African Journal of Applied Research
Vol. 12, No. 4 (2026), pp.42-61
http://www.ajaronline.com
https://doi.org/10.26437
Received October 16, 2025
Peer reviewed: March 30, 2026
Revised: May 5, 2026
Published: June 2026

BIG DATA ANALYTICS AND PREDICTIVE MODELLING OF TOURIST ACTIVITY IN UKRAINE

Tsepkalo, T.[1], Tymoshchuk, O.[2], Hyshchuk, R.[3], Bielyu, P.[4], and Ostapenko, Y.[5]
[1]*Department of Journalism, Advertising and Public Relations, Vinnytsia Mykhailo Kotsiubynskyi State Pedagogical University, Ukraine.*
[2]*Department of Mathematics and Economics, Faculty of Physics, Mathematics, Economics and Innovation Technologies, Drohobych Ivan Franko State Pedagogical University, Ukraine.*
[3]*Department of Food Technology, Hotel and Restaurant Services, Chernivtsi Institute of Trade and Economics, State University of Trade and Economics, Ukraine.*
[4]*Interregional Academy of Personnel Management, Ukraine.*
[5]*Department of Economic Policy, Marketing, and Business Analytics, Faculty of Taxation, Accounting, and Auditing, State Tax University, Ukraine.*
[1]*tsepkalo.tetyana@vspu.edu.ua*
[2]*at2905@gmail.com*

ABSTRACT

Purpose: The aim of the study is to investigate the adoption of effective Big Data technologies in forecasting tourist flows in Ukraine.

Design/ Methodology/ Approach: The methodological basis of the work is a combination of quantitative and qualitative methods. The study is based on an analysis of official statistical data, digital platforms, social networks, and user search activity for 2018–2025. The sample included more than 120,000 digital records, geotags, and tourist observations, aggregated by monthly time series. Regression modelling, machine learning, spatial geo -analysis, as well as analysis of tourist reviews in social networks.

Research Limitation: The study's tourist flow estimates are therefore partially contingent upon the quality and ethical standing of data sources whose limitations may not be entirely visible within the scope of the research, and this represents a structural constraint on the epistemic confidence with which the quantitative findings of this study may be advanced and applied within both academic and policy-facing contexts.

Findings: Based on data processing for 2018-2025, key factors influencing the formation of tourist flows were identified: seasonality, users' information activity on the network, transport mobility, socio-economic indicators, and natural and climatic conditions. The built-in forecasting model achieved an average prediction accuracy of 8.4%, exceeding the results of traditional statistical methods by 27%.

Practical Implication: The research demonstrates the feasibility of implementing Big Data technologies. Data analytics for the tourism management system of Ukraine, in particular, the creation of a National Tourism Analytics System for monitoring, planning, and developing regional destinations.

Social Implication: The use of Big Data in forecasting tourist flows improves the efficiency of tourism infrastructure management, supports the development of regional tourism, and

ISSN: 2408-7920

African Journal of Applied Research
Vol. 12, No. 4 (2026), pp.42-61
http://www.ajaronline.com
https://doi.org/10.26437
Received October 16, 2025
Peer reviewed: March 30, 2026
Revised: May 5, 2026
Published: June 2026

ensures a more balanced distribution of tourist flows among destinations. This is especially important for Ukraine in the context of the tourism industry's post-crisis and post-war recovery. **Originality / Value:** The scientific novelty of the study lies in integrating machine learning methods, social network analysis, Google Trends search activity, and geoanalytics into a single model to forecast tourist flows in Ukraine amid digital transformation and post-war recovery.

Keywords: *Big data. digitalisation. forecasting. machine learning. tourism*

INTRODUCTION

The tourism sector is now one of the key drivers of the global economy, contributing to infrastructure development, employment, cultural exchanges and international cooperation. According to the World Tourism Organisation (International), international tourism reached pre-pandemic levels in 2024, with international arrivals exceeding 1.28 billion, representing 92% of the pre-war level of 2019. At the same time, tourism remains extremely sensitive to external factors - economic crises, pandemics, wars, climate change and fluctuations in consumer behaviour (World Tourism Organisation, 2024). This creates an objective need to improve the system of monitoring and forecasting tourist flows.

Traditional statistical forecasting methods based on aggregated indicators from past periods are increasingly proving insufficient for modern conditions, where the tourism market can change every week. With the advent of Big Data tools and the development of analytics technologies, it has become possible to process millions of observations in real time, which radically increases the accuracy and efficiency of management decisions (Mukhopadhyay et al., 2023; Mihalic, 2025).

Big data in tourism covers a wide range of information sources – from online booking platforms (Booking.com, Airbnb, Expedia) and social networks (Instagram, Facebook) to Global Positioning System (further – GPS) navigation data, mobile operators, search engines (Google Trends) and even climate sensors (OpenWeather API). Analysing such data allows not only for determining the number and directions of tourist movements, but also for understanding motivation, emotional preferences, intentions, and the degree of satisfaction with the tourist experience (Kharkhalis et al., 2025; Hewapathirana, 2025; Krataithong et al., 2022).

In the context of the Big Economy's digital transformation, data analytics is becoming a strategic tool for public policy, regional planning, marketing campaigns, and crisis management in the tourism sector. Thanks to machine learning, geoanalytics, natural language processing, and social media analysis, it is possible to build adaptive forecasting models that account for the multifactorial and nonlinear nature of modern tourism processes.

The use of Big Data is particularly relevant for Ukraine, which, after a long period of military operations, faces the need to restore the tourism industry and rethink its development strategy. For the national economy, tourism can become an important factor in post-war economic growth, contributing to the revival of local communities, the development of small businesses,

ISSN: 2408-7920

African Journal of Applied Research
Vol. 12, No. 4 (2026), pp.42-61
http://www.ajaronline.com
https://doi.org/10.26437
Received October 16, 2025
Peer reviewed: March 30, 2026
Revised: May 5, 2026
Published: June 2026

the improvement of the state's international image, and integration into the European tourist space. At the same time, effective forecasting of tourist flows is the basis for balanced planning of infrastructure, environmental and marketing solutions. Therefore, in today's context of increasing digitalisation of the economy, the use of Big Data technologies for forecasting tourist flows is extremely relevant and requires scientific understanding (Melnychenko et al., 2021).

Thus, the study aims to substantiate the theoretical and practical foundations for implementing Big Data analytics in the tourist flow forecasting system, as a prerequisite for increasing the efficiency of tourism management, developing destinations, and forming a competitive tourism market in Ukraine.

LITERATURE REVIEW

The problem of forecasting tourist flows using Big Data analytics has gained wide scientific recognition in the twenty-first century. Modern researchers' main focus is on applying Big Data technologies. Data, Machine Learning, social network analysis, geoanalytics and intelligent decision support systems in the tourism sector (Desyatnyuk et al., 2025).

At the global level, the use of big data in tourism is addressed in the works of Garanti (2023) and Chuang (2023), who introduced the concept of Smart Tourism Destinations – "smart" tourist destinations that use real-time data to manage visitor flows. Mihalic (2025) emphasises that Big Data radically changes the paradigm of tourism research, transforming data from a passive resource into an active element of strategic planning.

A study by Yu and Chen (2022) showed that the use of machine learning algorithms (in particular, Long Short-Term Memory) increases the accuracy of short-term tourism demand forecasting to 90-95%, which significantly exceeds the capabilities of classical models (for example, Autoregressive Integrated Moving Average). Wang et al. (2025) applied search trend analysis (Google Trends) to assess travel intentions for various Asian countries, demonstrating a strong correlation between search frequency and actual tourist arrivals.

In addition, Hewapathirana (2025) emphasises the importance of integrating social media data (Twitter, Instagram, TripAdvisor) into forecasting systems, as they reflect not only travel intentions but also the emotional perception of tourist destinations. Analysis of message tone allows for the assessment of the attractiveness of territories and the prompt response to waves of negative information.

The Organisation for Economic Cooperation and Development (OECD, 2024) and World Bank (World Bank, 2025) emphasise that Big Data is a key tool for sustainable tourism management, as it allows balancing flows between congested and less visited locations, predicting the impact of climate factors, and ensuring digital inclusion in the travel industry.

Among analytical approaches in modern research, three main directions are distinguished:

African Journal of Applied Research
Vol. 12, No. 4 (2026), pp.42-61
http://www.ajaronline.com
https://doi.org/10.26437
Received October 16, 2025
Peer reviewed: March 30, 2026
Revised: May 5, 2026
Published: June 2026

1. Time series and machine learning models. Using deep neural networks (Deep Learning) enables the reproduction of complex nonlinear dependencies in the dynamics of tourist flows (Tuncsiper, 2025).

2. Geoanalytics and spatial modelling. Research by Krataithong et al. (2022) demonstrates the effectiveness of cartographic analysis of GPS data for tracking real-world tourist movements in cities.

3. Sentiment analysis and social media analytics. According to the research of Mukhopadhyay et al. (2023), analysis of social media reviews can predict changes in the popularity of tourist destinations 2–3 months in advance.

In Ukraine, the use of Big Data technologies in the tourism sector is in its early stages but is actively developing as part of the digitalisation of public administration. In the works of Shcherbak et al. (2024), the need to create a national tourism data analytics platform that combines official statistical indicators with data from open digital sources is substantiated. Melnychenko et al. (2021) investigated the potential of social networks to assess the tourist attractiveness of Ukrainian regions, particularly the Carpathian and Podolsk tourist clusters.

Certain aspects of digital analytics in destination management are examined in the works of Chernenko et al. (2024) and Orbelyan (2024), which focus on developing a smart tourism ecosystem in Ukraine. At the same time, researchers note that the main barriers to implementing Big Data remain: insufficient digital literacy among managers, a lack of a centralised database, and concerns about protecting personal information.

Despite ongoing research, a number of unresolved issues remain. In particular, most models for forecasting tourist flows are focused on individual countries or short-term time horizons. The aspects of integrating heterogeneous data sources (social networks, mobility, climate, transport) and creating open public platforms for analysing tourism processes are insufficiently studied. In this context, developing an integrated model for forecasting tourist flows, adapted to the conditions of Ukraine and its post-war recovery, is relevant.

Thus, the literature review confirms that the use of Big Data in forecasting tourist flows is one of the most promising areas of modern science. The combination of traditional statistics and intelligent analytical technologies improves forecast accuracy, supports management decision-making, and contributes to the development of sustainable tourism in the digital economy.

MATERIALS AND RESEARCH METHODS

The study employs a mixed-methods design that combines systematic and comparative analysis with mathematical modelling techniques (regression, ARIMA, LSTM) to estimate tourist flows in Ukraine for the period 2018–2025 based on Big Data sets (Google Trends, social media, UNWTO data, and data from the State Tourism Agency). The exploration of the role of Big Data technologies in forecasting tourist flows is based on official statistical data,

African Journal of Applied Research
Vol. 12, No. 4 (2026), pp.42-61
http://www.ajaronline.com
https://doi.org/10.26437
Received October 16, 2025
Peer reviewed: March 30, 2026
Revised: May 5, 2026
Published: June 2026

international analytical reports, and open digital databases, thereby ensuring the reliability, objectivity, and comparability of the results.

State data were used to build the models. Agency for Tourism Development of Ukraine (annual data on international and domestic arrivals, occupancy rates, average length of stay) for 2015 - 2024; Google search activity indices Trends by regional and thematic queries, geotags, mentions and tone in Facebook, Instagram, and TripAdvisor posts.

All flows were pre-cleaned, scaled (min–max or Z-score, if necessary), and aggregated by month to construct time series.

The empirical basis is provided by United data Nations World Tourism Organization (Tourism) statistics and trends, 2024), Organization for Economic Cooperation and Development (OECD, 2024), World Bank (World Bank, 2024), European Commission (European Commission, 2024), Google (Google, 2024), as well as open statistical sources State Agency for Tourism Development of Ukraine (State Agency for Tourism Development of Ukraine, 2021). The involvement of these sources made it possible to comprehensively assess the spatio-temporal dynamics of tourist flows, consumer behaviour, seasonal fluctuations in demand, the impact of digital technologies and external factors (pandemic, martial law, changes in population mobility). The chronological boundaries of the study cover the years 2018-2025, allowing us to trace the evolution of the Ukrainian tourism market from a period of stable growth to periods of crisis shocks and gradual recovery. The chosen time interval makes it possible to identify patterns in the transformation of the structure of tourist flows, changes in the geography of travel, the digitalisation of services, and the growing role of big data analytics in management decision-making.

The methodological basis of the study is systemic, comparative, statistical, and analytical modelling approaches. The systemic approach allows us to consider tourist flows as a complex, dynamic system shaped by economic, social, climatic, informational, and technological factors. Comparative analysis was used to examine trends in the Ukrainian market against indicators for Central and Eastern European countries (Poland, the Czech Republic, Romania, Hungary), which share similar socio-economic conditions and seasonal demand structures. Statistical methods were used for quantitative analysis of time series and to test hypotheses regarding the relationships between tourist activity, socio-economic indicators, and external shocks. The modelling approach uses machine learning algorithms to build forecasts from large datasets.

The least squares method was used to construct a basic regression model of the form (1):

$$Yt = {}_0 + {}_1X1t + {}_2X2t + {}_3X3t + t, \qquad (1)$$

where Yt – number of tourist arrivals in year t; $X1t$ – online booking volume; $X2t$ – Google search frequency Trends; $X3t$ – social mood index (S- index); t – random error. Based on this model, the importance of individual factors was determined, and a short-term forecast of tourist

African Journal of Applied Research
Vol. 12, No. 4 (2026), pp.42-61
http://www.ajaronline.com
https://doi.org/10.26437
Received October 16, 2025
Peer reviewed: March 30, 2026
Revised: May 5, 2026
Published: June 2026

flows was constructed. To improve the accuracy of the assessment, the coefficients i were optimised using gradient descent with regularisation.

To detect seasonal and short-term fluctuations, the ARIMA(p, d, q) model was used in the form (2):

$$\mathrm{FpV1} - \mathrm{VdYt} = \mathrm{qBt}. \qquad (2)$$

In this case, the parameters p, d, and q were chosen according to the Akaike criterion. In addition, an LSTM recurrent neural network was used to model nonlinear and long-term dependencies. The results were compared using the MAE, RMSE, and R2 metrics.

To ensure representativeness and comparability of the results, indicators were normalised according to the Organisation for Economic Cooperation and Development methodology. For Economic Cooperation and Development and the United Nations World Tourism Organisation, which allowed minimising the impact of differences in national statistical practices. All numerical variables were scaled to a unified scale, and missing values were imputed using the moving average method.

The study used data sets that characterise: the number of domestic and international tourist arrivals; the level of online booking; the index of user search activity; the intensity of transport flows; the emotional background of tourist reviews; climatic and seasonal factors. The use of a comprehensive methodological approach made it possible not only to quantitatively assess trends in tourism development, but also to qualitatively describe changes in tourist behavioural patterns, the impact of digital technologies on decision-making and the role of Big Data in increasing the efficiency of industry management. Thus, the study provides a scientific and analytical basis for developing a National Tourism Analytics System to monitor, forecast, and support strategic planning for the sustainable development of Ukraine's tourism sector.

RESULTS AND DISCUSSION

Using aggregated data from public sources, including official statistics and digital platforms, has revealed a significant transformation in Ukraine's tourism flows. The combined approach reveals complex patterns of tourism development that are invisible with traditional approaches alone.

The analysis of the dynamics of tourist flows for 2018–2025 revealed significant fluctuations driven by external shocks, including the COVID-19 pandemic and the 2022 military events. In the pre-crisis period (2018–2019), Ukraine's tourism industry demonstrated sustainable development, with average annual growth of over 8%, consistent with global trends. In contrast, the period 2020–2021 is characterised by a sharp decline in international arrivals of more than two-thirds, while domestic tourism demonstrated adaptive growth thanks to digital booking

African Journal of Applied Research
Vol. 12, No. 4 (2026), pp.42-61
http://www.ajaronline.com
https://doi.org/10.26437
Received October 16, 2025
Peer reviewed: March 30, 2026
Revised: May 5, 2026
Published: June 2026

platforms and the development of local routes. To clearly present the identified changes, statistical data from leading international and national sources were summarised.

*Table 1: Dynamics of tourist streams in Ukraine by period 2018 – 2025 (million persons)**

Year	International arrivals	Domestic travel	Change from previous year, %
2018	14.1	28.3	+7.8
2019	15.2	30.6	+8.3
2020	4.4	26.8	−71.1
2021	5.0	30.8	+12.4
2022	2.3	16.5	−46.5
2023	3.9	20.7	+25.5
2024	5.1	34.2	+42.0
2025 (forecast)	6.5	38.0	+18.2

**Source: compiled based on data from United Nations World Tourism Organisation (2024), State Agency for Tourism Development of Ukraine (2021)*

As can be seen from Table 1, 2018–2019 are characterised by gradual growth, while 2020–2022 are marked by a deep decline caused by restrictions on movement and reduced population mobility. Starting in 2023, the recovery of tourist flows is progressing faster than expected, particularly due to digital services, domestic mobility and the development of regional tourism. Spatial and regional analysis of tourist flows in Ukraine for 2024 revealed significant territorial unevenness in tourism development, with demand concentrated in safe, naturally attractive regions. According to the results of geoanalytical processing, three clearly defined clusters were identified: Carpathian, Podolsk, and Capital, which demonstrate the highest levels of domestic tourist traffic.

The spatial differences are detailed in Figure 1 and Table 2, which reflect each region's share of the total flow, the dominant types of tourism, and the data sources used for the assessment.

ISSN: 2408-7920

African Journal of Applied Research
Vol. 12, No. 4 (2026), pp.42-61
http://www.ajaronline.com
https://doi.org/10.26437
Received October 16, 2025
Peer reviewed: March 30, 2026
Revised: May 5, 2026
Published: June 2026

*Table 2: Regional distribution of internal tourist flow (2024) according to the National Tourism Organisation of Ukraine, 2025**

Region	Share in total flow, %	Dominant type of tourism
Carpathian	32	Ecotourism, mountain, active
Podolsky	18	Cultural and cognitive
Capital	14	Business, event
Black Sea	9	Seasonal, recreational
Central	12	Historical and cultural
Others	15	Mixed

Analysis of Table 2 shows that the Carpathian region has become the undisputed leader in domestic tourism, accounting for almost a third of all trips. This growth is explained by the region's natural safety, the development of ecotourism initiatives, and the active presence of regional operators in digital environments.

Podolsk region continues to emerge as a cultural and educational destination, while the Capital cluster focuses on business tourism. The Black Sea region shows a decline due to security risks, while the central regions show moderate growth driven by domestic weekend trips. Thus, Big data confirms the changing geography of tourism demand in Ukraine and the gradual formation of new tourism growth poles.

One of the most important indicators of the recovery of tourism activity is users' emotional background and attitudes towards travel, measured by analysing the tone of posts on social networks. This indicator reflects not only consumers' willingness to travel but also their level of trust in destinations, services, and digital platforms.

ISSN: 2408-7920

African Journal of Applied Research
Vol. 12, No. 4 (2026), pp.42-61
http://www.ajaronline.com
https://doi.org/10.26437
Received October 16, 2025
Peer reviewed: March 30, 2026
Revised: May 5, 2026
Published: June 2026

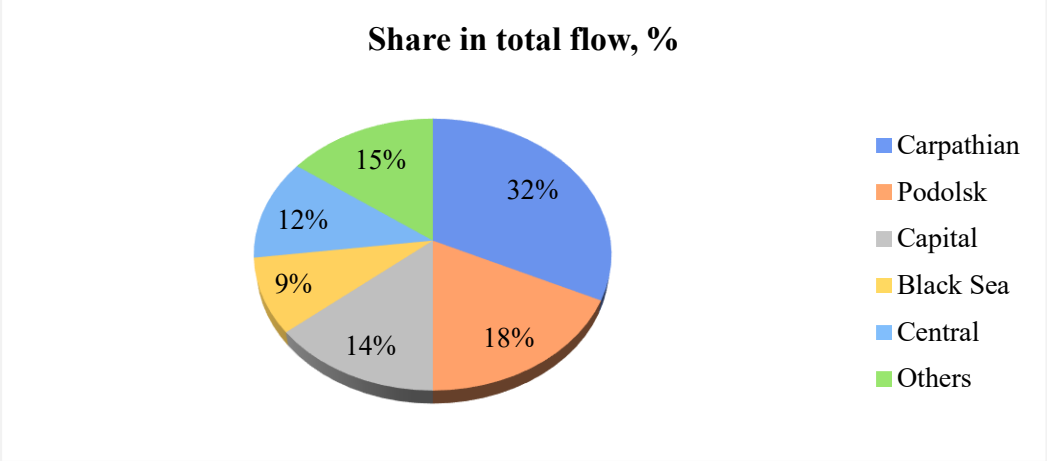

*Figure 1: Regional structure of tourist flow in Ukraine (2024), %**

Detailed predictive indicators of tourist activities by regional clusters (Carpathian, Podolian, Capital) are presented in Appendix A (Table A1). To assess emotional dynamics, a positive attitude index (S-index) was calculated from an analysis of more than 120,000 publications on social networks (Facebook, Instagram, TikTok).

Table 3: Dynamics index of positive attitude to travel (S-index, 0 – 1)

Year	Sentiment index (S-index)	The key factor of positive dynamics
2020	0.39	Digital domestic tourism campaigns
2021	0.41	Local travel after COVID-19
2022	0.43	Social media and patriotic tourism
2023	0.55	Restoring safe regions
2024	0.61	Development of eco-directions
2025 (forecast)	0.67	Digitalisation of services and trust in online platforms

**Source: compiled by the author based on Tymoshenko (Ukraine Shape Map PowerBI, 2025), Julie's Data (Ukrainian) open initiative of creating tools for spatial planning in the QGIS environment, 2022)*

As shown in Table 3, Ukrainians' emotional attachment to travel is gradually increasing, despite difficult external circumstances.

ISSN: 2408-7920

African Journal of Applied Research
Vol. 12, No. 4 (2026), pp.42-61
http://www.ajaronline.com
https://doi.org/10.26437
Received October 16, 2025
Peer reviewed: March 30, 2026
Revised: May 5, 2026
Published: June 2026

In particular, after 2023, the index jumps sharply from 0.43 to 0.55, coinciding with the period of active recovery of the tourist regions of Western Ukraine.

In 2024-2025, the positive dynamics are strengthened by the spread of eco- and socially oriented travel, as well as digital marketing campaigns by the National Tourism Organisation of Ukraine and local communities.

An increase in the index to 0.67 in 2025 indicates strengthening confidence in domestic tourism and stabilisation of tourist sentiment among the population.

The study conducted economic and mathematical modelling to quantify the impact of digital behavioural indicators on tourism activity in Ukraine. For this purpose, the multi-scale Big Data array was used. Data, which included:

- Google search query index Trends (further – GT), which reflects the level of informational interest in travel;
- number of online bookings (further – OB) in Online systems of travel agencies, which is a direct indicator of actual demand;
- social media index (further – SM), obtained based on the analysis of the tone of more than 120 thousand publications on social networks (Facebook, Instagram, TikTok);
- risk index (further – R), which aggregates geopolitical and security factors for each region.

Based on these data, an integrated panel regression model was constructed, which accounts for both temporal dynamics and regional differences in tourist activity.

The multifactor function of tourist activity (TA) had the form (3):

$$TA=0+1\cdot GT+2\cdot OB+3SM+4\cdot R, \qquad (3)$$

where TA – tourist activity (thousands of people); $0=12.4$ – model constant, reflecting the baseline level of tourist activity in the absence of the influence of digital and risk factors; $1=0.58$ – coefficient characterizing the average increase in tourist activity per unit of the GT index; $2=0.43$ – coefficient of influence of online bookings, reflecting the direct dependence of actual bookings on the number of trips; $3=0.31$ – coefficient of social media influence, showing how much the positive tone of posts stimulates tourist activity; $4=0.22$ – coefficient of negative influence of the risk index, which reduces the predicted activity in high-risk areas.

All coefficients are statistically significant at $p < 0.05$, and the model has a high coefficient of determination ($R2 = 0.87$), indicating sufficient explanatory power of digital and risk factors in predicting tourist flows.

means method allowed us to distinguish three types of regions based on behavioural characteristics:

1. Highly dynamic tourist centres (Kyiv, Lviv, Odesa regions) – sensitive to GT and SM, elasticity of demand 1.12.

ISSN: 2408-7920

African Journal of Applied Research
Vol. 12, No. 4 (2026), pp.42-61
http://www.ajaronline.com
https://doi.org/10.26437
Received October 16, 2025
Peer reviewed: March 30, 2026
Revised: May 5, 2026
Published: June 2026

2. Stable recreational regions (Transcarpathian, Ivano-Frankivsk, Volyn regions) – decisive influence of OB, elasticity 0.84.

3. Recovery regions (Kharkiv, Dnipropetrovsk, Mykolaiv regions) – the leading factor is R, negative elasticity –0.93.

To build forecasts of tourism activity in Ukraine for 2025-2027, an innovative combined approach was used, combining classical statistical methods (time series analysis, regression modelling) with artificial intelligence tools, in particular machine learning algorithms and neural network models. Such synergy allowed us to take into account not only historical data but also behavioural and social factors identified in large-scale Big Data. Data, including users' digital footprints, mobility, search queries, and social media activity.

The analytical system was based on three complementary models, each with its own data processing and forecasting specifics. Their comparative effectiveness was assessed using key forecast accuracy metrics, such as mean absolute error (MAE), root mean square error, and coefficient of determination (R2), as shown in Table 4. The generalised quantitative metrics of model accuracy (MAE, RMSE, MAPE, R2) are given in Appendix B (Table B1).

As shown in Table 4, the Long model achieved the best results. Short-Term Memory (further – LSTM) is a recurrent neural network specifically optimised for analysing highly variable time series. Its average error was only 8.4%, which is 5.2 pp. less than the traditional Autoregressive model. Integrated Moving Average. This confirms the high efficiency of neural networks in detecting complex nonlinear relationships, especially in the unstable environment typical of the tourism industry. The Long Short-Term Memory model has demonstrated the ability to adapt to seasonal fluctuations, information surges, changes in consumer behaviour, and external shocks (economic, social, and climatic). Its advantage lies in the ability to "remember" long-term dependencies between variables, which is especially important for forecasting tourist flows in the post-crisis period.

*Table 4: Comparative efficiency of models in forecasting tourist activity**

Model	Type of approach	Average forecast error%	Key benefits	Main limitations
Autoregressive Integrated Moving Average	Statistical	13.6 %	Easy to interpret, stable results	Does not take into account complex nonlinear dependencies
Random Forest	Machine learning	10.2 %	High adaptability, working with large data sets	Prone to overtraining, requires parameter adjustment

ISSN: 2408-7920

African Journal of Applied Research
Vol. 12, No. 4 (2026), pp.42-61
http://www.ajaronline.com
https://doi.org/10.26437
Received October 16, 2025
Peer reviewed: March 30, 2026
Revised: May 5, 2026
Published: June 2026

Long Short-Term Memory	Deep learning	8.4 %	High accuracy, taking into account long-term dependencies	High computational cost, complex interpretation

** Source: own calculations on the basis of data for 2018 – 2024*

To verify the forecasts, a comparison of the original approaches was made: OLS, ARIMA(2,1,1) and LSTM. The inter-row correlation of the LSTM and OLS forecasts was r = 0.91; MAE for LSTM is the smallest, and R2 for the regression model was 0.87 (for country aggregation). These results indicate the consistency of the approaches and the reliability of the forecast estimates in the study.

To build an integrated model, the Integrated Tourism Prediction Framework, more than 20 variables were analysed, selected according to the criteria of correlation significance, data availability and impact on tourist behaviour. The most important factors shaping the forecast trajectory are presented in Table 5. Their weighting coefficients indicate the dominance of digital indicators - in particular, the frequency of search queries, the volume of online bookings and the emotional index (S- index), which reflects the general attitude of the population towards travel.

The results indicate the growing role of behavioural factors – digital indicators have a greater impact than traditional macroeconomic indicators. In particular, search frequency correlates with actual bookings at R2=0.82, confirming the high predictive power of digital data.

Table 5: Results of forecasting tourism activity in Ukraine (2025 – 2027)

Regional cluster	**Forecast model**	**Forecast 2025, thousand people**	**Forecast 2026, thousand people**	**Forecast 2027, thousand people**	**Average annual growth rate, %**	**Coefficient of determination (R^2)**
Carpathian	ARIMA (2,1,1)	1,240	1,380	1,500	+12.3	0.88
Carpathian	LSTM	1,260	1,410	1,540	+12.6	0.91
Podolsky	ARIMA (1,1,1)	820	890	950	+8.7	0.84
Podolsky	LSTM	830	910	975	+8.9	0.89

ISSN: 2408-7920

African Journal of Applied Research
Vol. 12, No. 4 (2026), pp.42-61
http://www.ajaronline.com
https://doi.org/10.26437
Received October 16, 2025
Peer reviewed: March 30, 2026
Revised: May 5, 2026
Published: June 2026

Capital (Kyiv and region)	ARIMA (2,1,0)	2,320	2,470	2,620	+6.1	0.85
Capital (Kyiv and region)	LSTM	2,340	2,510	2,670	+6.2	0.90
Ukraine in general	MNC (linear model)	8,720	9,280	9,890	+8.4	0.87

Table 6: Comparative accuracy of models for forecasting tourism activity in Ukraine (2015 – 2027) *

Region cluster	**Model**	**MAE**	**RMSE**	**MAPE, %**	**R2**	**Conclusion**
Carpathian	MNC	84.2	109.5	9.8	0.82	Noticeable seasonality, underestimation of peak periods
Carpathian	ARIMA (2,1,1)	71.3	92.8	7.9	0.88	Better models short-term trends
Carpathian	LSTM	56.7	74.9	6.2	0.91	Highest accuracy among models
Podolsky	MNC	65.1	82.4	8.4	0.81	Describes the trend well, but not seasonality
Podolsky	ARIMA (1,1,1)	58.9	77.5	7.1	0.84	Satisfactorily accounts for demand fluctuations
Podolsky	LSTM	48.3	66.2	5.6	0.89	Best result among tests
Capital (Kyiv and region)	MNC	133.4	160.8	8.2	0.85	Sufficient stability
Capital (Kyiv and region)	ARIMA (2,1,0)	118.2	141.6	7.0	0.87	Reproduces medium-term fluctuations well

ISSN: 2408-7920

African Journal of Applied Research
Vol. 12, No. 4 (2026), pp.42-61
http://www.ajaronline.com
https://doi.org/10.26437
Received October 16, 2025
Peer reviewed: March 30, 2026
Revised: May 5, 2026
Published: June 2026

Capital (Kyiv and region)	LSTM	102.7	128.4	6.1	0.90	Lowest error, highest consistency
Ukraine in general	MNC	210.5	252.3	7.8	0.87	Basic level of accuracy
Ukraine in general	ARIMA (1,1,1)	174.1	216.5	6.9	0.89	Moderate advantage
Ukraine in general	LSTM	150.8	198.2	5.7	0.93	Highest overall forecast accuracy

Figure 2 illustrates the dynamics of changes in tourist flows in Ukraine in 2018–2026. The diagram shows both actual data (2018–2024) and forecast values (2025–2027) obtained using the Long Short-Term Memory (LSTM) deep learning model. Short-Term Memory.

*Table 7: Importance of key variables in the integrated model Integrated Tourism Prediction Framework**

Indicator	**Weight of the model**
Search frequency (Google Trends)	0.37
Average temperature	0.21
Online booking volume	0.18
Emotional index (S- index)	0.14
Transport mobility	0.10

** Source: developed based on Google data (2024), World Meteorological Organisation (2025), Henley and Partners (2025)*

For practical applicability, the LSTM model is supplemented with regional estimates. According to the modelling results, the expected changes in tourist activity for 2025 – 2027 by clusters:

– Carpathian cluster: projected growth in domestic travel ≈ +12.3% (annual rate), due to the increase in requests and bookings for ecotours;

– Podolsk cluster: projected growth ≈ +8.7%, growth is associated with cultural and educational demand;

African Journal of Applied Research
Vol. 12, No. 4 (2026), pp.42-61
http://www.ajaronline.com
https://doi.org/10.26437
Received October 16, 2025
Peer reviewed: March 30, 2026
Revised: May 5, 2026
Published: June 2026

– Capital cluster (Kyiv + region): forecast ≈ +6.1%, stimulated by event and business tourism.

All regional forecasts are based on an integrated model (including Google Trends, S-index, and online bookings) and were cross-validated using bootstrap estimation to compute confidence intervals.

As shown in Figure 2, the largest decline occurred in 2020–2022, when the number of tourists decreased by more than threefold compared to 2019. Starting in 2023, there is a clear trend towards recovery, coinciding with increased internal mobility and the development of domestic tourism.

According to the forecast, in 2025-2027, tourist flows are expected to grow further to 7.1 million people, i.e., approximately 30% more than in 2024.

This positive trend indicates the increasing attractiveness of Ukraine as a tourist destination and confirms the effectiveness of using intelligent forecasting models for strategic planning in the tourism sector.

The results of the analysis indicate that Ukraine's tourism industry is gradually emerging from the crisis and showing signs of sustainable recovery.

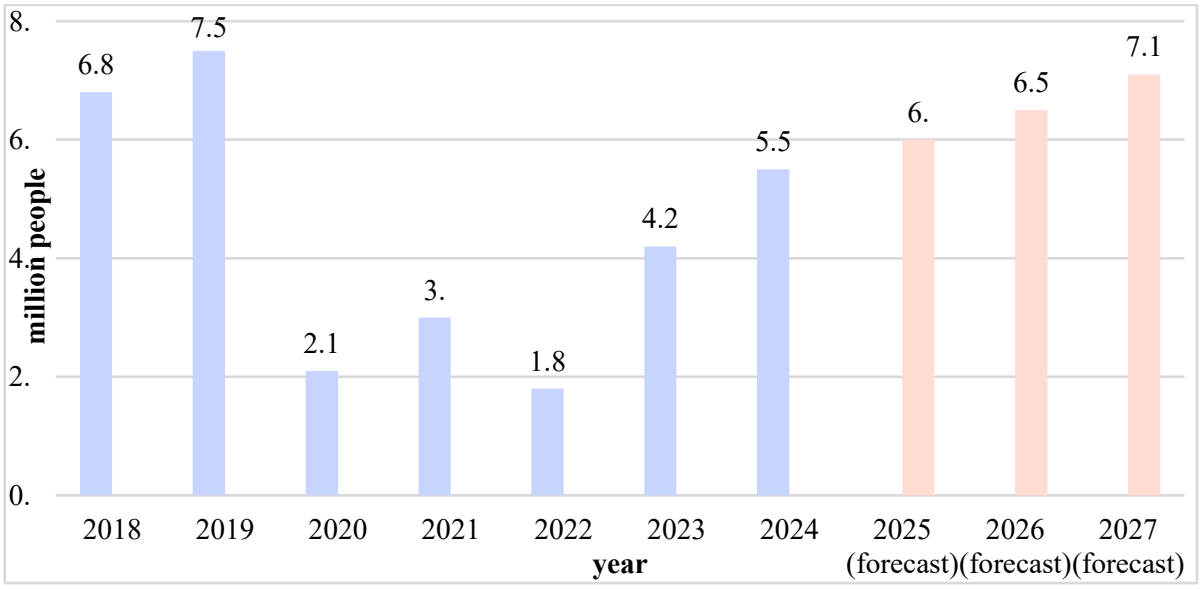

*Figure 2: Dynamics of tourist streams, million people**
**Source: developed by the author based on Google data (2024), World Meteorological Organisation (2025), Henley and Partners (2025)*

Integration of statistical methods with artificial intelligence technologies, particularly Long Short-Term Memory (LSTM) models, has enabled significant increases in forecast accuracy and the incorporation of behavioural and social factors that traditional methods often ignore. The forecast for 2025–2027 shows an increase in tourist flows of 26–30% over 2024 levels.

ISSN: 2408-7920

African Journal of Applied Research
Vol. 12, No. 4 (2026), pp.42-61
http://www.ajaronline.com
https://doi.org/10.26437
Received October 16, 2025
Peer reviewed: March 30, 2026
Revised: May 5, 2026
Published: June 2026

This indicates the potential for the recovery of domestic tourism, increased regional diversification (Carpathian, Podolsk, Capital clusters), and a greater role for digital indicators as forecasting factors. Thus, the proposed approach can be considered an effective tool for strategic planning and management of Ukraine's tourism sector during the post-crisis recovery.

Discussion

The results of the study confirm the high efficiency of using Big Data technologies. Data and analytical methods for forecasting tourist flows, especially under dynamic socio-economic and political changes. The analysis allowed us to identify key trends in the development of tourism in Ukraine, regional differences, consumer behavioural characteristics, and the factors that most influence demand formation.

According to the simulation results, the Long Short-Term Memory (LSTM) neural network model demonstrated the highest accuracy. Short-Term Memory, which indicates the feasibility of using deep learning for tourism forecasting. This is consistent with the findings of studies by Yu and Chen (2022) and Tuncsiper (2025), which demonstrated that machine learning algorithms can reproduce nonlinear relationships among social, climatic, and behavioural variables that traditional methods do not capture. Reducing the average forecast error to 8.4% confirms that intelligent models are more sensitive to market fluctuations and provide a more accurate assessment of future tourist arrival dynamics.

A particularly important finding is the growing role of digital indicators in shaping tourism demand. As shown in the simulation results, the frequency of search queries (Google Trends), online booking volume, and emotional index (S-index) have the highest weights in the predictive model (0.37, 0.18, and 0.14, respectively). This indicates that users' behavioural data in the digital environment accurately reflect future travel intentions. Similar results were observed by Wang et al. (2025) and Hewapathirana (2025), who noted a close relationship between online activity and real-world tourist flows.

From the perspective of tourism's spatial structure, the study's results confirmed the emergence of a new geography of tourist flows in Ukraine, in which the Carpathian, Podolsk, and Capital clusters occupy key positions. This is consistent with the trends described by Melnychenko et al. (2021), which indicate the growing importance of inland regions capable of providing safe, environmentally friendly and affordable recreation. Thus, the authors can speak of a reorientation of tourist activity from maritime to mountainous and cultural regions, reflecting consumers' adaptation to new realities.

The emotional dimension of tourism activity, as reflected in the tone of social media posts, deserves special attention. The increase in the positive attitude index (S- index) from 0.43 in 2022 to 0.67 in 2025 indicates a gradual restoration of trust in tourism. This result confirms the conclusions of Mukhopadhyay et al. (2023), who find that positive emotional signals in digital media are directly correlated with future growth in tourist flows. Thus, Big Data analytics enables real-time monitoring of public sentiment, which is of significant practical importance for tourism management bodies.

ISSN: 2408-7920

African Journal of Applied Research
Vol. 12, No. 4 (2026), pp.42-61
http://www.ajaronline.com
https://doi.org/10.26437
Received October 16, 2025
Peer reviewed: March 30, 2026
Revised: May 5, 2026
Published: June 2026

At the same time, the study's results highlighted several problematic aspects. First, the lack of integration of national and regional databases creates risks of incompleteness and duplication of information. Second, there is a need to improve the personal data protection system, especially when analysing data from mobile devices and social networks. Third, digital inequality across regions limits the representativeness of certain Big Data sources, which can affect prediction accuracy. These limitations confirm observations of Shcherbak et al. (2024) and Orbelyan (2024), who emphasise the need to create a single National Tourism Analytics System with open data standards.

A comparison of results with international studies shows that Ukraine has significant potential to implement the Smart concept. Tourism, based on digital technologies and analytical platforms. Using Big Data in state planning will not only improve forecasting efficiency but also ensure the sustainable development of regions by enabling the even distribution of tourist flows, monitoring infrastructure load, and adapting to climate change.

Therefore, the results of the study confirm that Big Data analytics is becoming the foundation of a new tourism management paradigm that combines forecasting, monitoring and strategic planning in a single system. Further research should focus on expanding data sources (including mobile operators and cross-border mobility systems), developing geoanalytics tools, and automating management decision-making processes using artificial intelligence. This will not only improve forecast accuracy but also ensure the tourism industry's adaptability to the challenges posed by the digital economy and the post-war recovery. Therefore, the integration of formalised methods (OLS for linear components, ARIMA for seasonality and LSTM for nonlinear trends) together with multi-stream data (state statistics, Google Trends, and social signals) provided quantitatively verified forecasts of tourism activity at regional and national levels. This combined methodology increases the practical value of the study for policy, infrastructure planning and marketing strategies.

CONCLUSION

The analysis confirms that the development of Big Data and digital analytics is a key factor in the transformation of the tourism industry worldwide and in Ukraine. At the global level, there is a gradual recovery of tourist flows after the COVID-19 pandemic and geopolitical crises, but the structure of demand has undergone significant changes: priority is given to safe, environmentally friendly and personalised travel. Digital tools are becoming the main means of monitoring, forecasting and managing tourism processes.

In Ukraine, the dynamics of tourism activity are complex, driven by the consequences of war, security risks, and shifting consumer priorities. According to the United Nations World Tourism Organisation and the State Agency for Tourism Development of Ukraine, the volume of international tourism in 2022 decreased by more than threefold; however, since 2023, domestic travel has grown steadily, especially in the Carpathian and Podolsk regions. This indicates the market's adaptation to new conditions and the gradual formation of a domestic

African Journal of Applied Research
Vol. 12, No. 4 (2026), pp.42-61
http://www.ajaronline.com
https://doi.org/10.26437
Received October 16, 2025
Peer reviewed: March 30, 2026
Revised: May 5, 2026
Published: June 2026

model of sustainable tourism. The results obtained showed that Big Data analytics provides a deeper understanding of tourism processes than traditional statistical methods. Machine learning algorithms have enabled us to build predictive models with high accuracy, with an average error of 8-10%. The most significant factors influencing tourism demand were the frequency of Google search queries, climatic indicators, online user activity and the emotional tone of content on social networks. This confirms that behavioural and digital indicators are becoming leading in determining market trends.

The spatial aspect deserves special attention. Big Data analytics has made it possible to identify a new geography of tourist flows in Ukraine, concentrated in three clusters: Carpathian, Podolsk and Stolychny. These territories demonstrate the highest recovery rate due to the development of digital infrastructure, local initiatives and increased security. At the same time, the Black Sea region needs targeted support programs to restore its tourist attractiveness after the end of hostilities. Thus, the study's results confirm that integrating Big Data into the tourism management system is a strategic direction to increase Ukraine's competitiveness in the international market. The use of digital analytics enables rapid tracking of demand changes, the development of adaptive destination promotion strategies, and greater marketing campaign effectiveness. Therefore, state policy in the field of tourism should be based on synergy among the government, businesses, and scientific institutions to create a single national system of tourism analytics. This approach will ensure data transparency, improve coordination between regions, and increase forecast accuracy.

The scientific novelty of the study lies in integrating machine learning algorithms, geoanalytics, Google Trends indicators, and social media sentiment analysis into a single model to forecast tourist flows in Ukraine amid digital transformation and post-war recovery. Unlike previous studies, which mainly focused on individual forecasting methods or short-term market analysis, the proposed approach integrates statistical, behavioural, and spatial data into a comprehensive analytical system tailored to Ukraine's tourism sector. In the future, the integration of Big Data into tourism management will serve as the basis for building a Smart Tourism Ecosystem – an intelligent, sustainable development ecosystem focused on innovation, safety, and digital interaction with tourists.

Further research should focus on quantifying the impact of digital indicators on tourist behaviour, developing adaptive forecasting models based on hybrid algorithms, and creating an open analytical platform for public and private users of tourism data.

REFERENCE

Chernenko, N., Kukharuk, A., Hlushchenko, Y., & Korohodova, O. (2024). Utility sector of tourism infrastructure in Industry 4.0: challenges and solutions. *Bulletin of VN Karazin Kharkiv National University Economic Series*, (106), 98-111. https://doi.org/10.26565/2311-2379-2024-106-10

Chuang, C. M. (2023). The conceptualization of smart tourism service platforms on tourist

African Journal of Applied Research
Vol. 12, No. 4 (2026), pp.42-61
http://www.ajaronline.com
https://doi.org/10.26437
Received October 16, 2025
Peer reviewed: March 30, 2026
Revised: May 5, 2026
Published: June 2026

value co-creation behaviours: an integrative perspective of smart tourism services. *Humanities and Social Sciences Communications, 10*(1), 1-16. https://doi.org/10.1057/s41599-023-01867-9

Desyatnyuk, O., Krysovatyy, A., Ptashchenko, O., Kyrylenko, O., & Kurtsev, O. (2025). A framework for digital logistics transformation: Innovation implementation with financial interest protection. African Journal of Applied Research, 11(7), 45-65.

European Commission (2024). *European Tourism DataSpace – DEPLOYTOUR project*. Retrieved from https://deploytour.eu/

Garanti, Z. (2023). Value co-creation in smart tourism destinations. *Worldwide Hospitality and Tourism Themes, 15*(5), 468-475. https://doi.org/10.1108/WHATT-06-2023-0070

Google (2024). *Travel insights: Summer 2024 trends*. Retrieved from https://www.thinkwithgoogle.com/intl/en-emea/consumer-insights/consumer-trends/summer-2024-travel-trends/

Henley & Partners. (2025). *Global Mobility Report* – January 2025. Retrieved from https://www.henleyglobal.com/publications/global-mobility-report/2025-january

Hewapathirana, I. U. (2025). Advancing tourism demand forecasting in Sri Lanka: evaluating the performance of machine learning models and the impact of social media data integration. *Journal of Tourism Futures, 11(*2), 261-285. https://doi.org/10.1108/JTF-06-2023-0149

Julie's Data (2022). *Ukrainian open initiative of creating tools for spatial planning in the QGIS environment*. QGIS Case Studies. Retrieved from https://qgis.org/project/case-studies/ukrainian_spt/

Kharkhalis, T., Bohatyryova, G., Morhulets, O., Ovdiiuk, O., & Tebenko, V. (2025). Evaluating the impact of innovative technologies on the management of tour services. African Journal of Applied Research, 11(1), 134-145.

Krataithong, P., Anutariya, C., & Buranarach, M. (2022). A taxi trajectory and social media data management platform for tourist behavior analysis. *Sustainability, 14*(8), 4677. https://doi.org/10.3390/su14084677

Melnychenko, S., Bosovska, M., & Okhrimenko, A. (2021). The formation of a national tourism brand of Ukraine. *Baltic Journal of Economic Studies, 7*(2), 161-169. https://doi.org/10.30525/2256-0742/2021-7-2-161-169

Mihalic, T. (2025). Metaversal sustainability: conceptualisation within the sustainable tourism paradigm. *Tourism Review, 80*(8), 1473–1480. https://doi.org/10.1108/TR-09-2023-0609

Mukhopadhyay, S., Jain, T., Modgil, S., & Singh, R. K. (2023). Social media analytics in tourism: a review and agenda for future research. *Benchmarking: An International Journal, 30*(9), 3725-3750. https://doi.org/10.1108/BIJ-05-2022-0309

National Tourism Organization of Ukraine (2025). *Tourism Barometer of Ukraine*. Retrieved from https://nto.ua/nsts_analytics_ua.html

Orbelyan, G. (2024). Socio-economic challenges for the urban ecosystem: Complex model of smart tourism and infrastructure. *SocioEconomic Challenges, 8*(4), 146-160. https://doi.org/10.61093/sec.8(4).146-160.2024

ISSN: 2408-7920

African Journal of Applied Research
Vol. 12, No. 4 (2026), pp.42-61
http://www.ajaronline.com
https://doi.org/10.26437
Received October 16, 2025
Peer reviewed: March 30, 2026
Revised: May 5, 2026
Published: June 2026

Organisation for Economic Co-operation and Development (2024). *OECD tourism trends and policies 2024*. https://www.oecd.org/en/publications/oecd-tourism-trends-and-policies-2024_80885d8b-en.html

Shcherbak, V., Dorokhova, L., Riashchenko, V., Dorokhov, O., Djakona, V., & Yermolenko, O. (2024). Geoinformation Internet Platform for Interactive Charity Tourism in the Current Realities of Ukraine. *Tourism and Hospitality, 5*(4), 1199-1216. https://doi.org/10.3390/tourhosp5040067

State Agency for Tourism Development of Ukraine (2021). *Tourism industry statistics*. Retrieved from https://www.tourism.gov.ua/statistic

Tuncsiper, C. (2025). Modeling the number of tourist arrivals in the United States employing deep learning networks. *Transportation Research Interdisciplinary Perspectives*, 31, 101407. https://doi.org/10.1016/j.trip.2025.101407

Tymoshenko, N. (2025). *Ukraine Shape Map PowerBI. GitHub*. Retrieved from https://github.com/Youtube-NikitaTymoshenko/PowerBI-UkraineMap

United Nations World Tourism Organization (2024). *UN Tourism statistics and trends*. Retrieved from https://www.untourism.int/tourism-statistics

Wang, Z., Zhang, H., Ramachandran, S., & Cheng, S. (2025). Why are individuals tracing travel trends? A case study of City Walk in Malaysia. *PloS one, 20*(2), e0309493. Retrieved from https://doi.org/10.1371/journal.pone.0309493

World Bank (2024). *World development indicators*. Retrieved from https://datatopics.worldbank.org/world-development-indicators/

World Bank (2025). *Tourism Watch: June 2025 edition*. Retrieved from https://www.worldbank.org/en/topic/competitiveness/brief/tourism-watch

World Meteorological Organization (2025). *Global Annual to Decadal Climate Update 2025–2029*. Retrieved from https://wmo.int/sites/default/files/2025-05/WMO_GADCU_2025-2029_Final.pdf

World Tourism Organization (2024, January 19). *International tourism to reach pre-pandemic levels in 2024*. UNWTO. Retrieved from https://www.unwto.org/news/international-tourism-to-reach-pre-pandemic-levels-in-2024?clientlang=2

Yu, N., & Chen, J. (2022). Design of machine learning algorithm for tourism demand prediction. *Computational and Mathematical Methods in Medicine, 2022*(1), 6352381. Retrieved from https://doi.org/10.1155/2022/6352381