برای بهینه سازی عملکرد یا بهبود یک پارامتر استفاده می شود و به رایانه ها امکان یادگیری بدون برنامه ریزی صریح را می دهد. یادگیری ماشین زمینه ای از مطالعه است که به استفاده از الگوریتم های محاسباتی برای تبدیل داده های تجربی به مدل هایی می پردازد که قابلیت استفاده دارند. زمینه یادگیری ماشینی از آمار سنتی و کاربردهای هوش مصنوعی نشئت گرفته است. از تلاش های شرکت های بزرگ مانند گوگل، مایکروسافت، فیس بوک، آمازون و غیره، یادگیری ماشینی به یکی از داغ ترین موضوعات علوم محاسباتی در دهه گذشته تبدیل شده است. از طریق فرایندهای تجاری آنها، حجم عظیمی از داده ها جمع آوری شده و خواهد شد و بنابراین فرصتی را برای تقویت مجدد رویکردهای آماری و محاسباتی برای تولید خودکار مدل های مفید از داده ها فراهم کرده است (Witten et al., 2011).

شبکه های اجتماعی

امروزه فناوری های اطلاعاتی و ارتباطی به تمامی فضاها، سازمانی و فردی در جوامع بشری نفوذ کرده و شیوه زندگی روزمره افراد را تا حدود درخور ملاحظه ای تحت تأثیر قرار داده است. بهره گیری از این فناوری ها، به مثابه وقوع یک «چرخش پارادایمی» در زندگی فردی و اجتماعی انسان است که در آن، حتی مفاهیم کهن نیز با هندسه معرفتی جدیدی شناخته می شوند (Dawit Negussie, 2023). شبکه های اجتماعی از گروه هایی عموماً فردی یا سازمانی تشکیل شده اند که از طریق یک یا چند نوع وابستگی به هم متصل اند و در بستر یک جامعه اطلاعاتی پیچیده، کارکرد مؤثر شبکه همگرا را تصویر می کنند و موفقیت و محبوبیت آنها به دلیل داشتن رنگ و بوی اجتماعی است. در واقع، شبکه های اجتماعی نسل جدیدی از وبسایت ها هستند که این روزها در کانون توجه کاربران شبکه جهانی اینترنت قرار گرفته اند (Dwivedi)

1. Douglas Laney
2. Mark A. Beyer
3. Machine Learning
4. Artificial Intelligence

et al., 2023). این‌گونه سایت‌ها بر مبنای تشکیل اجتماعات برخط فعالیت می‌کنند و هریک دسته‌ای از کاربران اینترنتی را با ویژگی خاصی گردهم می‌آورند. به سبب حجم ارتباطات گسترده‌ای که در محیط این شبکه‌ها امکان پذیر است و همچنین امکان برقراری ارتباط با مشخصات غیرواقعی، این سایت‌ها سهم بسزایی از کل زمان صرف‌شده در اینترنت در جهان را به خود اختصاص داده‌اند. مصرف‌کنندگان در فضاهایی مانند فیس‌بوک، یوتیوب یا توئیتر زمان زیادی را صرف به اشتراک‌گذاری تجربیات یکدیگر پیرامون برند در این فضای ارتباطی جدید می‌کنند. از آنجاکه در فضای این شبکه‌های اجتماعی، محتوا به دست کاربران و به صورت تعاملی ایجاد می‌شود، قدرت افراد و اجتماعات برای اثرگذاری بر برندهای موجود یا حتی جاودانه کردن برندهای قدیمی در حال افزایش است. الیسون^۱ از سه عنصر اصلی برای تعریف شبکه‌های اجتماعی استفاده کرده است: (۱) کاربران باید قادر به ساخت پروفایل شخصی، عمومی یا نیمه‌خصوصی خود در آن باشند؛ (۲) کاربران باید قادر به برقراری ارتباط با دیگران در آن شبکه باشند؛ (۳) کاربران باید بتوانند رفتار و کارهای سایر کاربران را تعقیب کنند (Boyd & Ellison, 2007).

تجزیه و تحلیل کلان‌داده‌ها

تفسیر و محاسبه نظام‌مند داده‌ها با استفاده از روش‌های آماری به منزله تجزیه و تحلیل شناخته می‌شود. تجزیه و تحلیل از ریاضیات، آمار و هوش مصنوعی برای کمک به بررسی و تفسیر داده‌هایی استفاده می‌کند که اغلب درک‌ناپذیرند تا تصمیم‌های بهتری به کمک آنها گرفته شود. در واقع تجزیه و تحلیل کلان‌داده‌ها نوعی آشکارسازی الگوها، روندها و بینش‌ها از داده‌های نامرتب عظیم است. برای نمونه متن‌کاوی، نظریه گراف اجتماعی، نظرکاوی، تجزیه و تحلیل تأثیر اجتماعی، تحلیل احساسات، تجزیه و تحلیل آماری، تجزیه و تحلیل ریسک سایبری و موارد دیگر برخی از رویکردهای متنوع تجزیه و تحلیل کلان‌داده‌ها در رسانه‌های اجتماعی هستند (Theodorakopoulos & Halkiopolus et al., 2023). شرکت‌های مختلف ممکن است از نتایج تجزیه و تحلیل کلان‌داده‌ها برای بهبود راهبردهای تولید یا بازاریابی خود برای رقابت در دنیای تجارت دیجیتال استفاده کنند. برای مثال، تجزیه و تحلیل رسانه‌های اجتماعی ممکن است به کسب‌وکارها کمک کند تا نظرات کاربران را درباره محصولات خود دریافت کنند، که در نتیجه می‌تواند برای ایجاد تغییرات و کسب ارزش بیشتر از برند خود استفاده شود (Skielse & Ayele, 2017).

روش‌شناسی تحقیق

هدف این تحقیق بررسی مطالعاتی است که در سال‌های اخیر در حوزه گردشگری با بهره‌گیری از داده‌های کلان انجام شده است. با توجه به هدف پژوهش، که شامل پاسخ به سؤالات متمرکز بر طبقه‌بندی و سازمان‌دهی پژوهش‌های انجام‌شده در حوزه گردشگری با محوریت داده‌های کلان است و در آن شناسایی، طبقه‌بندی و تحلیل ادبیات موجود مرتبط با موضوع خاصی صورت خواهد گرفت، پترسن و همکاران (2008) در این راستا روش مرور نگاشت نظام‌مند (SMS^۲) را توصیه می‌کنند. مطالعه نگاشت نظام‌مند شکل خاصی از مرور ادبیات است که روش مرور نظام‌مند را تکمیل می‌کند (Banaeianjahromi & Smolander, 2016). در حالی که مرور نظام‌مند سؤال تحقیق را به تفصیل بررسی می‌کند (Petersen, 2008)، مطالعه نگاشت نظام‌مند وسیله‌ای برای طبقه‌بندی و خلاصه کردن اطلاعات موجود درباره سؤال تحقیق به روشی بی‌طرفانه است (Wendler, 2012). به عبارت دیگر، مطالعه نگاشت نظام‌مند با هدف بررسی موضوعات مرتبط با شناسایی، تجزیه و تحلیل و سازمان‌دهی اهداف، روش‌ها و محتوای اولیه، مروری بر یک زمینه تحقیقاتی خاص ارائه می‌دهد (Kitchenham et al., 2011). در این روش، فعالیت‌های مختلف متوالی با دنبال کردن مجموعه‌ای از وظایف مستقل به رسیدن به هدف نهایی در این سیستم منجر می‌شود. در این روش، اغلب به جمع‌آوری و بررسی مقالات علمی از یک حوزه موضوعی خاص برای پاسخ به برخی از سؤالات از پیش تعیین شده می‌پردازیم. راهبرد پشت این روش، یافتن و ارزیابی همه مقالات قابل اجرا برای رسیدگی به مشکلات خاصی است که حول یک موضوع خاص وجود دارد (Di

1. Ellison
2. Systematic Mapping Study

(Francesco et al., 2017). از این رو، در این تحقیق به شش مرحله مختلف برای رسیدن به نتیجه مطلوب رسیدیم: ۱. تعیین هدف تحقیق؛ ۲. تعیین سؤالات تحقیق؛ ۳. راهبرد تحقیق؛ ۴. معیارهای انتخاب؛ ۵. انتخاب پژوهش‌های مناسب؛ ۶. تحلیل نتایج. در شکل ۱، فرایند یک مطالعه نگاهت نظام‌مند مشاهده می‌شود (Schmerl et al., 2017).



شکل ۱: فرایند مطالعه نگاهت نظام‌مند

هدف تحقیق پیش‌رو

هدف این پژوهش بررسی میزان بهره‌گیری از کلان‌داده در تحقیقات علمی در زمینه گردشگری، با شناسایی شکاف‌های پژوهشی و پیشرفت‌های آتی و طراحی دستور کار برای تحقیقات آینده است؛ از این رو تجزیه و تحلیل ادبیات علمی موجود به منظور توصیف استفاده از کلان‌داده‌ها با توجه به رویکردهایی مانند فرایندها، روش‌ها و ابزارها از دیدگاه محققان و متخصصان شکل می‌گیرد؛ البته باید به این نکته توجه داشت که قابلیت تنظیم و تنوع در داده‌های اجتماعی بزرگ نیازمند تجزیه و تحلیل داده‌ها، یادگیری ماشینی و علم داده در حوزه تحلیل داده است. هدف دیگر این تحقیق، نشان دادن این موضوع است که امروز تا چه میزان رسانه‌های اجتماعی منبع خوبی برای کلان‌داده‌ها براساس آمارهای واقعی هستند.

سؤالات تحقیق

براساس هدف این پژوهش، سؤالات زیر طراحی شده‌اند تا برای فرمول‌بندی و سازمان‌دهی فرایند پژوهش ما را کمک کنند:

- ۱) چه منبعی از کلان‌داده‌ها در مطالعات گردشگری استفاده می‌شود؟
- ۲) چه رویکردها و ابزارهایی برای تحلیل کلان‌داده‌های استخراج‌شده استفاده نمی‌شوند؟
- ۳) نوع و دسته‌بندی کلان‌داده‌هایی که این ابزار و روش‌های تحلیل را پشتیبانی می‌کنند کدام‌اند؟

راهبرد تحقیق

به منظور یافتن پاسخ سؤالات مطرح‌شده در قسمت قبل، یک راهبرد خاص برای جست‌وجو و یافتن منابع علمی مختلف مبتنی بر کلیدواژه‌های مدنظر تدوین شد؛ از این رو از منابع مختلف و مرتبط به یافتن مقالاتی پرداخت که در عناوین آن‌ها کلیدواژه‌هایی چون کلان‌داده‌ها، داده‌های شبکه‌های اجتماعی، تحلیل داده‌ها و حتی شبکه‌های اجتماعی وجود داشت؛ با این توضیح که زمینه این تحقیق‌ها باید با حوزه گردشگری مرتبط می‌بود. با توجه به محدودیت زمان انتشار، کیفیت و قابلیت دسترسی عناوین انتخابی مجبور به انتخاب شش منبع دیجیتال برای یافتن مقالات

علمی شدیم: کتابخانه آنلاین وایلی^۱، اسپرینگر^۲، ساینس دایرکت^۳، ای‌تریپل‌ای^۴، کتابخانه دیجیتال ای‌سی‌ام^۵ و اسکوپوس^۶ که اغلب از طریق سایت ریسرچ گیت^۷ به صورت رایگان و تمام‌متن دسترس‌پذیر بودند.

معیارهای انتخاب

معیارهای متفاوتی برای انتخاب مقالات مدنظر قرار گرفت که در ادامه ذکر شده‌اند:

معیار ۱: تحقیق مدنظر با مقوله گردشگری، تحلیل داده‌های کلان و شبکه‌های اجتماعی مرتبط باشد.

معیار ۲: مقاله مدنظر به منزله مقاله علمی و به زبان انگلیسی منتشر شده باشد.

معیار ۳: متن کامل مقاله به صورت آزاد در دسترس باشد.

معیار ۴: مقاله پژوهشی یا مروری باشد.

انتخاب پژوهش مناسب

در فرایند انتخاب پژوهش مناسب از نرم‌افزار تحت وب اس‌ال‌آر تول^۸ استفاده شده است. این نرم‌افزار به مدیریت فرایند SLR کمک می‌کند و تمرکز آن بر روی افزایش کیفیت و کارایی فرایند جست‌وجو و انتخاب نتایج حاصله از موتورهای جست‌وجوی پایگاه‌های مقالات مختلف است؛ برای مثال ابتدا در موتور جست‌وجوی پایگاه ساینس دایرکت به یافتن کلیدواژه‌هایی مانند Big data و Tourism پرداختیم که حاصل آن ۲۹۹ مقاله علمی مختلف بود. فهرست خروجی این جست‌وجو به صورت فایل BIBTEXT وارد نرم‌افزار اس‌ال‌آر تول شد و در مرحله اول با کمک این نرم‌افزار کلیدواژه‌های مدنظر در تمام این ۲۹۹ مقاله منتخب بررسی و ۵۸ مقاله در این فاز انتخاب شد. در فاز بعدی، عناوین این ۵۸ مقاله بررسی و بسته به ارتباط با موضوع و معیارهای مدنظر، ۳۵ مقاله تأیید شد. در فاز بعد، چکیده این مقالات بررسی و ۲۰ مقاله انتخاب شد. در فاز نهایی، محتوای کل مقاله مطالعه شد و پس از تطبیق با معیارهای انتخابی، فقط ۸ مقاله تأیید شد. شکل ۲ نمایی از خروجی این نرم‌افزار برای پایگاه مقالات ساینس دایرکت را نشان می‌دهد.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

1. Wiley Online Library
2. Springer Link
3. Science Direct
4. IEEE Xplore
5. ACM Digital Library
6. Scopus
7. ResearchGate
8. ww.slr-tool.com

شکل ۲: نمونه خروجی نرم افزار SLR TOOL

The screenshot displays the SLR Tool interface. At the top, there is a navigation bar with 'SLRTool', 'Projectlist', 'About', and 'Contact'. A user is logged in as 'Hello a.sadeqnia@gmail.com!'. The main content area shows a search for 'Project big data' with various filters and a list of results. The first result is 'Big data analytics and sustainable tourism: A comprehensive review and network based analysis for potential future research' by Agrawal, Rohit; Wankhede, Vishal A; Kumar, Anil; Luthra, Sunil; Huisingh, Donald. The second result is 'Big Data analytics for forecasting tourism destination arrivals with the applied Vector Autoregression model' by Liu, Yuan-Yuan; Tseng, Fang-Mei; Tseng, Yi-Heng.

با بررسی سایر پایگاه‌ها، از ۸۱۹ پژوهش گردآوری شده در فاز اول و طی کردن مراحل بالا به ۲۰ مقاله نهایی رسیدیم که بیشترین ارتباط را با موضوع انتخابی و معیارهای ذکر شده داشتند. جدول ۲ آمار دقیق تری از فرایند انتخاب مقالات را با توجه به پایگاه مقالات ارائه می کند.

جدول ۱: تعداد مقالات انتخابی براساس منابع گردآوری

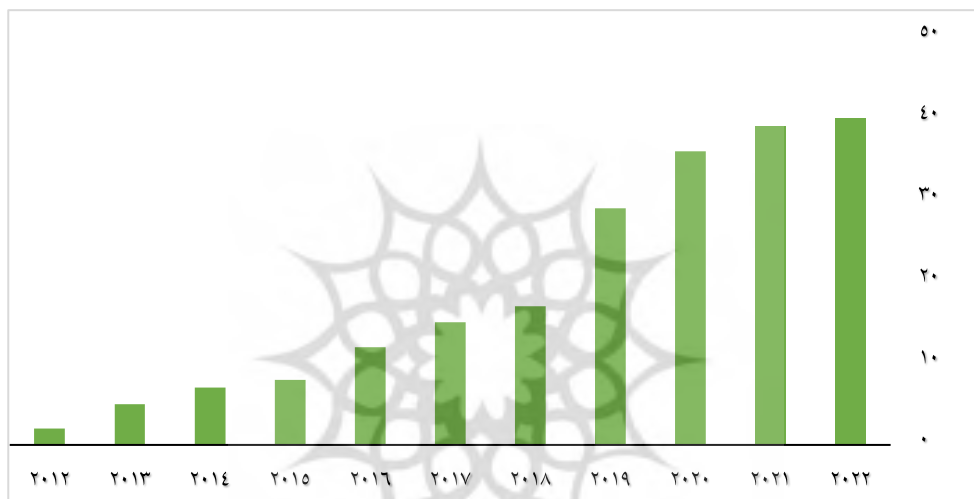
منبع گردآوری	تعداد مقالات ابتدایی	تعداد مقالات نهایی از نرم افزار SLR Tool
Wiley Online Library	۳۵	۱
Springer Link	۱۵۰	۳
Science Direct	۲۹۹	۸
IEEE Xplore	۱۴۴	۳
ACM Digital Library	۱۰۲	۳
Scopus	۸۹	۲

یافته‌ها

در این بخش، نتیجه‌ای از یافته‌های این پژوهش را براساس تحلیل داده‌های به‌دست‌آمده از سایر پژوهش‌های گردآوری‌شده به‌منظور رسیدن به پاسخ سؤالات تحقیق ارائه می‌کنیم.

ویژگی پژوهش‌های انتخابی

شکل زیر تعداد پژوهش‌های منتشرشده براساس سال انتشار حول موضوع بهره‌گیری از کلان‌داده‌ها در صنعت گردشگری را نشان می‌دهد. براساس این نمودار، دو نتیجه کلی برداشت می‌شود: اول این‌که پژوهش بر روی موضوع استفاده از کلان‌داده در صنعت گردشگری هنوز در ابتدای راه خود قرار دارد و دوم اینکه در سه سال اخیر، رشد و توجه شایان توجهی بر روی این موضوع در حال شکل‌گیری است و می‌توان پیش‌بینی کرد در سال‌های پیش‌رو نیز این رشد به‌صورت نمایی ادامه خواهد داشت.



شکل ۳: تعداد مقالات مرتبط با استفاده از کلان‌داده‌ها در زمینه گردشگری براساس سال انتشار

از منظر نوع تحقیق، اغلب پژوهش‌ها (۵۸ درصد) تحقیق‌های کمی بودند (برای نمونه در آن‌ها از روش‌هایی مانند تحلیل آماری، مدل‌سازی ریاضی، تحلیل فضایی و داده‌کاوی‌های عددی استفاده شده بود). همچنین، ۲۵ درصد نیز تحقیق‌های کیفی به‌شمار می‌آیند (که در آن‌ها از پرسش‌نامه‌های کیفی و تحلیل تشریحی و کیفی بهره‌گرفته شده بود) و حدود ۵ درصد نیز ترکیبی از این دو روش استفاده کرده بودند. بقیه پژوهش‌ها تحقیقاتی نظری و با ادبیات مروری به‌شمار می‌آیند. با بررسی عناوین و چکیده مقالات انتخابی و قراردادن آن‌ها در نرم‌افزار وردکلودز^۱ در شکل ۴، به ترسیم ابرکلمات مرتبط با این مقالات پرداختیم. در این شکل، کلمات پرتکرار با اندازه‌ای برجسته‌تر نمایان هستند که به این معنی است که مقوله کلان‌داده و گردشگری عمدتاً با کلیدواژه‌هایی مانند پیش‌بینی، تقاضا، توسعه پایدار، مقصد و عملکرد بررسی و مطالعه شده است.

1. WordClouds.com

محتوای تولیدشده به دست کاربر (UGC)¹

توسعه فناوری و شکل‌گیری شبکه‌های اجتماعی و وبسایت‌های مختلف گردشگران را برای به اشتراک گذاشتن تجربیات سفر خود هرچه بیشتر تشویق می‌کنند و باعث تسهیل فرایند اشتراک‌گذاری سفر می‌شوند (Cong et al., 2014)؛ برای مثال در عصر گردشگری مدرن، شبکه‌های اجتماعی ابزاری برای به اشتراک‌گذاری تجربیات و عکس‌ها و نظرات و عقاید افرادی است که یک مقصد یا جاذبه گردشگری خاص را تجربه کرده‌اند و این رسانه‌های اجتماعی در قالب پلتفرم، به گردشگران اجازه می‌دهند که اطلاعات دریافت کنند و همچنین افکار و توصیه‌های مثبت یا منفی خود را درباره محل اقامت، مکان‌ها، جاذبه‌ها یا خدمات به اشتراک بگذارند (Van der Zee & Bertocchi, 2018). همین ردپای گردشگران داده‌هایی تولید می‌کند که هم برای گردشگران جدید مفید است و هم برای تحلیلگران کسب‌وکار و مقصد گردشگری یا برای سیاست‌گذاری در سطح کلان، تا رفتار گردشگران به منظور بهبود شرایط و تجربیات فعلی هم در مرحله قبل از سفر، هم در حین سفر و هم پس از سفر ارزیابی و به‌منزله ابزاری برای سنجش معتبر در نظر گرفته شود (Liu et al., 2019).

داده‌های تعاملی کاربر

تراکنش‌های گردشگران را می‌توان منبعی برای گردآوری داده‌های کلان در نظر گرفت. به این صورت که در مقاصدی که کارت‌های اعتباری و پرداخت‌های دیجیتال استفاده می‌شوند، می‌توان به راحتی رفتار خرید محصولات و خدمات گردشگری کاربران را رصد و علاوه بر آن، این آمار را بین گردشگران خارجی و داخلی تفکیک کرد (Palop et al., 2019). منبع دوم داده‌های کلان تعاملی که گردشگران ناخودآگاه تولید می‌کنند داده‌های تلفن همراه و رومینگ و همچنین موقعیت مکانی گردشگران در طول سفر است؛ مثلاً از این داده‌ها می‌توان برای ایجاد یک الگوی حرکتی از نقل مکان گردشگران در طول سفر بهره جست (Raun et al., 2016).

داده‌های غیرکاربر

عکس‌های ماهواره‌ای گرفته‌شده از سطوح زمین در مقاصد گردشگری مختلف و تغییرات اقلیمی و زیست‌بوم در محیط آن مقاصد، که اطلاعات زمانی و مکانی متنوعی را خلق می‌کند، از جنس این داده‌ها هستند که گردشگران در تولید این داده‌ها نقشی ندارند (Tussyadiah & Miller, 2019).

رویکردها

در این بخش از تحقیق، به سؤال دوم این پژوهش درباره رویکردهای اتخاذشده در راستای بهره‌گیری از کلان‌داده‌ها در مطالعات مرتبط با گردشگری خواهیم پرداخت و ابزار و روش‌های استفاده‌شده در برخی از پژوهش‌های انتخابی را ارزیابی خواهیم کرد. جدول ۴ درباره هر گروه از پژوهش‌های انجام‌شده و تعداد آن‌ها براساس رویکردهای به‌کارگرفته شده توضیحی را بیان می‌کند.

1. User Generated Content

جدول ۴: رویکردهای به کار گرفته شده در مقالات

منبع گردآوری کلان داده‌ها	رویکرد تحلیل آماری	رویکرد هوش مصنوعی و یادگیری ماشین
محتوای تولیدشده به دست کاربر	۴۳	۲۶
داده‌های تعاملی کاربر	۲	۲
داده‌های غیرکاربر	۱۵	۶

نوع و دسته‌بندی کلان داده‌ها

در این قسمت از پژوهش، به یافتن پاسخ سؤال سوم این تحقیق خواهیم پرداخت. طبق یافته‌ها و بررسی مقالات انتخابی از نقطه نظر تحلیلی می‌توان چهار دسته‌بندی مختلف برای کلان داده‌ها در نظر گرفت که عبارت‌اند از: ۱. تحلیل متنی که با داده‌های متنی و کاراکتری و حتی رشته‌ای از کلمات سروکار دارد؛ ۲. تحلیل عکس و تصویر؛ ۳. تحلیل صوت با استفاده از یادگیری ماشین که به استخراج حس از صدا کمک می‌کند؛ ۴. تحلیل فیلم و ویدئو که آخرین فناوری در تحلیل کلان داده‌هاست. از نقطه نظر هدف تحلیل کلان داده‌ها نیز چهار دسته‌بندی وجود دارد: (۱) تحلیل پیش‌گویانه که از الگوریتم‌های یادگیری ماشین برای توسعه مدل پیش‌بینی خود بهره می‌گیرد. در این مدل، پیش‌بینی براساس سوابق و تاریخچه تحلیل داده‌های قبلی ارائه می‌شود؛ (۲) تحلیل توصیفی که با تجزیه و تحلیل وضعیت حال یا گذشته نقص‌ها را شناسایی می‌کند. در این روش، اتفاقات مختلف رصد شده و گزارشی بر آن اساس ارائه خواهد شد؛ (۳) تحلیل تجویزی که در آن، وضعیت‌های مختلف بررسی می‌شوند و بهترین راه‌حل تجویز می‌شود؛ (۴) تحلیل تشخیصی که برای رسیدن به بهترین نتایج مداوم به کار گرفته می‌شود تا بهبود در فرایند نهایی شکل گیرد. در این روش، از داده‌کاوی و یافتن هم‌بستگی بین داده‌ها استفاده می‌شود (Rahman & Reza, 2022). در دو جدول ۵ و ۶، مقالات انتخابی در هر یک از دسته‌بندی‌های ذکر شده مشخص شده‌اند.

جدول ۵: تعداد مقالات انتخابی براساس نوع تحلیل کلان داده‌ها

نوع تحلیل کلان داده‌ها	تعداد مقالات
تحلیل متنی	۱۴
تحلیل عکس و تصویر	۳
تحلیل صوت	۱
تحلیل فیلم و ویدئو	۲

تجزیه و تحلیل یافته‌ها

با بررسی نتایج به‌دست‌آمده از سؤالات تحقیق مشخص شد که اولاً استفاده از تحلیل کلان‌داده‌ها در مقالات علمی رشد چشمگیری داشته و ثانیاً در بیشتر پژوهش‌های صورت‌گرفته حول موضوع گردشگری، با استفاده از تحلیل کلان‌داده‌ها، بیشترین منبع گردآوری کلان‌داده‌ها به داده‌های تولیدشده به‌دست کاربران با رویکرد تحلیل آماری و پس از آن تحلیل با هوش مصنوعی و یادگیری ماشین تعلق دارد. این به آن معنی است که شبکه‌های اجتماعی، که نقش انتشار داده‌های تولیدشده به‌دست کاربر را برعهده دارند، نقش پررنگ‌تری در تحقیقات علمی بازی می‌کنند و روش سنتی جمع‌آوری پرسش‌نامه جای خود را به بررسی نظرات واقعی کاربران در شبکه‌های اجتماعی حول موضوعی خاص خواهد داد. بنابراین در پاسخ به سؤال اول این پژوهش، که پژوهشگر اغلب از چه منبعی از کلان‌داده‌ها در مطالعات گردشگری استفاده می‌کند، باید گفت یافته‌های تحقیق نشان‌دهنده‌ی علاقه پژوهشگران به استفاده از کلان‌داده‌های تولیدشده به‌دست کاربران در شبکه‌های اجتماعی است؛ چراکه دسترسی به آن‌ها آسان‌تر و تحلیل آن‌ها توسط هوش مصنوعی چالش کمتری دارد. در پاسخ به سؤال دوم این پژوهش که چه رویکردها و ابزارهای تحلیل کلان‌داده‌های استخراج‌شده استفاده می‌شود، یافته‌های پژوهش حاکی از پیروگی رویکرد تحلیل آماری بر محتوای تولیدشده به‌دست کاربر است؛ چراکه با توجه به پاسخ سؤال اول، اغلب کلان‌داده‌های تولیدشده به‌دست کاربران به شکل متن و نظر در شبکه‌های اجتماعی وجود دارد و گردآوری این نظرات می‌تواند رفته‌رفته جای پرسش‌نامه‌های پژوهش‌های علمی را بگیرد. در پاسخ به سؤال سوم این پژوهش (نوع و دسته‌بندی کلان‌داده‌هایی که این ابزار و روش‌های تحلیل را پشتیبانی می‌کنند) باید گفت یافته‌های پژوهش نشان می‌دهند تحلیل متنی داده‌های گردآوری‌شده با هدف تحلیل پیش‌گویانه بیشترین نمونه استفاده را در مقالات انتخاب‌شده داشته است؛ از این رو یکی از پرکاربردترین و پرچالش‌ترین جنبه‌ها در مدیریت مقاصد گردشگری، یعنی مدیریت حجم بالای ورودی در فضول اوج، تا حدودی پیش‌بینی‌پذیر خواهد شد. بررسی ابرکلمات این موضوع را تأیید می‌کند که اغلب کاربرد این مقالات در راستای پیش‌بینی تقاضای گردشگری یک مقصد و رسیدن به الگویی برای پیش‌بینی و در گام دوم بررسی شرایط توسعه پایدار مقاصد گردشگری با بهره‌گیری از تحلیل کلان‌داده‌هاست. علاوه بر این، نتایج نشان می‌دهند که در غالب تحقیقاتی که پژوهشگر میل به استفاده از کلان‌داده‌ها را در آن دارد، بیشتر از تحلیل داده‌های کمی استفاده شده و کمتر به تحلیل کیفی روی آورده شده است.

نتیجه‌گیری



هدف این مقاله بررسی میزان بهره‌گیری از کلان‌داده در تحقیقات علمی در زمینه گردشگری منتشرشده از سال ۲۰۱۶ تا ۲۰۲۲ بود که بر این اساس، سؤالاتی طراحی شد و سعی شد در راستای رسیدن به هدف به این سه سؤال پاسخ داده شود. نتایج به‌دست‌آمده از تحلیل یافته‌های این پژوهش مبین آن است که بیشتر پژوهشگرانی که در حوزه گردشگری با استفاده از تحلیل کلان‌داده‌ها پژوهش کرده‌اند علاقه بیشتری به استفاده از کلان‌داده‌های تولیدشده به‌دست کاربران در شکل داده‌های متنی و نظرات دارند. همچنین اغلب هدف استفاده از آن‌ها برای پیش‌بینی و پیش‌بینی روند یا حجم ورودی یا تقاضای فصلی در صنعت گردشگری بوده است. شاید دلیل اصلی این امر سهولت تحلیل داده‌هایی است که به شکل متن در شبکه‌های اجتماعی به‌دست کاربران تولید شده است. با وجود این، کلان‌داده‌ها انواع و اشکال متفاوتی دارند که بسته به نیاز محقق می‌توان از آن‌ها بهره گرفت؛ اما شاید اطلاع‌نداشتن پژوهشگران از سایر روش‌های تحلیل کلان‌داده‌ها آن‌ها را به آسان‌ترین روش که تحلیل داده‌های کلان از جنس کیفی است نزدیک‌تر می‌کند.

محدودیت‌ها: گرچه استفاده از کلان‌داده‌ها در تحقیقات آکادمیک پیشرفت چشمگیری داشته است، همچنان چالش‌هایی درباره کیفیت داده‌ها، هزینه داده‌ها و نگرانی‌های حفظ حریم خصوصی کاربران وجود دارد. علاوه بر این، چندین موضوع درباره قابلیت اطمینان به داده‌های تولیدشده به‌دست کاربر مانند جعلی بودن نظرات درباره یک جاذبه گردشگری برای رسیدن به تعداد کاربر بیشتر وجود دارد. همچنین با توجه به نگرانی‌های حفظ حریم خصوصی، رومینگ تلفن همراه و داده‌های ترانکس کاربران به‌طور گسترده در تحقیقات گردشگری استفاده نشده‌اند، اما

با وجود مزایا، دسترسی به آن‌ها پرهزینه و دشوار است؛ چراکه معمولاً شبکه‌های اپراتور تلفن همراه، همچنین ارائه‌دهندگان راهکارهای تجاری مایل به اشتراک‌گذاری اطلاعات خصوصی کاربران خود نیستند. برای غلبه بر این چالش‌ها، همکاری متقابل بین دانشگاه، دولت و صنایع ضروری می‌شود. این همکاری نه تنها در دسترس بودن داده‌ها و کاهش هزینه داده‌ها برای تحقیقات گردشگری با استفاده از داده‌های بزرگ را تضمین می‌کند، بلکه در ازای آن، به مسائل عملی نیز می‌پردازد. راه دیگری که برای غلبه بر این چالش به ذهن می‌رسد استفاده از داده‌های شبکه‌های اجتماعی است؛ زیرا به افراد اجازه می‌دهد دیدگاه‌های خود را دربارهٔ رویداد، محصول، ابزار و موضوعات دیگر ارائه دهند و دسترسی به این داده‌ها آسان‌تر است و باید در تحقیقات آتی از مزایای این حجم عظیم داده استفاده شود.

پیشنهادها برای تحقیقات آتی: با توجه به ابعاد و اشکال متنوع و گسترده کلان‌داده‌ها و همچنین روش‌ها و ابزارهای گوناگون که روزبه‌روز نیز به تنوع آن‌ها افزوده می‌شود، لزوم بهره‌گیری از این تنوع در پژوهش‌های علمی حس می‌شود، اما شکافی که در این پژوهش مورد توجه قرار گرفته است نبود اطلاع کافی پژوهشگران از این ابزار و تنوع گوناگون نوع داده‌هایی است که به‌صورت روزافزون در عرصه فناوری اطلاعات در حجم بالا تولید می‌شود. یکی از مهم‌ترین قسمت‌ها پس از جمع‌آوری کلان‌داده‌ها، قسمت تجزیه و تحلیل این داده‌های کلان و سپس تصمیم‌گیری داده‌محور براساس نتایج این تحلیل‌هاست. از این رو پیشنهاد می‌شود پژوهشگران قبل از انجام تحقیق و انتخاب ابزار و روش، به شناخت و اطلاع از این روش‌ها بپردازند و پس از تسلط کافی در زمینه تحلیل کلان‌داده‌ها، بهینه‌ترین روش را بر مبنای نیاز خود انتخاب کنند تا بیشترین بهره را از پتانسیل موجود در استفاده از کلان‌داده‌ها در حوزه گردشگری کسب کنند.

ORCID:

Jafar Ahangaran	 http://orcid.org/0009-0006-0387-475X
Abbas Sadeghnia	 http://orcid.org/0000-0002-2314-4678

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

- Álvarez-García, J., Durán-Sánchez, A., del Río-Rama, M. D. L. C., & Simonetti, B. (2023). Big data and tourism research: measuring research impact. *Quality & Quantity*, 57(Suppl 3), 271-292. <http://orcid.org/10.1007/s11135-020-01044-z>.
- Amudhavel, J., Padmapriya, V., Gowri, V., Lakshmi Priya, K., Kumar, K. P., & Thiagarajan, B. (2015, March). Perspectives, motivations and implications of big data analytics. In *Proceedings of the 2015 International Conference on Advanced Research in Computer Science Engineering & Technology (ICARCSET 2015)* (pp. 1-5). <http://orcid.org/10.1145/2743065.2743099>
- Ayele, W. Y., & Juell-Skielse, G. (2017). Social media analytics and internet of things: Survey. In *Proceedings of the 1st International Conference on Internet of Things and Machine Learning* (pp. 1-11). DOI: 10.1145/3109761.3158379
- Banaeianjahromi, N., & Smolander, K. (2016). What do we know about the role of enterprise architecture in enterprise integration? A systematic mapping study. *Journal of Enterprise Information Management*, 29(1), 140-164. <http://orcid.org/10.1108/JEIM-12-2014-0114>
- Boyd, D. M., & Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. *Journal of computer mediated Communication*, 13(1), 210-230. <https://doi.org/10.1111/j.1083-6101.2007.00393.x>
- Chen, L. Z. (2017). Big data for smart tourism: A multi-perspective analysis. *Tourism Management*, 1(3), 134-143. <https://doi.org/10.52255/smarttourism.2021.1.3.2>
- Cong, L., Wu, B., Morrison, A. M., Shu, H., & Wang, M. (2014). Analysis of wildlife tourism experiences with endangered species: An exploratory study of encounters with giant pandas in Chengdu, China. *Tourism management*, 40, 300-310. <http://orcid.org/10.1016/j.tourman.2013.07.005>
- Dawit Negussie, F, Fikeresiyon, Fekede., Gobezu, G. (2023). SOCIAL MEDIA AND TOURISM: A SYSTEMATIC LITERATURE REVIEW. *GAP iNTERDISCIPLINARITIES - A GLOBAL JOURNAL OF INTERDISCIPLINARY STUDIES* ,6(4), 40-46. <http://orcid.org/10.47968/gapin.640001>
- Di Francesco, P., Malavolta, I., & Lago, P. (2017). Research on architecting microservices: Trends, focus, and potential for industrial adoption. In *2017 IEEE International conference on software architecture (ICSA)* (pp. 21-30). IEEE. <http://orcid.org/10.1109/ICSA.2017.24>
- Dwivedi, SH., Tripathi, H., Kumar, V., & Vishwakarma, S. (2023). A STUDY ON IMPACT OF SOCIAL MEDIA IN PROMOTING TOURISM. *International Journal of Development Research*, 13, 14-24. <http://orcid.org/10.56556/jssms.v3i1.685>
- George, G., Osinga, E. C., Lavie, D., & Scott, B. A. (2016). Big data and data science methods for management research. *Academy of Management Journal*, 59(5), 1493-1507. DOI: <https://doi.org/10.5465/amj.2016.4005>

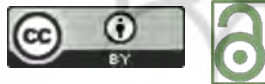
- Gretzel, U., Sigala, M., Xiang, Z., & Koo, C. (2015). Smart tourism: foundations and developments. *Electronic markets*, 25, 179-188. <http://orcid.org/10.1007/s12525-015-0196-8>
- Kaushal, V., & Srivastava, S. (2021). Hospitality and tourism industry amid COVID-19 pandemic: Perspectives on challenges and learnings from India. *International journal of hospitality management*, 92, 102707. <http://orcid.org/10.1016/j.ijhm.2020.102707>
- Kitchenham, B., Budgen, D. (2011). Using mapping studies as the basis for further research- A participant-observer case study. *Information and Software Technology*, 53(6), 638-651. <http://orcid.org/10.1016/j.infsof.2010.12.011>
- Li, X. W. (2018). The role of big data in promoting sustainable tourism development. *Journal of Sustainable Tourism*, 12(11), 265-281. <http://orcid.org/10.1080/13683500.2021.1974358>
- Liu, Y., Huang, K., Bao, J., & Chen, K. (2019). Listen to the voices from home: An analysis of Chinese tourists' sentiments regarding Australian destinations. *Tourism Management*, 71, 337-347. <http://orcid.org/10.1016/j.tourman.2018.10.004>
- Neuhofer, B., Buhalis, D., & Ladkin, A. (2015). Technology as a catalyst of change: Enablers and barriers of the tourist experience and their consequences. In *Information and Communication Technologies in Tourism 2015: Proceedings of the International Conference in Lugano, Switzerland, February 3-6, 2015* (pp. 789-802). Springer International Publishing. http://orcid.org/10.1007/978-3-319-14343-9_57
- Park, D., Kim, W. G., & Choi, S. (2019). Application of social media analytics in tourism crisis communication. *Current Issues in Tourism*, 22(15), 1810-1824. DOI:10.1080/13683500.2018.1504900
- Petersen, K., Feldt, R., Mujtaba, S., & Mattsson, M. (2008, June). Systematic mapping studies in software engineering. In *12th International Conference on Evaluation and Assessment in Software Engineering (EASE) 12* (pp. 1-10). <http://orcid.org/10.14236/ewic/EASE2008.8>
- Rahman, M. S., & Reza, H. (2022). A systematic review towards big data analytics in social media. *Big Data Mining and Analytics*, 5(3), 228-244. <http://orcid.org/10.26599/BDMA.2022.9020009>
- Raun, J., Ahas, R., & Tiru, M. (2016). Measuring tourism destinations using mobile tracking data. *Tourism Management*, 57, 202-212. <http://orcid.org/10.1016/j.tourman.2016.06.006>
- Ritter, T., & Pedersen, C. L. (2020). Digitization capability and the digitalization of business models in business-to-business firms: Past, present, and future. *Industrial Marketing Management*, 86, 180-190. <http://orcid.org/10.1016/j.indmarman.2019.11.019>
- Romero Palop, J. D. D., Murillo Arias, J., Bodas-Sagi, D. J., & Valero Lapaz, H. (2019). Determining the usual environment of cardholders as a key factor to measure the evolution of domestic tourism. *Information Technology & Tourism*, 21(1), 23-43. <http://orcid.org/10.1007/s40558-018-0130-y>

- Schmerl, B., Kazman, R., Ali, N., Grundy, J., & Mistrik, I. (2017). Managing trade-offs in adaptable software architectures. In *Managing trade-offs in adaptable software architectures* (pp. 1-13). Morgan Kaufmann. <http://orcid.org/10.1016/B978-0-12-802855-1.00011-3>
- Sharma, P., Leung, T., Kingshott, R., Davicik, N., Cardinal, S. (2020). Managing uncertainty during a global pandemic: An international business perspective. *Journal of Business Research*, 116, 188-192. <http://orcid.org/10.1016/j.jbusres.2020.05.026>
- Sigala, M. (2017). Tourism and COVID-19: Impacts and implications for advancing and resetting industry and research. *Journal of Business Research*, 117, 312-321. doi.org/10.1016/j.jbusres.2020.06.015
- Soluk, J., Kammerlander, N., & De Massis, A. (2021). Exogenous shocks and the adaptive capacity of family firms. *exploring behavioral changes and digital technologies in the COVID-19 pandemic*, 51(4), 364-380. <http://orcid.org/10.1111/radm.12471>
- Theodorakopoulos, L., & Halkiopolus, C. (2023). Applying Big Data Technologies in Tourism Industry: A Conceptual Analysis. *International Conference of the International Association of Cultural and Digital Tourism*. SpringerLink. <http://orcid.org/10.1088/1742-6596/1648/4/042101>
- Turkpenova, J., & Abitova, G. A. (2023). Challenges and prospects in big data analytics: a comprehensive review of developments, hurdles, and future research directions. *Bulletin of Shakarim University Technical Sciences*, 5, 11-7. [https://doi.org/10.53360/2788-7995-2023-3\(11\)-7](https://doi.org/10.53360/2788-7995-2023-3(11)-7)
- Tussyadiah, I., & Miller, G. (2019). Perceived impacts of artificial intelligence and responses to positive behaviour change intervention. In *Information and Communication Technologies in Tourism 2019: Proceedings of the International Conference in Nicosia, Cyprus, January 30–February 1, 2019* (pp. 359-370). Springer International Publishing. http://orcid.org/10.1007/978-3-030-05940-8_28
- UNWTO. (2021). BIG DATA FOR BETTER TOURISM POLICY, MANAGEMENT, AND SUSTAINABLE RECOVERY FROM COVID-19. *Co-publication of the Asian Development Bank and the World Tourism Organization*. <http://orcid.org/10.22617/SPR210438-2>
- Van der Zee, E., & Bertocchi, D. (2018). Finding patterns in urban tourist behaviour: A social network analysis approach based on TripAdvisor reviews. *Information Technology & Tourism*, 20(1-4), 153-180. <http://orcid.org/10.1007/s40558-018-0128-5>
- Wang, D. L. (2016). Information technology and tourist experience: Cultural differences in technology acceptance. *Tourism Management*, 53, 49-61. <http://orcid.org/10.3390/su151813349>
- Wendler, R. (2012). The maturity of maturity model research: A systematic mapping study. *Information and software technology*, 54(12), 1317-1339 <http://orcid.org/10.1016/j.infsof.2012.07.007>

- Witten, I. H., Frank, E., & Hall, M. A. (2011). Data mining: Practical machine learning tools and techniques.- The Morgan Kaufmann ser. in data management systems. doi:10.1186/1475-925X-5-51
- Xiang, Z., Du, Q., Ma, Y., & Fan, W. (2017). A comparative analysis of major online review platforms: Implications for social media analytics in hospitality and tourism. *Tourism Management*, 58, 51-65.. <http://orcid.org/10.1016/j.tourman.2016.10.001>
- Xu, F., Nash, N., & Whitmarsh, L. (2020). Big data or small data? A methodological review of sustainable tourism. *Journal of Sustainable Tourism*, 28(2), 144-163.. DOI: 10.1080/09669582.2019.1631318
- Youssef, A. B., Boubaker, S., Dedaj, B., & Carabregu-Vokshi, M. (2021). Digitalization of the economy and entrepreneurship intention. *Technological Forecasting and Social Change*, 164, 120043.. <http://orcid.org/10.1016/j.techfore.2020.120043>

COPYRIGHTS

©2024 by the authors. Published by University of Science and Culture. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 International (CC BY 4.0) <https://creativecommons.org/licenses/by/4.0/>



پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی