

Research Article

Garima Malik*, Dharmendra Singh

Twitter sentiment analysis: An estimation of the trends in tourism after the outbreak of the Covid-19 pandemic

<https://doi.org/10.2478/ejthr-2023-0004>

received June 6, 2021; accepted January 16, 2023

Abstract: The novel coronavirus pandemic drastically impacted economies in 2020, including travel and tourism. The nationwide lockdown to curb infection has led people to use social media such as Twitter to express their opinions and share information on several issues. This paper focuses on sentiment analysis using Indian tourist Tweets during Covid-19 using Python and the maximum likelihood method to determine the parameters. The sentiment analysis yielded valuable insights into which sites will revive quickly; the analysis of the Tweets using sentiment analysis will help in predicting the revival of the tourism sector after the pandemic. The various algorithm scores will help in predicting the best ways to enhance the customer experience at various sites in the tourism sector. Further, the study will enable the tourism sector to design a revival strategy in the Covid-19 pandemic.

Keywords: Tourism; sentiment analysis; Covid-19 pandemic; logistic regression; Python; revival strategy

1 Introduction

SARS-COV2, a novel coronavirus, is the causative agent for Covid-19, which has driven a global pandemic and has caused great health concerns. This virus originated in 2019 in the animal wet market of Wuhan city in China, and it was highly contagious, spreading easily through person-to-person contact. By the beginning of 2020, the virus had spread across most parts of Europe, America, and the Asian sub-continent (Guan et al., 2020; Mao et al., 2020;

Mizumoto & Chowell, 2020; Shereen et al., 2020; Sohrabi et al., 2020). There has been a fight against this pandemic worldwide and nations have opted for lockdowns to curb its impact (Lau et al., 2020; Saadat et al., 2020; Setiati & Azwar, 2020; *The Lancet*, 2020). This major lockdown impacted economies drastically and every sector is working toward revival and trying to cope with the dent created by this pandemic. Out of all the sectors, the most severely impacted is the tourism sector, where a lot has to be done to come back to the position which existed before the pandemic (Bakar & Rosbi, 2020; Gössling et al., 2020; Higgins-Desbiolles, 2020; Lapointe, 2020; World Tourism Organization, 2020). The tourism sector, the primary focus of this study, was an exception: the role of IT was marginal and only acted as a supplementary service to the core service (Albeshr & Ahmad, 2015; Edghiem & Mouzoughi, 2018; Nafei, 2019; Sinclair-Maragh, 2014).

Covid-19 related to travel restrictions created an atmosphere of fear that badly hit the tourism sector across the world (Ulak, 2020). In the Indian economy, contrary to the countries where tourism contributes a significant percentage of their national incomes, the e-commerce industry, automobile sector, information technology and software services, rail system, and many other industries played a vital role in the contribution of national income in the third quarter of 2019-20, when the Covid-19 pandemic knocked at the door of India (Bansal et al., 2020). With the lockdown, a tremendous loss has been reported in the respective sectors, leading to a loss in the GDP (Osterholm & Olshaker, 2020). After the process of unlocking the economies, it has been predicted that domestic tourism in OECD countries, which accounts for around 75% of the tourism economy, is expected to recover more quickly (Sigala, 2020). The adverse impact of this pandemic has been felt on the entire tourism ecosystem and restarting and restoration of tourist destinations requires a collaborative approach. There is a need to build confidence among travelers, to stimulate them to visit a destination with the government's tourism-specific measures (Assaf & Scuderi, 2020). So, the main objective of this research is

*Corresponding author: Garima Malik, Amity Business School, Amity University, U.P. India, Email: gkmalik@amity.edu
Dharmendra Singh, Modern College of Business and Science, Muscat, Oman

to focus on sentiment analysis using Indian tourist tweets during Covid-19, using Python and the maximum likelihood method to determine the parameters.

2 Global Tourism and Pandemic

Global tourism has been widely exposed to many crises that have been influencing the tourism sector, especially from 2000 till today like the 9/11 terrorist attacks, the SARS outbreak (2003), the MERS outbreak, the global economic crisis in 2008/2009, and now Covid-19 as depicted in Figure 1. The relationships between pandemics and travel are central to understanding health security and global change (World Travel & Tourism Council, 2018). Globally the tourism sector has been considerably affected, with a downsize of 22% in the first quarter of 2020 which could further reach 60–80% in comparison to the travel figures of 2019 (World Tourism Organization, 2019).

3 Review of literature

Technological changes related to the internet, including smartphones and tablets, have revolutionised the tourism industry from a brick-and-mortar and person-to-person service industry to a heavily digitally supported and omnipresent travel service network. Travelers have access to online platforms to provide feedback and make recommendations for other travellers (Neidhardt et al., 2017; Yang et al., 2017; Ye et al., 2009). In addition to professional systems, online social media, such as Twitter, Instagram, Facebook, Four Square, Sina Weibo, and Google Plus, play a significant role in creating electronic word-of-mouth (Confente, 2015; García-Pablos et al., 2016; Leung et al., 2013; Phillips et al., 2017). Importantly, online social media, travel professional websites and platforms, and

blogs present inexpensive means to gather rich, authentic, and unsolicited data on travelers' opinions. Besides, sentiment can also be modelled by machines for automation, and integration across various applications (Choi et al., 2007; Rabanser & Ricci, 2005). Previous research used Twitter data to discover people's reactions to coronavirus-related issues such as lockdown, safety, job loss, mental health-related issues, and sentiments related to their resort visits (Barkur et al., 2020; Rajput et al., 2020; Cinelli et al., 2020; Samuel et al., 2020).

Sentiment analysis refers to the use of computational linguistics and natural language processing to analyze text and identify its subjective information. The sentiment analysis has the task of detecting and classifying opinions and feelings expressed in the text. Whilst there are many challenges to be solved before sentiment analysis programs attain expert human-level performance (Cambria et al., 2017), they give results that correlate positively with user ratings and so it is reasonable to use sentiment analysis scores when user ratings are absent (López-Barbosa et al., 2015). Sentiment analysis can also give additional insights into the important components of an offering. For example, an analysis of restaurant reviews found that context was important to diners in addition to those more widely recognised: food, service, price, and ambiance (Gan et al., 2017). Sentiment analysis has become a standard component of the social media analysis toolkit of marketers and customer relationship managers in large organisations (Hofer-Shall, 2010). These may purchase relevant data (e.g., Facebook posts mentioning a hotel) and gain access to a suite of sentiment analysis and other tools to analyze it in a platform like Pulsar (pulsarplatform.com). Understanding how sentiment analysis works is therefore important for those working in tourism and the wider marketing sector.

Tourists can benefit from sentiment analysis through computing systems that provide recommendations. The software can do this by extracting the key aspects of potential destinations (e.g., hotels), extracting reviews about these destinations, processing the reviews with aspect-based sentiment analysis, and then summarising the overall sentiment expressed about the customers' potential choices. A web user wishing to decide between three hotels might be presented with a graphic showing their key aspects (e.g., breakfast quality, cleanliness, view) and then be shown the average rating of customers on those aspects (Schmunk et al., 2013). In theory, this would save customers the task of reading reviews if, for example, they were only concerned with one aspect of a hotel (e.g., children's entertainment, bar, beach) and so overall ratings that considered all aspects of a hotel would not be rele-



Figure 1: International Tourist Arrivals By Region In Q1 2020

Source: Report of World Tourism Organization

vant. Whilst this would seem to be a useful service, there does not seem to be a successful commercial example so far. A promising but so far also apparently unsuccessful application is to overlay a map with icons representing tourism services, color-coded by sentiment (Cresci et al., 2014). Other applications have also attempted to exploit sentiment in tourist reviews to generate improved recommendations for them (e.g., Dong and Smyth, 2016). An unusual research application of sentiment analysis analysed the sentiment of tourism-related press coverage for countries and cities, generating an overall map of destinations by sentiment (Scharl et al., 2008). Even, as the global community faces the Covid-19 pandemic together, Twitter is helping people find reliable information, connect with others, and follow what's happening in real time.

4 Research Methodology

To conduct the sentiment analysis and opinion-mining process in which a piece of writing is analysed computationally positively, negatively, or neutrally, the tweets were fetched from Twitter using Python. The lockdown and work-from-home model led to a wide-scale increase in the socialisation of people on the internet via Facebook, Twitter, etc. These platforms gave people opportunities to express their feelings and sentiments. The tweets were fetched during March, April, and May when most of the economies were under lockdown due to the Covid-19 pandemic. Tweepy, TextBlob, and Corpora were installed on Python (Sai & Balachander, 2020), and with a registered app through a Twitter account, the tweets through Twitter API were fetched. Tweepy is a Python client for official Twitter API, TextBlob was used for cleaning and processing the fetched tweets, and the Sentiment classifier (D'Andrea et al., 2015) was then used to classify the tweets as positive, negative, or neutral (Da Silva et al., 2014; Pujazon-Zazik & Park, 2010; Shelar & Huang, 2018). Finally, logistic regression was incorporated to analyse the fetched tweets, since it explains the relationship between one dependent binary variable and one or more independent variables. In other words, it is a classification algorithm that predicts the occurrence of an event in binary as well as multiple predicting variables, which is categorical (Da Silva et al., 2014). The maximum likelihood method has been used while estimating the logistic regression, as it determines the parameters that are likely to produce the observed data and assumes mass function (Allison, 2008; Friedman et al., 2000; Pregibon, 1981; Spitznagel, 2007).

The sigmoid function, also called the S-curve, has been used as it can take up any real value and map it into values ranging from 0 to 1 (Lipovetsky, 2010; Pampel & Denney, 2011), as seen in Figure 2.

A confusion matrix has been used to evaluate the performance of the classified model. A confusion matrix is “a table that is often used to describe the performance of a classification model (or ‘classifier’) on a set of test data for which the true values are known. It allows the visualisation of the performance of an algorithm” (Schulz et al., 2011; Ting, 2017). See Figure 3 & Figure 2.

where,

TP is true positive
TN is true negative
FN is false negative
FP is false positive

Receiver operates the characteristic curve. Figure 3 is a “graphical plot that illustrates the diagnostic ability of a binary classifier system as its discrimination threshold is varied and the ROC curve is created by plotting the true positive rate against the false-positive rate at various threshold settings”. The higher the value under the area of the curve the better the predictive power of the model. Considering the different threshold value, the ROC curve is built, but which is never beyond 1 (Figure 4). The curve helps to determine the threshold values too (Bradley, 1997; Faraggi & Reiser, 2002).

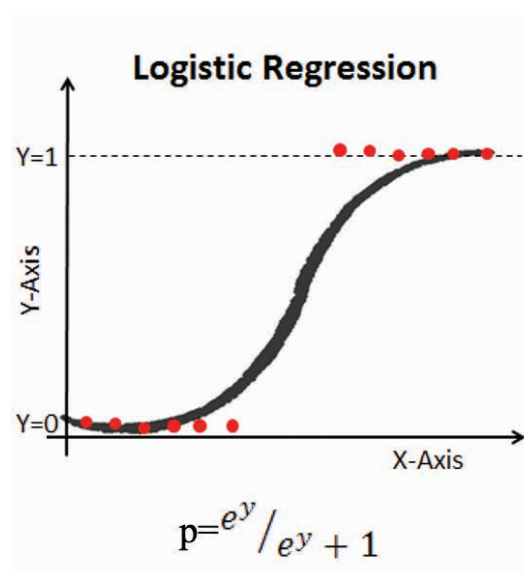


Figure 2: Equation1–Sigmoid Function

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

Figure 3: Confusion matrix

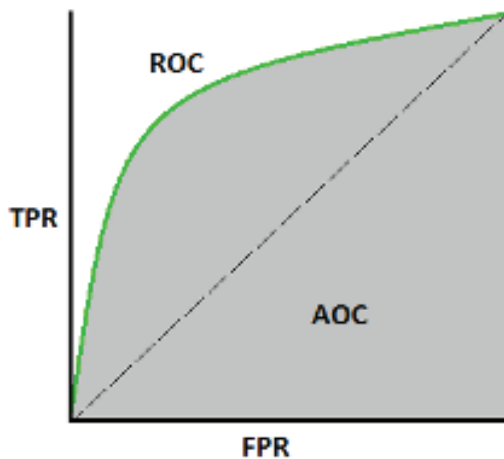


Figure 4: ROC Curve

5 Results and discussion

The tweets' data were processed using the various libraries of Python by installing the needed data to process the tweets. The data was checked for any missing values and outliers to get accurate predictions. From the *sklearn linear model* logistic regression was imported, as shown in Figure 5. The tweet data was run for logistic regression on the test and train data with the train size=.70, as shown in Figure 5.

The confusion matrix was created with the following values

```
array([[632, 86],
       [287, 190]])
```

```
File Edit View Insert Cell Name Windows Help
Python 3.0
In [35]: from sklearn.model_selection import train_test_split
In [36]: X_train, X_test, Y_train, Y_test = train_test_split(X, Y, train_size=0.7, random_state=100)
In [37]: from sklearn.linear_model import LogisticRegression
In [38]: lr = LogisticRegression()
In [39]: lr.fit(X_train, Y_train)
C:\Users\Nar\Anaconda2\lib\site-packages\sklearn\linear_model\logistic.py:402: FutureWarning: Default solver will be changed to 'lbfgs'. In #.22, Specify a solver to silence this warning.
FutureWarning:
Out[39]: LogisticRegression(C=1.0, class_weight=None, dual=False, fit_intercept=True,
intercept_scaling=1, l1_ratio=None, max_iter=100,
multi_class='warn', n_jobs=None, penalty='l2',
random_state=None, solver='warn', tol=0.0001, verbose=0,
warm_start=False)
In [40]: y_pred = lr.predict(X_test)
In [41]: y_pred
```

Figure 5: Logistic Regression on Python

The classification report depicted an accuracy of .687. The precision value (measures the ability of the classifier not to label as positive a sample that is negative) was .688 and the recall value (measures the proportion of positives that are correctly identified) was .398. while the F1 score was .50. F1 Score is the weighted average of precision and recall. Therefore, this score takes both false positives and false negatives into account. Intuitively it is not as easy to understand as accuracy, but F1 is usually more useful than accuracy, especially if you have an uneven class distribution. This F1 score was quite good compared to the training data, and the accuracy of positive and negative sentiments are quite a bit better than that of the neutral. In the results, 1 and 0 represent positive and negative sentiments respectively (Figures 6 and 7).

The receiver operating characteristic curve's value was higher under the area of the curve, signifying better predictive power of the model (Bradley, 1997; Faraggi & Reiser, 2002) as shown in Figure 8. All this clearly proved the reliability and validity of the model and the algorithms.

From Table 1 it can be understood that any data where the sentiment score was less than or equal to 0.50 is considered as a positive tweet and a score holding less than 0.49 has been considered negative. Looking at each keyword one by one, trekking has a positive sentiment with 0.06K likes and 97.71K re-tweets, which shows a high performance of the tourism site where trekking is available. Similarly, the keyword nature has a positive sentiment score: though there are no likes, it has been re-tweeted at 1238.4k. The performance is high, likewise, for the keywords travelgram, museums, and adventure tour. On the other hand, keywords spiritual tourism and nostalgia bear very low sentiment scores (0.12 and 0.36, respectively) which was considered as negative, though 5310.9K re-tweets showed low performance. Other keywords such as memorials, travel life, travel diaries, mountains, cruise, antitravel, holiday wanderlust, and beaches, performed moderately.

```

In [16]: from sklearn.metrics import confusion_matrix
In [16]: (confusion_matrix(Y_test,Y_pred))
Out[16]: array([[147,  80],
              [147, 204]])
In [17]: from sklearn.metrics import accuracy_score
In [17]: accuracy = accuracy_score(Y_test,Y_pred)
In [18]: accuracy
Out[18]: 0.507861887006105
In [19]: from sklearn.metrics import precision_score
In [19]: precision_score(Y_test,Y_pred)
Out[19]: 0.500481971039405
    
```

Figure 6: Confusion matrix on Python

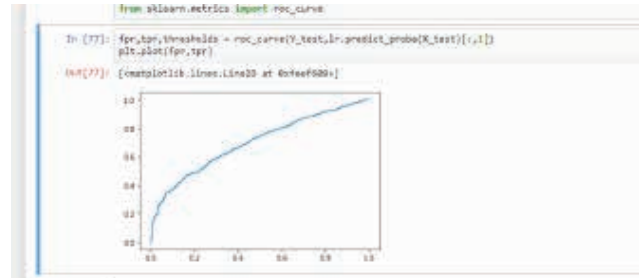


Figure 8: ROC curve on Python

```

Out[19]: 0.500481971039405
In [20]: from sklearn.metrics import precision_score
In [20]: precision_score(Y_test,Y_pred)
Out[20]: 0.500481971039405
In [21]: from sklearn.metrics import recall_score
In [21]: recall_score(Y_test,Y_pred)
Out[21]: 0.507861887006105
In [22]: from sklearn.metrics import F1_score
In [22]: F1_score(Y_test,Y_pred)
Out[22]: 0.504181123093981
    
```

Figure 7: Confusion matrix–accuracy on Python

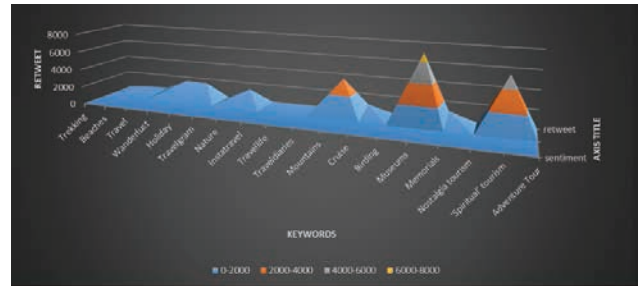


Figure 9: Surface Diagram of Sentiment Analysis

From the sentiment analysis, it can be clearly seen that nature, trekking adventure tours, and visits to the museum will be the most sought-after places by tourists once the economies open their boundaries for tourism. The tourism sector must thus focus on developing the sites where the tourist can experience adventure, and further sites where trekking can be possible should be

developed, so that trekking coupled with adventure can draw more tourists to these destinations. Museums have been the most liked and visited, as they create an experience and build adventure, like time traveling into the past. The tourist sector should consider this point, take care of museums well, and also create further experiences by creating IT-enabled services. The IT-enabled services

Table 1: Sentiment Analysis

Keywords	sentiment	likes	re-tweet	Performance
Trekking	0.52	0.06	97.71	High
Beaches	0.2	0.04	325.68	Moderate
Travel	0.34	0.01	439.42	Moderate
Wanderlust	0.24	0.81	1521.9	Moderate
Holiday	0.3	0.15	1455.17	Moderate
Travelgram	0.55	0.5	16.04	High
Nature	0.58	0	1238.4	High
Insta travel	0.44	0.59	5.83	Moderate
Travel life	0.43	1.43	9.33	Moderate
Travel diaries	0.44	0.58	6.78	Moderate
Mountains	0.38	0.04	3471.22	Moderate
Cruise	0.25	0.12	822.25	Moderate
Birding	.32	0.99	66.11	Moderate
Museums	0.51	0.03	6895.4	High
Memorials	0.27	0.85	965.5	Moderate
Nostalgia tourism	0.3636	2.367	3.41	Low
'Spiritual' tourism	0.12	0.12	5310.9	Low (many religious places are closed till date)
Adventure Tour	0.54	1.26	77.6	High

may create a virtual experience by adding more adventure to their tours. The next-ranked destinations, such as memorials, mountains, cruises, holiday wanderlust, and beaches, should be considered by the tourism sector to enhance customer experience according to the sentiment analysis depicted in Figure 9.

6 Conclusion

This paper analysed tweets in detail with the help of sentiment analysis to highlight their usefulness to the tourism sector. The analysis in particular presents the road map to the revival of the tourism sector after the pandemic. The various algorithm scores predicted which enhancements of the tourism sector will improve the customer experience. The sentiment analysis of the tweets revealed that the majority of the tweets were positive and there was significant enthusiasm for the revival of the sector. The findings suggest that the tourism sector could leverage the customer sentiment expressed in the tweets in order to understand their needs and expectations and to enhance their customer experience. The findings from the analysis could be used to plan promotional activities and to understand the sentiments of target customers. Overall, this analysis has provided valuable insights into the sentiments of the people regarding the tourism sector and the road map to its revival. The paper also highlighted the need for changes to existing tourism policies and regulations to ensure the safety of tourists and employees. The analysis also discussed the importance of social media marketing and how to use it to promote tourism in the post-Covid era. All in all, this paper provides valuable insights and a practical road map to the revival of the tourism industry.

7 Implications and limitations

The study contributes a lot to the tourism literature in many respects. Firstly, the study contributes methodologically, in the way in which the tweets have been fetched and the sentiment analysis has been conducted, which is an innovative and novel way of predicting the footprint in the various sites of tourism. Next, the literature contributes to the hospitality sector concerning the occupancy of hotels, resorts, and other properties (McCartney & Pao Cheng Pek, 2018; Philander & Zhong, 2016). This study also looks at other tourism sites, such as trekking sites, beaches, spiritual sites, and so on, which is an altogether new dimension that has been analysed for the tourism

sector. Further, the paucity of research on the tourism sector, especially in the Indian sub-continent, has also led to the motivation to take up this study, so the study will subsequently contribute to the generalisation of findings in the Asian sub-continent, because of the similarity in the culture. Again, the extant studies have undertaken sentiment analysis in the developed countries, and the scarcity of studies in the context of developing countries has stimulated us to undertake the study, so the study contributes well to the literature of the tourism sector in developing strategies to draw tourists of the fast-developing countries as major income in the tourism sector is contributed by the fast-developing countries (Crane et al., 1997; Lu et al., 2018; Mazumder et al., 2009; O'Sullivan & Jackson, 2002). Though it has been predicted that there will be a rapid revival of the tourism sector, which has been proved in this study, which sites of tourism have to be considered to enhance the customer experience has not been predicted by prior studies. This again is a substantial contribution towards the tourism literature.

From the tourism perspective, the sentiment analysis enables tourism organisations to keep track of tweets and scale these to useful opinions, which could help them in developing their revival strategies during the global pandemic. The sentiment analysis helps marketers to know the trajectory of public opinion, which would enable them to be ready with the right kind of service mix to get maximum advantage. Further, these analyses would also help managers in developing short-term strategies to cope with the pandemic. The sentiment analysis would also help managers in understanding and quantifying the customer's perception so that better strategies can be developed to revive the tourism sector.

With all its advantages, the paper also has certain limitations; it would be unethical not to mention them. The first limitation is data fetching; there are limited tweets that can be fetched at a certain point of time; some may be incorrect or biased, and it can create a data bubble. Not only this, more keywords such as vacation, honeymoon, and family trip, may widen the array of studies. Subsequently, the data being smaller in number can sometimes be miscategorised. Future studies can undertake bigger data sets, improving the content validity of these classifier algorithms and establishing guidelines for managers for sample size. Further studies can also be undertaken stressing the demographic profile of the Twitter accounts, which could widen the scope of the study by giving a real-time picture of the sentiment analysis. Future studies can also be considered for traditional market research tools, which can give new insights into the issues, though tweets can act as complementary tools to these research tools.

This may enhance the scope and draw more concrete results. Last but not least, the tweets are from a developing economy, India; results may become more concrete to generalise if more developing countries are included in the study and comparative analysis of these nations may give rise to new dimensions in tourism literature.

Bionotes

Garima Malik is a doctoral scholar of Marketing at Xavier School of Management, XLRI Jamshedpur, India. She has more than fifteen years of academic experience. She is also associated with Amity Business School, Noida, India. Her area of research includes tourism and hospitality, gamifications, IT-enabled services, and customer relationship management. Her research papers are published in various leading journals such as *Information Technology & Tourism International*, *Journal of Healthcare Management*, *Journal of Global Marketing*, *Innovative Marketing*, *Pacific Asia Journal of the Association for Information Systems*, *Journal of Science and Technology Policy Management*, among others. ORCID: 000-0002-3892-8299

Dharmendra Singh is an Assistant Professor at Modern College of Business and Science, Muscat (Oman). He possesses a rich professional experience of over 22 years in the field of finance. He holds a Ph.D. (Finance), professional certifications like Certified Financial Planner (CFP), associate diploma (life insurance) from Insurance Institute of India, and CFA. He has published a number of articles and research papers in various reputed international and refereed journals indexed in ABDC and Scopus. His research area includes banking, corporate finance, and financial markets. He also serves as a reviewer for select journals. He is a member of the Financial Planning Standard Board (FPSB), Indian Econometric Society, and Insurance Institute of India. ORCID: 0003-0966-6530

References

- [1] Albeshr, H., & Ahmad, S. Z. (2015). Service innovation by Dubai International Airport: the battle to remain competitive. *Esmerald Emerging Markets Case Studies*, 5(1), 1–23.
- [2] Allison, P. D. (2008, March 16-19). *Converge Failures in Logistic Regression* [Paper Presentation]. SAS Global Forum 2008, San Antonio, Texas.
- [3] Assaf, A., & Scuderi, R. (2020). COVID-19 and the recovery of the tourism industry. *Tourism Economics*, 26(5), 731–33.
- [4] Bakar, N. A., & Rosbi, S. (2020). Effect of Coronavirus disease (COVID-19) to tourism industry. *International Journal of Advanced Engineering Research and Science*, 7(4), 61–78.
- [5] Bansal, C. N. K., Sharma, S., & Gautam, A. (2020). A Study on Impact of COVID-19 a Global Pandemic on Indian Economy: With special Context to Goods and Service Tax. *Journal of Xidian University*, 14(3), 1809–18.
- [6] Barkur, G., Vibha, & Kamath, G. B. (2020). Sentiment analysis of nationwide lockdown due to COVID 19 outbreak: Evidence from India. *Asian Journal of Psychiatry*, 51(2020), 1–2.
- [7] Bradley, A. P. (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition*, 30(7), 1145–59.
- [8] Cambria, E., Poria, S., Gelbukh, A., & Thelwall, M. (2017). Sentiment Analysis Is a Big Suitcase. *IEEE Intelligent Systems*, 32(6), 74–80.
- [9] Choi, S., Lehto, X. Y., & Morrison, A. M. (2007). Destination image representation on the web: content analysis of Macao travel related websites. *Tourism Management*, 28(1), 118–29.
- [10] Cinelli, M., Quattrocchi, W., Galeazzi, A., Valensise, C. M., Brugnoli, E., Schmidt, A. L., Zola, P., Zollo, F., & Scala, A. (2020). The COVID-19 social media infodemic. *Scientific Reports*, 10, 1–10.
- [11] Confente, I. (2015). Twenty-Five Years of Word-of-Mouth Studies: A Critical Review of Tourism Research. *International Journal of Tourism Research*, 17, 613–24.
- [12] Crane, R., Danieri, A., & Harwood, S. (1997). The Contribution of Environmental Amenities to Low-income Housing: A Comparative Study of Bangkok and Jakarta. *Urban Studies*, 34(9), 1495–512.
- [13] Cresci, S., D'Errico, A., Gazzé, D., Lo Duca, A., Marchetti, A., & Tesconi, M. (2014, May 28-30). *Tour-pedia: a Web Application for Sentiment Visualization in Tourism Domain* [Paper Presentation]. 9th Language Resources and Evaluation Conference (LREC 2014), Reykjavik.
- [14] Da Silva, N. F. F., Hruschka, E. R., & Hruschka Jr, E. R. (2014). Tweet sentiment analysis with classifier ensembles. *Decision Support Systems*, 66, 170–79.
- [15] D'Andrea, A., Ferri, F., Grifoni, P., & Guzzo, T. (2015). Approaches, Tools and Applications for Sentiment Analysis Implementation. *International Journal of Computer Applications*, 125(3), 26–33.
- [16] Dong, R., & Smyth, B. (2016, October 21 – November 2). *Personalized Opinion-Based Recommendation*. *International conference on case-based reasoning* [Paper Presentation]. 24th International Conference on Case-Based Reasoning, Atlanta, GA, USA.
- [17] Edghiem, F., & Mouzoughi, Y. (2018). Knowledge-advanced innovative behaviour: a hospitality service perspective. *International Journal of Contemporary Hospitality Management*, 30(1), 197–216.
- [18] Faraggi, D., & Reiser, B. (2002). Estimation of the area under the ROC curve. *Statistics in Medicine*, 21(20), 3093–06.
- [19] Friedman, J., Hastie, T., & Tibshirani, R. (2000). Additive Logistic Regression: A Statistical View of Boosting. *The Annals of Statistics*, 28(2), 337–407.
- [20] Gan, Q., Ferns, B. H., Yu, Y., & Jin, L. (2017). A Text Mining and Multidimensional Sentiment Analysis of Online Restaurant Reviews. *Journal of Quality Assurance in Hospitality & Tourism*, 18(4), 465–92.

- [21] García-Pablos, A., Lo Duca, A., Cuadros, M., Linaza, M. T., & Marchetti, A. (2016). Correlating languages and sentiment analysis on the basis of text-based reviews. In A. Inversini & R. Schegg (Eds.), *Information and Communication Technologies in Tourism* (pp. 565-577). Springer International Publishing.
- [22] Gössling, S., Scott, D., & Hall, C. M. (2020). Pandemics, tourism and global change: a rapid assessment of COVID-19. *Journal of Sustainable Tourism*, 29(1), 1–20.
- [23] Guan, W.-j., Ni, Z.-y., Hu, Y., Liang, W.-h., Ou, C.-q., He, J.-x., Liu, L., Shan, H., Lei, C.-l., Hui, D. S. C., Du, B., Li, L.-j., Zeng, G., Yuen, K.-Y., Chen, R.-c., Tang, C.-l., Wang, T., Chen, P.-y., Xiang, J., Li, S.-y., Wang, J.-l., Liang, Z.-j., Peng, Y.-x., Wei, L., Liu, Y., Hu, Y.-h., Peng, P., Wang, J.-m., Liu, J.-y., Chen, Z., Li, G., Zheng, Z.-j., Qiu, S.-q., Luo, J., Ye, C.-j., Zhu, S.-y., & Zhong, N.-s. (2020). Clinical characteristics of coronavirus disease 2019 in China. *New England Journal of Medicine*, 382(18), 1708–20.
- [24] Higgins-Desbiolles, F. (2020). Socialising tourism for social and ecological justice after COVID-19. *Tourism Geographies*, 22(3), 610–23.
- [25] Hofer-Shell, Z. (2010). The Forrester wave: listening platforms, Forrester Research.
- [26] Lapointe, D. (2020). Reconnecting tourism after COVID-19: the paradox of alterity in tourism areas. *Tourism Geographies*, 22(3), 633–38.
- [27] Lau, H., Khosrawipour, V., Kocbach, P., Mikolajczyk, A., Schubert, J., Bania, J., & Khosrawipour, T. (2020). The positive impact of lockdown in Wuhan on containing the COVID-19 outbreak in China. *Journal of Travel Medicine*, 27(3), 1–7.
- [28] Leung, D., Law, R., van Hoof, H., & Buhalis, D. (2013). Social media in tourism and hospitality: A literature review. *Journal of Travel & Tourism Marketing*, 30(1–2), 3–22.
- [29] Lipovetsky, S. (2010). Double logistic curve in regression modeling. *Journal of Applied Statistics*, 37(11), 1785–93.
- [30] López-Barbosa, R. R., Sánchez-Alonso, S., Sicilia-Urban, M. A. (2015). Evaluating hotels rating prediction based on sentiment analysis services. *Aslib Journal of Information Management*, 67(4), 392–407.
- [31] Lu, C.-L., Chen, S.-T., & Kuo, H.-I. (2018). International tourism demand in Asia: before and after the economic crisis. *Asia Pacific Journal of Tourism Research*, 23(11), 1073–85.
- [32] Mao, L., Jin, H., Wang, M., Hu, Y., Chen, S., He, Q., Chang, J., Hong, C., Zhou, Y., Wang, D., Miao, X., Li, Y., & Hu, B. (2020). Neurologic Manifestations of Hospitalized Patients With Coronavirus Disease 2019 in Wuhan, China. *JAMA Neurology*, 77(6), 683–90.
- [33] Mazumder, M. N. H., Ahmed, E. M., & Al-Amin, A. Q. (2009). Does Tourism Contribute Significantly to the Malaysian Economy? Multiplier Analysis Using I-O Technique. *International Journal of Business and Management*, 4(7), 146–59.
- [34] McCartney, G., & Pao, R. P. C. (2018). An Examination of Sina Weibo Travel Blogs' Influence on Sentiment towards Hotel Accommodation in Macao. *Journal of China Tourism Research*, 14(2), 146–57.
- [35] Mizumoto, K., & Chowell, G. (2020). Estimating Risk for Death from Coronavirus Disease, China, January–February 2020. *Emerging Infectious Diseases*, 26(6), 1251–56.
- [36] Nafei, W. A. (2019). Organizational Ambidexterity as a Mediating Variable between Psychological Capital and Strategic Success: An Applied Study. *Case Studies Journal*, 8(11), 1–23.
- [37] Neidhardt, J., Rümmele, N., & Werthner, H. (2017). Predicting happiness: user interactions and sentiment analysis in an online travel forum. *Information Technology Tourism*, 17(16), 101–19.
- [38] Osterholm, M. T., & Olshaker, M. (2020, March 21). Facing covid-19 reality: A national lockdown is no cure. *The Washington Post*. <https://www.washingtonpost.com/opinions/2020/03/21/facing-covid-19-reality-national-lockdown-is-no-cure/>
- [39] O'Sullivan, D., & Jackson, M. J. (2002). Festival Tourism: A Contributor to Sustainable Local Economic Development?. *Journal of Sustainable Tourism*, 10(4), 325–42.
- [40] Pampel, F. C., & Denney, J. T. (2011). Cross-national sources of health inequality: education and tobacco use in the World Health Survey. *Demography*, 48(2), 653–74.
- [41] Philander, K., & Zhong, Y. (2016). Twitter sentiment analysis: Capturing sentiment from integrated resort tweets. *International Journal of Hospitality Management*, 55(2016), 16-24.
- [42] Phillips, P., Barnes, S., Zigan, K., & Schegg, R. (2017). Understanding the Impact of Online Reviews on Hotel Performance: An Empirical Analysis. *Journal of Travel Research*, 56(2), 235–49.
- [43] Pregibon, D. (1981). Logistic Regression Diagnostics. *The Annals of Statistics*, 9(4), 705–24.
- [44] Pujazon-Zazik, M., & Park, M. J. (2010). To tweet, or not to tweet: gender differences and potential positive and negative health outcomes of adolescents' social internet use. *American Journal of Men's Health*, 4(1), 77–85.
- [45] Rabanser, U., & Ricci, F. (2005). Recommender Systems: Do They Have a Viable Business Model in e-Tourism?. In A. J. Frew (Ed.), *Information and Communication Technologies in Tourism 2005* (pp. 160–71). Springer.
- [46] Rajput, N. K., Grover, B. A., & Rathi, V. K. (2020). Word frequency and sentiment analysis of twitter messages during coronavirus pandemic. *A Preprint*, 7(2), 13–26.
- [47] Saadat, S., Rawtani, D., & Hussain, C. M. (2020). Environmental perspective of COVID-19. *Science of The Total Environment*, 728, 1–6.
- [48] Sai, J. V., & Balachander, B. (2020). Sentimental Analysis of Twitter Data Using Tweepy and Textblob. *International Journal of Advanced Science and Technology*, 29(3), 10922–929.
- [49] Samuel, J., Ali, G. G. M. N., Rahman, M. M., Esawi, E., & Samuel, Y. (2020). Covid-19 public sentiment insights and machine learning for tweets classification. *Information*, 11(6), 1–22.
- [50] Scharl, A., Dickinger, A., & Weichselbraun, A. (2008). Analyzing News Media Coverage to Acquire and Structure Tourism Knowledge. *Information Technology Tourism* 10(1), 3–17.
- [51] Schmunk, S., Höpken, W., Fuchs, M., & Lexhagen, M. (2013). Sentiment Analysis: Extracting Decision-Relevant Knowledge from UGC. In Z. Xiang & I. Tussyadiah (Eds.), *Information and communication technologies in tourism 2014* (253–65). Springer.

- [52] Schulz, A., Damer, N., Fischer, M., & Stiefelhagen, R. (2011, August 30 – September 2). *Combined Head Localization and Head Pose Estimation for Video-Based Advanced Driver Assistance Systems* [Conference Session]. 33rd Annual Symposium of the German Association for Pattern Recognition in Conjunction With the GfKI 2011 Symposium, Frankfurt, German.
- [53] Setiati, S., & Azwar, M. K. (2020). COVID-19 and Indonesia. *Acta Medica Indonesiana*, 52(1), 84–89.
- [54] Shelar, A., & Huang, C.-Y. (2018, December 12-14). *Sentiment analysis of twitter data* [Paper Presentation]. 2018 International Conference on Computational Science and Computational Intelligence, Las Vegas.
- [55] Shereen, M. A., Khan, S., Kazmi, A., Bashir, N., & Siddique, R. (2020). COVID-19 infection: Emergence, transmission, and characteristics of human coronaviruses. *Journal of Advanced Research*, 24, 91–98.
- [56] Sigala, M. (2020). Tourism and COVID-19: Impacts and implications for advancing and resetting industry and research. *Journal of Business Research*, 117, 312–21.
- [57] Sinclair-Maragh, G. M. (2014). Resort-based or resource-based tourism? A case study of Jamaica. *Emerald Emerging Markets Case Studies*, 4(2), 1–19.
- [58] Sohrabi, C., Alsafi, Z., O'neill, N., Khan, M., Kerwan, A., Al-Jabir, A., Losifidis, C. & Agha, R. (2020). World Health Organization declares global emergency: A review of the 2019 novel coronavirus (COVID-19). *International Journal of Surgery*, 76, 71–76.
- [59] Spitznagel Jr, E. L. (2007). 6 Logistic Regression. *Handbook of Statistics*, 27(2007), 187–209.
- [60] *The Lancet*. (2020, April 25). *India under COVID-19 lockdown*. [https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(20\)30938-7/fulltext#%20](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)30938-7/fulltext#%20)
- [61] Ting, K. M. (2017). Confusion Matrix. *Encyclopedia of Machine Learning and Data Mining*, 260.
- [62] Ulak, N. (2020). A Preliminary Study of Novel Coronavirus Disease (COVID-19) Outbreak: A Pandemic Leading Crisis in Tourism Industry of Nepal. *Journal of Tourism and Hospitality Education*, 10, 108–31.
- [63] World Tourism Organization. (2019). International Tourism Highlights, 2019 Edition. <https://www.e-unwto.org/doi/pdf/10.18111/9789284421152>
- [64] World Tourism Organization. (2020). Tourism and COVID-19. In UNWTO Tourism Highlights.
- [65] World Travel & Tourism Council. (2018). *About*. <https://wtcc.org/>
- [66] Yang, Y., Mao, Z., & Tang, J. (2017). Understanding Guest Satisfaction with Urban Hotel Location. *Journal of Travel Research*, 57(1), 1–17.
- [67] Ye, Q., Zhang, Z., & Law, R. (2009). Sentiment classification of online reviews to travel destinations by supervised machine learning approaches. *Expert Systems with Applications*, 36, 6527–535.