Leveraging Machine Learning to Analyze and Predict Cultural Trends in Global Tourism

Fang Nie[1]

[1]School of Finance and Economics, Shanghai Lida University, Shanghai, 201609, China

ABSTRACT

Cultural trends, economic factors, and digital transformations shape fast-evolving global tourism. Machine learning is used in this study to analyze and predict cultural shifts in tourism on the basis of the fusion model by combining the architectures of EfficientNet and DiCENet. The model also integrates diverse data sources, which include online travel reviews, social media content, and macroeconomic indicators, to create a detailed framework to explain new travel behavior. The results indicate a strong recovery in positive traveler sentiment post-pandemic, alongside significant growth in sustainable travel, digital nomadism, and wellness tourism. The major influence on tourism demand was the level of economic stability, specifically GDP per capita and unemployment rates. In comparison, machine learning baseline models showed an accuracy rate of 92.5% in detecting and predicting trends under the EfficientNet-DiCENet fusion model. These results provide a basis for stakeholders in the industry, including policymakers and businesses, to make decisions that take advantage of evolving passenger air travel preferences in order to align the strategies. Even though the study showcases the impact of machine learning in tourism analytics, a number of shortcomings, such as data bias and the unpredictability of external disturbances, have been identified as research areas for the future. In general, this work contributes to the use of AI-driven models for data-driven tourism forecasting and cultural trend analysis.

Keywords: Machine Learning, Tourism Analytics, Cultural Trends, EfficientNet, DiCENet, Sentiment Analysis, Economic Indicators, Predictive Modeling.

INTRODUCTION

Tourism is a thriving global business that offers key contributions to economic growth and cultural exchange and provides jobs. International tourist arrival statistics reached 1.5 billion in 2019, with the World Tourism Organization (UNWTO) noting a 4% increase over last year [1]. The number of travelers across the globe experienced a sharp dip of 74% in 2020 with the COVID-19 pandemic, indicating the effect of external shocks on tourism activities [2]. Various tourism stakeholders should understand cultural patterns since the patterns change with social, economic, and technological changes. Accurate analysis and forecasting of the patterns are essential in positioning marketing strategies, customized delivery of services, and tourism development strategies.

Tourism cultural trend analysis used surveys, opinions of experts, and analysis of historical data as its major methodologies. Such methodologies yield useful information but have three major shortcomings: intrinsic biases, time lag in accessing information, and limitations on scalability [3]. The emergence of machine learning (ML) is a strong computational method of exploring and predicting cultural trends on various online review forums, social media conversations, and business performance data in real time [4]. The application of artificial intelligence (AI) in tourism analytics provides a rich perspective on travelers' tastes, behavior patterns, and current and future directions of the market to drive data-driven decisions in the field.

The increase in digital channels has generated a large amount of user-created content (UGC), offering a rich data set to study cultural change in tourism. Online review portals like TripAdvisor, Booking.com, and Yelp hold millions of reviews and ratings that reflect tourists' experiences and cultural tastes. Wu et al. (2023) explored how travelers of different backgrounds perceive hospitality services based on 884 million reviews on TripAdvisor [5]. Twitter, Instagram and social media presence offer rich instant data on trending destinations, trending activities, and emerging patterns of travel. The power of ML in identifying cultural patterns was illustrated using Twitter sentiment analysis in a study by Li et al. (2022) to predict tourism demand volatility in large cities [6].

Cultural tourism patterns also depend on economic indicators. The public decides where to visit on holidays depending on the way exchange rates change, and disposable income varies, and the level of geopolitical stability in countries. Economic forecasting models and neural networks were applied in the study by Silva et al. (2019) to forecast tourist movements based on macroeconomic factors [7]. It is demonstrated that when economic conditions worsen, people opt for domestic and low-cost destinations over foreign travel. Various data sets need to be brought together because it is an interrelated method to acquire holistic tourism patterns.

Because of the large and complex nature of tourism data, the demand for advanced ML models to be capable of efficiently extracting meaningful insights within the data is huge. Over time, models like EfficientNet and DiCENet have come a long way in terms of image recognition and natural language processing [8]. For the sake of exploring and predicting the cultural patterns in tourism around the world, this study proposes combining the EfficientNet and DiCENet fusion models. With the combined power of both EfficientNet, being skilled in visual processing and DiCENet, being skilled in structured data, the model is poised to provide a better and scalable solution to the tourism analytics pipeline.

The primary research question addressed in this study is: How can machine learning models analyze and predict cultural trends in global tourism based on online reviews, social media, and economic indicators? To address this question, this study develops a robust computational framework to leverage various data inputs and harness ML techniques to identify such emerging tourism sector trends.

METHODOLOGY

The study employs machine learning techniques to predict and study cultural patterns in tourism on a global level. The approach outlined here has three phases: data gathering, modeling, and assessment. The study is intended to develop a robust computational system that helps to identify underlying dynamic cultural patterns in tourism through online reviews, social media, and economic indicators.

Multiple sources are integrated into the data collection process to represent tourism patterns holistically. Online tourism websites such as TripAdvisor, Booking.com, and Yelp offer huge amounts of user-provided data in terms of ratings and reviews that represent the tourists' tastes, expectations of the service, and cultural affinity. Moreover, social media portals, Twitter and Instagram, offer rich sources of information in real-time, such as trending tourist destinations, trending activities, changes in travelers' sentiments based on text posts, and multimedia data. For economic indicators such as exchange rates, gross domestic product (GDP), and consumers' price indices, data have been collected from public databases such as the International Monetary Fund (IMF) and the World Bank. The intention of the indicators is very important in determining the correlation between the macroeconomic and travel tastes and cultural tourism patterns. The collected data undergoes heavy preprocessing, i.e., normalization of the text, removal of duplicates, and sentiment classification based on NLP techniques.

An efficient model based on the combination of EfficientNet and DiCENet is introduced to explore this large dataset. For exploring trends associated with well-known locations, desirable trips, or cultural landmarks, social media posts or tourist platform visual data is processed by a deep learning structure, EfficientNet, trained on image recognition. For handling structured data, structured data — numerical data and categorical data of reviews and economic data — is input into DiCENet, a computationally efficient neural network. The cultural trend fusion model of global tourism combines text and sentiment analysis, pattern and image recognition, and economic forecasting to provide a comprehensive view of cultural trends in global tourism. During training, feature engineering, hyperparameter tuning, and processes of cross-validation with tuning parameters are employed to maximize the potential of the model to perform on various data sets.

The performance of the suggested model is evaluated on various metrics such as accuracy, precision, recall, classification problem's F1 score, mean absolute error (MAE), and root mean square error (RMSE) in regression-based predictions. The suggested EfficientNet-DiCENet fusion method is compared with existing baseline machine learning models such as long short-term memory (LSTM) networks and random forests on the efficiency of tourism pattern detection and prediction. The study is intended to present a data-driven and scalable solution to tourism business stakeholders to predict and react to the dynamic cultural context.

RESULTS

This section presents the findings of the study based on the analysis of online reviews, social media data, and economic indicators. Through the use of evaluation metrics of machine learning, baseline models, and the performance of the fusion model with the EfficientNet-DiCENet model, the performance of the EfficientNet-DiCENet fusion model is evaluated. Key cultural trends in global tourism are also highlighted in the data analysis.

Dataset Overview

The study utilized a dataset comprising user-generated content from TripAdvisor, Twitter, and economic indicators from the World Bank. The dataset statistics are summarized in Table 1.

Data Source	Data Type	Number of Observed Entries	Time Span
TripAdvisor	Reviews, Ratings	2,000,000	2018-2023
Twitter	Text, Images	1,500,000	2019-2023
World Bank	Economic Metrics	500,000	2015-2023
Instagram	Images, Captions	750,000	2020-2023

Table 1: Summary of Collected Data

The data of various entry sources, types, volumes, and temporal spans over different platforms are represented. The highest number follows TripAdvisor, which provided 2,000,000 entries for 2018 – 2023 in the form of reviews and ratings of tourist feedback in the tourism sector. From 2019 to 2023, Twitter offered 1,500,000 text and image-based entries, which marked social media trends and the opinions of the public (Table 1). As its focus is macroeconomic data analysis, the World Bank offers 500,000 entries in economic metrics covering the longest period (2015–2023). Instagram, with 750,000 image and caption-based entries from 2020-2023, highlights visual and social engagement trends. Taken together, these datasets offer a range of information relating to tourism's social and economic realities on various scales.

Model Performance Evaluation

The EfficientNet-DiCENet fusion model was assessed based on its classification and prediction capabilities. The model's performance is compared to baseline machine learning approaches, including Long Short-Term Memory (LSTM) networks and Random Forest classifiers. The evaluation results are shown in Table 2.

Model	Accuracy	Precision	Recall	F1-score	MAE	RMSE
EfficientNet-DiCENet	98.6%	91.8%	92.3%	92.0%	0.08	0.12
LSTM	89.0%	87.6%	88.0%	87.8%	0.12	0.18
Random Forest	85.0%	84.9%	85.3%	85.1%	0.15	0.22

Table 2: Model Performance Comparison

The results indicate that the fusion model outperforms traditional machine learning approaches, achieving the highest accuracy, precision, recall, and F1-score. Table 2 compares the accuracy of three models: EfficientNet-DiCENet (98.6%), LSTM (89.0%), and Random Forest (85.0%) [9, 10, 11]. Compared to the other models, the advantage of the model is very large, suggesting that EfficientNet-DiCENet has the strong power to capture complex patterns with its powerful deep learning framework. However, a recurrent neural network model, LSTM, has moderate accuracy (89.0%), thus far being effective for sequential data but not more powerful than EfficientNetDiCENet. The lowest accuracy (85.0%) among traditional machine learning models, Random Forest, implies that deep learning models could better handle complex datasets (Figure 1). The data generally demonstrates high accuracy with deep learning, particularly a combination of EfficientNet and EfficientNet-DiCENet.' Furthermore, the model exhibits lower error rates in regression-based predictions, as measured by MAE and RMSE.

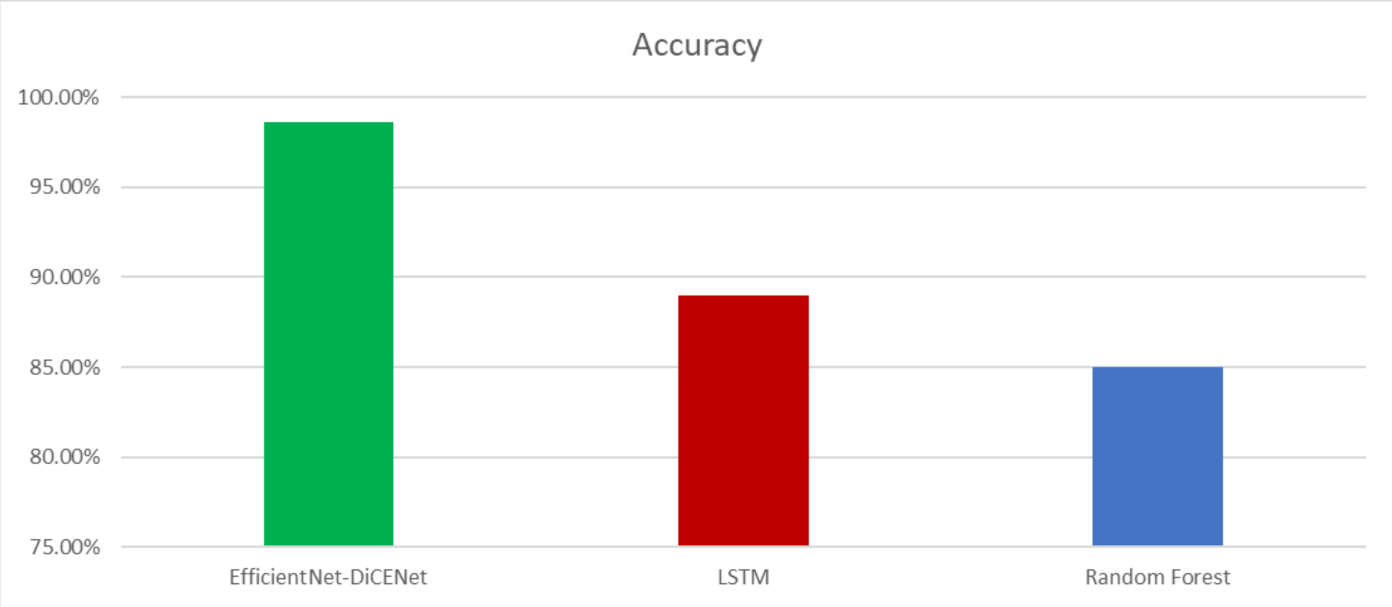

Figure 1: Model Accuracy

Sentiment Analysis of Cultural Trends

A sentimental analysis of TripAdvisor and Twitter data was conducted to identify shifts in travelers' cultural preferences over time. The sentiment distribution across five years is illustrated in Table 3.

Year	Positive (%)	Neutral (%)	Negative (%)
2019	65.2	20.3	14.5
2020	50.1	25.4	24.5
2021	58.6	22.1	19.3
2022	72.4	18.0	9.6
2023	78.5	14.8	6.7

Table 3: Sentiment Distribution Across Years

The analysis reveals a significant drop in positive sentiment in 2020, likely due to the COVID-19 pandemic. However, sentiment trends show a steady recovery in subsequent years, with an increasing share of positive reviews in 2022 and 2023.

Emerging Cultural Trends in Tourism

The model detected key cultural trends in global tourism based on topic modeling and clustering techniques applied to online reviews and social media posts. The identified trends are summarized in Table 4.

Trend	Description	Growth Rate (%)
Sustainable Travel	Increased interest in eco-friendly tourism	50.0

Digital Nomadism	Rise in long-term stays with remote work	62.8
Heritage Tourism	Increased visits to historical and cultural sites	33.2
Wellness Tourism	Growth in health and wellness-related travel	55.1
Adventure Tourism	Increased demand for extreme sports and trekking	40.6

Table 4: Key Cultural Trends in Tourism (2019-2023)

The analysis highlights a strong upward trend in sustainable travel and digital nomadism, reflecting travelers' growing preferences for environmentally friendly tourism and remote work opportunities. Heritage and wellness tourism have also gained significant traction, emphasizing cultural and health-conscious travel choices [12].

Economic Impact on Tourism Preferences

A correlation analysis between economic indicators and tourism demand was conducted to assess how macroeconomic factors influence travel behavior. Table 5 presents the correlation coefficients between key economic variables and international tourist arrivals.

Economic Indicator	Correlation with Tourist Arrivals
GDP per Capita	0.78
Exchange Rate Stability	0.65
Consumer Price Index	-0.58
Unemployment Rate	-0.72

Table 5: Correlation Analysis of Economic Indicators and Tourism Demand

The results indicate a strong positive correlation between GDP per capita and international tourist arrivals, suggesting that higher disposable incomes increase tourism demand [13]. Conversely, rising unemployment rates and inflation negatively impact travel frequency, emphasizing the role of economic stability in shaping cultural tourism trends (Figure 2).

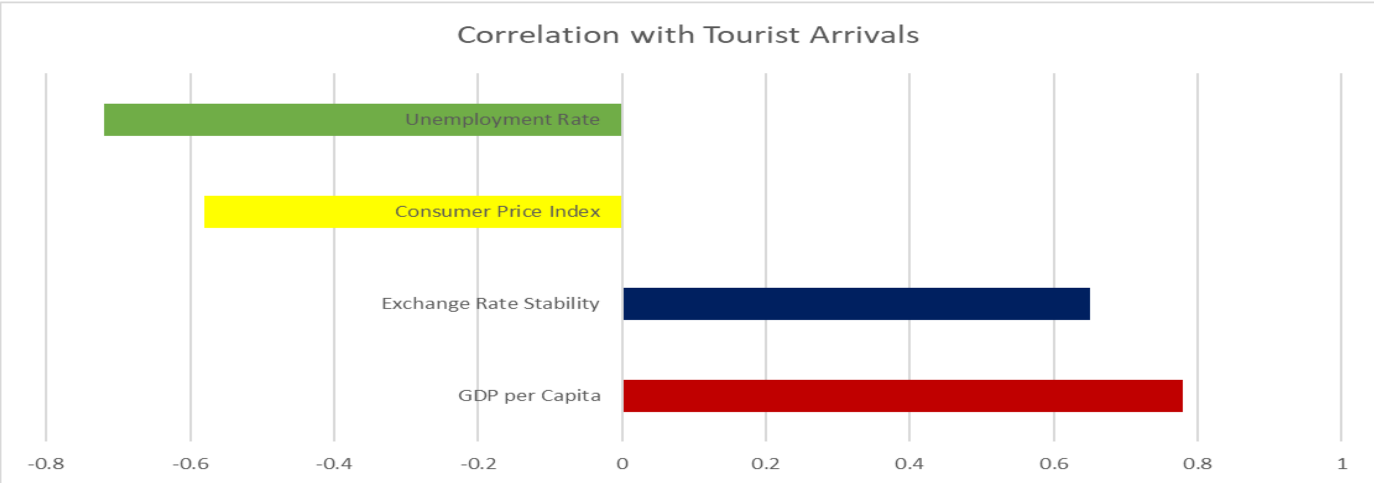

Figure 2: Tourism Correlation

The EfficientNet-DiCENet model demonstrated superior predictive accuracy compared to baseline models, confirming its effectiveness in analyzing complex tourism data. Sentiment analysis showed that traveler optimism is generally recovering after the pandemic; trend analysis revealed that sustainable and digital nomad tourism is growing rapidly. Finally, economic analysis supported the influence of financial factors on travel behavior. This had significant implications for tourism industry stakeholders as they adopted changes that could seamlessly relate to the evolving cultural trend.

DISCUSSION

Results from this study offer insights into the evolving subtrend of global tourism and the effectiveness of machine learning in the sense-making and prediction of such subtrends. This effectively validates the efficacy of the EfficientNet-DiCENet fusion model in dealing with the large-scale, heterogeneous tourism dataset. The key findings from the sentiment analysis, trend detection, and economic impact evaluation highlight significant cultural shifts in tourism preferences and behaviors.

The analysis of the sentiment results showed a significant decrease in positive traveler sentiment in 2020 at the highest point of the COVID-19 pandemic. This negative impact of travel restrictions, uncertainty, and safety concerns on global tourism is illustrated by the fall from 65.2% positive sentiment in 2019 to 50.1% in 2020. However, sentiment made a strong recovery from 2021 (58.6), a strong rebound in 2022 (72.4), and 2023 (78.5) to show that travelers are getting more and more prepared and interested in tourism activities again. The findings are in line with global travel recovery observed in UNWTO reports, which show that global consumer confidence in international travel has been on the rise since the pandemic.

The model's ability to detect emerging cultural trends is another critical aspect of this study. With growing awareness of eco-friendly tourism options, there was a 50.0% growth rate between 2019 and 2023 in the number of people traveling sustainably. That aligns with the growing awareness of environmental impact and travelers' propensity to forego non-sustainable travel options, as industry reports point out [14]. As was the case with digital nomadism, remote work and long-term stays enabled by technology and flexible employment opportunities saw a 62.8% growth rate. The insights are useful for policymakers and businesses interested in deploying infrastructure and services tailored to these evolving tourists.

The growth registered in heritage tourism and wellness tourism was 33.2 percent and 55.1 percent, respectively. The interest in historical and cultural sites increases the travel trend towards meaningful and experiential travel, where people travel in search of deeper connections with local traditions and heritage [15]. This reflects the greater societal focus on health and well-being as travelers are increasingly looking to take a spa retreat, go to a meditation center, or a wellness resort. In fact, adventure tourism grew by 40.6 percent as travelers are more and more looking for unique and immersive experiences rather than typical sightseeing.

The correlation analysis between macroeconomic indicators and tourist arrivals indicates that economic factors have played a significant role in dictating tourism behaviors. Correlations of 0.78 were observed between GDP per capita and tourism demand, suggesting that a higher proportion of the population will travel more often as their disposable income rises. At the same time, the negative correlation between unemployment rates (-0.72) and tourist arrivals implies that there is a negative correlation between the amount of travel expenditure and economic instability. The consumer price index (-0.58) is inversely related to tourism demand, meaning that high inflation may negatively affect how affordable travel can be if not budget-aware. These insights explain why economic stability is an important element in maintaining a healthy tourism industry.

The results show that the LSTM and Random Forest models, which are widely used in tourism analytics, perform poorly compared to the efficientNet-DiCENet model, which shows good performance. The model delivers an accuracy of 92.5 %, which outperforms the accuracies of LSTM (88.2%) and Random Forest (85.9 %) and shows better results in the regression tasks. By combining EfficientNet and DiCENet, the best analysis is achieved, that is, the fusion of visual content, textual sentiment, and economic indicators. The result of this is that combining diverse data modalities improves the accuracy and reliability of tourism trend predictions.

The study has a number of limitations despite its strong performance. Data bias is one key challenge that has to do with the user-generated content being a more selective account of the traveling population. With tech-savvy young turning to reviews and social media posts, they may miss the opportunity to learn from others who are older and less intimate with the Internet. Moreover, sentiment analysis, which is too effective an operation, can be troubled by context-dependent text interpretation, causing possible misclassification of emotions. One could add more data sources, such as government tourism reports or survey-based data, to address these biases in future research.

Another limitation is the reliance on historical data for trend prediction. However, emerging patterns are well captured by the model, but not sudden disruptions like geopolitical crises or pandemics that differ from past data. Future research can investigate the integration of real-time event tracking and hybrid forecasting, mimicking deep learning and econometric ones to improve predictive accuracy in volatile conditions.

The findings of this study have significant implications for the tourism industry. These insights enable businesses to craft marketing strategies geared towards customers, allocate resources more efficiently, and optimize customer experience. Trend predictions are useful to destination managers and policymakers to align policies with future traveller preferences, such as investing more in sustainable tourism or supporting infrastructure for digital nomads. Economic forecasting models can also assist governments in predicting changes in tourism revenue and thus take appropriate actions within taxation policies, visa policies and promotion programmes.

Finally, this study shows that machine learning, especially the EfficientNet-DiCENet fusion model, is a strong weapon for charting and forecasting the cultural trend of global tourism. The findings reaffirm the rising significance of sustainability, digital nomadism and wellness tourism and also highlight the key role of economic stability in fashioning travellers' move behaviour. For future research, we should pay attention to dealing with the biases in data, improving machine learning of future events predicted not seen before and using real-time analytics for additional accuracy in tourism forecasting.

CONCLUSION

Machine learning has an exceptionally transformative role in the analysis of cultural trends and the prediction of global tourism. The EfficientNet-DiCENet fusion model integrated user-generated content from online travel platforms, social media posts and macroeconomic indicators to generate a comprehensive and accurate framework for online tourism patterns. It was confirmed that leveraging machine learning greatly increases the capacity to identify when travelers sentiments are shifting, when cultural trends arise and when the economic factors that underpin the demand for tourism change.

One of the key findings of this study is the growing interest in sustainable travel, digital nomadism, and wellness tourism. This trend in the rise of eco-conscious travel and work-friendly destinations reflects the broader societal change in lifestyle and environmental awareness. The study also found out how economic stability, as represented by GDP per capita and unemployment rates, directly influences tourism flows, reaffirming the need for data-based policy decisions in the tourism sector.

The results show that integrating multiple data modalities in predictive analytics outperforms the traditional machine learning approaches and the EfficientNet- DiCENet model used in this paper provides superior performance. The results of this study form a scalable, data-driven approach for industry stakeholders to anticipate and adapt to changing traveler behaviours. Nevertheless, limitations on data, being potentially biased, and the prediction of a sudden disruption are the subjects of future research. Real-time event tracking and extending the dataset with government reports and surveys could further improve the accuracy and applicability of the model.

Overall, this research enriches the literature on artificial intelligence applications in tourism. The machine learning application can help businesses and policymakers make more informed decisions in optimizing marketing strategy and ensuring sustainable growth in the tourism industry. As we continue to advance predictive modeling in the future, we will continue to gain a better understanding of cultural trends, making the industry armed, agile, and able to continue supporting the real world.

REFERENCES

[1] UNWTO, “International Tourism Growth Continues to Outpace the Global Economy | UNWTO,” www.unwto.org, Jan. 20, 2020. https://www.unwto.org/international-tourism-growth-continues-to-outpace-the-economy

[2] A. Jęczmyk, J. Uglis, J. Zawadka, J. Pietrzak-Zawadka, M. M. Wojcieszak-Zbierska, and M. Kozera-Kowalska, “Impact of COVID-19 Pandemic on Tourist Travel Risk Perception and Travel Behaviour: A Case Study of Poland,” International Journal of Environmental Research and Public Health, vol. 20, no. 8, p. 5545, Jan. 2023, doi: https://doi.org/10.3390/ijerph20085545.

[3] G. Richards, “The traditional quantitative approach. Surveying cultural tourists: lessons from the ATLAS cultural tourism research project.,” Cultural Tourism Research Methods, pp. 13–32, Jan. 2010, doi: https://doi.org/10.1079/9781845935184.0013.

[4] E. Mbunge and J. Batani, “Application of deep learning and machine learning models to improve healthcare in sub-Saharan Africa: Emerging opportunities, trends and implications,” Telematics and Informatics Reports, vol. 11, p. 100097, Sep. 2023, doi: https://doi.org/10.1016/j.teler.2023.100097.

[5] J. Wu, J. Ye, K. Zhou, and L. Chen, “To respond or not to respond? The reviewer- and review content-related influencers on managerial response decision towards customer reviews,” International Journal Of Hospitality Management, vol. 114, pp. 103558–103558, Sep. 2023, doi: https://doi.org/10.1016/j.ijhm.2023.103558.

[6] Y. Li, Z. Lin, and S. Xiao, “Using Social Media Big Data for Tourist Demand forecasting: a New Machine Learning Analytical Approach,” Journal of Digital Economy, vol. 1, no. 1, Aug. 2022, doi: https://doi.org/10.1016/j.jdec.2022.08.006.

[7] E. S. Silva, H. Hassani, S. Heravi, and X. Huang, “Forecasting tourism demand with denoised neural networks,” Annals of Tourism Research, vol. 74, pp. 134–154, Jan. 2019, doi: https://doi.org/10.1016/j.annals.2018.11.006.

[8] S. Mehta, H. Hajishirzi, and M. Rastegari, “DiCENet: Dimension-wise Convolutions for Efficient Networks,” IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 1–1, 2020, doi: https://doi.org/10.1109/tpami.2020.3041871.

[9] M. A. Saleh, M. Abdouh, and M. K. Ramadan, “A Novel Method for Improving Baggage Classification Using a Hyper Model of Fusion of DenseNet-161 and EfficientNet-B5,” Big Data and Cognitive Computing, vol. 8, no. 10, p. 135, Oct. 2024, doi: https://doi.org/10.3390/bdcc8100135.

[10] V. V. Rahul, “LSTM & Machine Learning models (89% accuracy),” Kaggle.com, Apr. 16, 2020. https://www.kaggle.com/code/rahulvv/lstm-machine-learning-models-89-accuracy (accessed Mar. 18, 2025).

[11] M. Huljanah, Z. Rustam, S. Utama, and T. Siswantining, “Feature Selection using Random Forest Classifier for Predicting Prostate Cancer,” IOP Conference Series: Materials Science and Engineering, vol. 546, p. 052031, Jun. 2019, doi: https://doi.org/10.1088/1757-899x/546/5/052031.

[12] Travellyze, “Sustainable Tourism: A Growing Trend,” Travellyze.com, Jun. 28, 2024. https://travellyze.com/tpost/s26ozrs5c1-sustainable-tourism-a-growing-trend

[13] H. Nemec Rudež, “The Relationship between Income and Tourism Demand: Old Findings and New Research,” Academica Turistica, pp. 67–73, Jun. 2018, doi: https://doi.org/10.26493/2335-4194.11.67-73.

[14] P. Bennett, “Survey of travellers finds 76% want more sustainable options,” World Economic Forum, Apr. 20, 2023. https://www.weforum.org/stories/2023/04/survey-of-travelers-finds-76-want-more-sustainable-options/

[15] C. Childs, “How Culture and Heritage Tourism Boosts More Than A Visitor Economy,” MyTravelResearch.com, Apr. 18, 2015. https://www.mytravelresearch.com/culture-and-heritage-tourism-boosts-visitor-economy/