Geojournal of Tourism and Geosites
ISSN 2065-0817, E-ISSN 2065-1198

Year XIX, vol. 66, no. 2spl, 2026, p.1603-1612
DOI 10.30892/gtg.662spl31-1792

DETECTING DECEPTION IN TOURISM AND GEOSITES: A SUPERVISED MACHINE LEARNING APPROACH

Richárd NÉMETH [1], Stephen AFRIFA [2], Vijayakumar VARADARAJAN [3], Ferenc ERDŐS [1*]

[1] Széchenyi István University, Department of Informatics, Faculty of Informatics and Electrical Engineering, Győr, Hungary; nemeth.richard@sze.hu (R.N.); erdosf@sze.hu (F.E.)
[2] North Carolina State University, Artificial Intelligence and Robotics Laboratory, Raleigh, North Carolina, United States; afrifastephen@tju.edu.cn (S.A.)
[3] University of Technology Sydney, Australia, Swiss School of Business and Management, Geneva, Switzerland; Vijayakumar. Varadarajan@uts.edu.au (V.V.)

Citation: Németh, R., Afrifa, S., Varadarajan, V., & Erdős, F. (2026). Detecting deception in tourism and geosites: A supervised machine learning approach. *Geojournal of Tourism and Geosites*, 66(2spl), 1603–1612. https://doi.org/10.30892/gtg.662spl31-1792

Abstract: The increasing circulation of misinformation and fake news in online environments poses growing and unpedictable risks to tourism destinations and geosites, where accurate communication is essential for visitor safety, heritage protection, and destination reputation. This problem is further intensified by the rapid and unchecked spread of digital content. This study develops a supervised machine-learning approach to identify deceptive content in tourism-related texts by combining domain-specific keywords extracted from the abstracts of the GeoJournal of Tourism and Geosites with a pre-labelled fake-news corpus. The text preprocessing phase utilized comprehensive Natural Language Processing techniques to clean the textual data, eliminate noise, and ensure high-quality feature extraction. Using Natural Language Processing techniques and five classification algorithms (logistic regression, Support Vector Machine, boosted decision tree, decision forest, and neural network), the models were trained and evaluated within the Microsoft Azure environment. Results show that simpler and more interpretable models, particularly boosted decision trees and logistic regression, outperform neural networks on the sparse, tourism-filtered dataset, achieving high accuracy, precision, and F1 scores. The performance evaluation was strictly validated through comprehensive confusion matrices and Receiver Operating Characteristic curves to compare overall classifier efficiency. The findings highlight the vulnerability of geosites to misinformation, especially regarding natural hazards, geomorphological features, and heritage narratives, and demonstrate the potential of AI-based tools to support reliable communication and decision-making in tourism management. The study provides a foundation for future multilingual and geospatially enhanced misinformation-detection systems that can strengthen the resilience and sustainability of tourism destinations and geosites.

Keywords: fake news, deception, classification, Azure, Machine Learning models, tourism, geosites, journal article

* * * * * *

INTRODUCTION

The tourism and geosite sector faces numerous challenges related to resource management and preserving natural and cultural heritage. These challenges cover a wide range of concerns, from the ecological impact of tourism activities and maintaining the quality of natural environments crucial for attracting visitors, to mitigating the risks posed by natural hazards to tourism infrastructure and visitor safety. For instance, issues such as visitor capacity management or preserving biodiversity in tourist areas are often discussed on social media platforms. These platforms, while serving as important channels for information dissemination about tourism and geosites, also present opportunities for the spread of misinformation, manipulation of public perception, and the proliferation of fake news.

Misinformation, disinformation, and fake news, which are increasingly prevalent in publications and social media posts in the field, have been shown to have a significant impact on tourism and the number of visitors, reputation, and heritage issues of geosites. Numerous scientific articles highlight that such information distortions affect tourists' attitudes, risk perception, and ultimately their choice of destination, and can distort tourists' decisions, potentially affecting the image and reputation of these regions and trust in tourism service providers in the long term. Among others, Tham & Chen (2022) point out that when fake news spreads rapidly, tourists often respond with distorted risk perception, avoiding certain destinations or choosing new locations based on a false sense of security. Misleading health, weather, or public safety information resulting from disinformation can cause serious long-term damage to a given region.

Misinformation, fraudulent "guerrilla marketing," or clickbait fake news at heritage sites can directly reduce visitor numbers, weaken local communities' trust in tourism operators, and even jeopardize the sustainability of the local

* Corresponding author

economy due to excessive or unjustified fears (Fedeli, 2020). Tourists' experiences have an effect on the reviews of platforms such as Tripadvisor or booking.com, distorting the ratings that tourists take into account (Akay, 2020; Dina & Juniarta, 2022). Although research on the impact of fake news is becoming increasingly widespread (see, for example Allcott & Gentzkow, 2017; Bastick, 2021; Karaş, 2024; Lazer et al., 2018), and there are numerous scientific works dealing specifically with the flow of authentic information, local heritage protection issues, cooperation between destination management organizations and tourism businesses, and many other issues in the context of tourism and geosites (Kaszás et al., 2025; Vasist & Krishnan, 2023), there is currently a lack of work supporting the detection of fake news in this field using machine learning algorithms. This paper aims to fill this gap by developing a supervised machine learning model for identifying fake news related to this domain.

A specialized dataset is constructed for model training, comprising labeled social media posts containing keywords relevant to the Geojournal of Tourism and Geosites' (GTG) thematic scope, focusing on resource management and heritage preservation. The training resulted in a complex model that integrates learning processes, machine learning algorithms, and best practices, using various data preparation, text processing, text analysis, and hyperparameter optimization techniques. It is important to clarify that our model does not classify GTG articles themselves. Instead, we extract a domain-specific keyword set from GTG abstracts and apply it to a global, not tourism-specific dataset to isolate social media post items typically relevant to tourism and geosites. The model therefore evaluates tourism-related fake posts within a general fake-news corpus. The primary metric for evaluating the model is prediction accuracy, but we also compared the reliability of the results with additional mathematical-statistical metrics such as precision, recall, and F1 score.

The aim of the research was to improve the reliability of online information about tourism and geosites by providing a mechanism for detecting and mitigating the propagation of misinformation. The intended outcome is to facilitate evidence-based decision-making, promote critical evaluation of online content, and increase public understanding of the complex factors influencing the management and preservation of tourism destinations and geosites.

GEOGRAPHICAL DIMENSIONS OF FALSE NEWS IN TOURISM AND GEOSITES

Cyberspace, especially social media, has become the primary source of information today, but unfortunately, it also provides opportunities for deception, manipulation, and influencing public opinion through questionable or outright false news appearing on online platforms. The misleading nature of this information overload threatens not only the safety of individuals but also the functioning of society, the existing order, and trust in the media (Idiongo, 2024).

The concept of fake news is highly complex and lacks a universally accepted definition (Gelfert, 2018), as it overlaps with other information disorders such as misinformation (false information shared without intent to harm, usually based on a misunderstanding – see Gelfert, 2018; Cook et al., 2015) and disinformation (false information deliberately created to deceive, often driven by political or financial goals) (Lazer et al., 2018). The most accepted approaches define fake news by intent, content authenticity, and dissemination methods (Walters, 2018).

The main threat of such information lies in its difficulty to verify and its strong influence on public opinion. This phenomenon is particularly critical in sectors where real-time, reliable information is crucial, such as tourism, where tourists often rely on online sources when assessing risks and natural hazards or searching for destinations (Akay, 2020).

It can be concluded that the rapid spread of misinformation and fake news has become a widespread challenge in many areas of society, and tourism is no exception. In tourism, and especially in specialized areas such as geotourism and geosites, misinformation can have serious consequences. Just think about false information on natural hazards, caves, geoparks, geomorphological features, and archaeological sites. In this chapter, we examine the spatial variability of these phenomena, highlighting how distorted or false information affects visitors' value judgments, safety, and the preservation of heritage sites. We illustrate these topics with regional examples, including volcano tourism, myths about archaeological sites, and conspiracy theories about geomorphological specialities.

Geotourism focuses on geological heritage sites such as caves, volcanoes, geoparks, and other unique geomorphological structures, and offers numerous opportunities for regions with natural monuments of regional, national, or international significance (Štrba, 2015). These sites attract visitors because of their scientific, cultural, and aesthetic value. However, the specialized and sometimes esoteric knowledge of the history and characteristics of these places makes them particularly vulnerable to misinformation (Kaszás et al., 2025). False or exaggerated information about geological hazards, rarities, or supernatural phenomena often spreads through social media, travel blogs, and informal word of mouth, creating a distorted image among tourists (Hilberts et al., 2024).

Misinformation about natural hazards has direct safety implications. For example, underestimating volcanic activity or seismic risks at certain tourist destinations can expose tourists to unforeseen dangers, jeopardizing visitor safety and undermining trust in travel organizers (Bonadonna et al., 2022).

Conversely, exaggerating threats can deter visitors, which can impact local economies that rely on geotourism. The problem is similarly acute in cave tourism, where myths or unfounded claims about hidden treasures, paranormal phenomena, or mysterious stories can lead to an increase in unauthorized visits and physical damage to fragile ecosystems. Sites such as the volcanoes of Vesuvius (Italy), Merapi (Indonesia), and Kīlauea (Hawaii) are prime targets for misinformation and conspiracy theories, ranging from denial of volcanic activity to completely irrational exaggerations (Brantley et al., 2019; Barberi et al., 2008). Local authorities are combating misinformation with scientific information campaigns, but the rapid spread of false claims on social media remains a difficult problem to address.

Many archaeological and geological sites, such as Stonehenge or the Nazca Lines, are victims of widespread myths and conspiracy theories (Holtorf, 2005), suggesting, for example, extraterrestrial origins or ancient high-tech civilizations.

While these narratives increase tourist interest, they often obscure fact-based heritage information, leading to a number of challenges. Narratives associated with unique geographical features undoubtedly create false impressions in visitors, leading to environmental destruction and unauthorized actions. The spread of false or misleading information has a serious impact on both conservation efforts and the reputation of geological sites. Misinformation can lead to inappropriate visitor behavior, illegal exploitation, or development projects that are incompatible with the preservation of the site. In addition, distorted narratives affect the image of the destination, negatively impacting visitor numbers or resulting in unsustainable situations that local infrastructure cannot accommodate, exacerbating tourism overload (Drápela, 2022).

THE ROLE OF MACHINE LEARNING IN TOURISM AND GEOSITES, AND IN IDENTIFYING FAKE NEWS

Parallel to the transformation of the digital information environment, the development of artificial intelligence has created new risks and opportunities. Artificial neural networks were inspired by early models describing how human neurons function (see Hodgkin & Huxley, 1952). After basic threshold-based neuron concepts, Rosenblatt's Perceptron provided the first learning-capable model and laid the foundation for later neural network architectures (Rosenblatt, 1958). Multi-layer perceptrons were already capable of modeling complex, nonlinear functions (Bishop, 2006), while the introduction of the backpropagation algorithm (Rumelhart et al., 1986) made the training of deep networks significantly more efficient. Figure 1 shows the main learning paradigms of Machine Learning. In the 21st century, high computing power and the availability of large amounts of data have enabled the rise of deep learning.

Convolutional neural networks (CNNs) have revolutionized computer vision (LeCun et al., 1998; Krizhevsky et al., 2012) and now play a key role in tourism-related applications, such as the automated analysis of satellite images, visual data indicating volcanic activity, or images taken by tourists. Recurrent neural networks (RNNs) have revolutionized text data processing (Greff et al., 2017), enabling the analysis of online reviews and tourist opinions. However, RNNs were effective in processing sequential data, but they struggled with long-term dependencies (Cho et al., 2014) – this problem was solved by LSTM and GRU architectures, which provided more stable learning for longer sequences (Hochreiter & Schmidhuber, 1997; Greff et al., 2017). The real paradigm shift came with the advent of Transformer models (Vaswani et al., 2017), which are able to map the relationships between large amounts of text through self-attention mechanisms, significantly improving context interpretation. Transformer-based models – such as BERT (Devlin et al., 2018) or the generative GPT family – are now widely used in disinformation detection, content classification, and automatic risk reporting. The power of modern neural networks lies in their ability to automatically recognize complex patterns and make highly accurate predictions (however, more complex models are not necessarily more accurate – see Rudin, 2019), thus playing an important role in enabling the tourism sector to quickly and efficiently verify false information. Based on these developments, the main objective of our research was to reliably identify these type of contents. We aim to achieve this using models based on machine learning algorithms that are capable of classifying news items and social media posts into true or false groups based on their content.

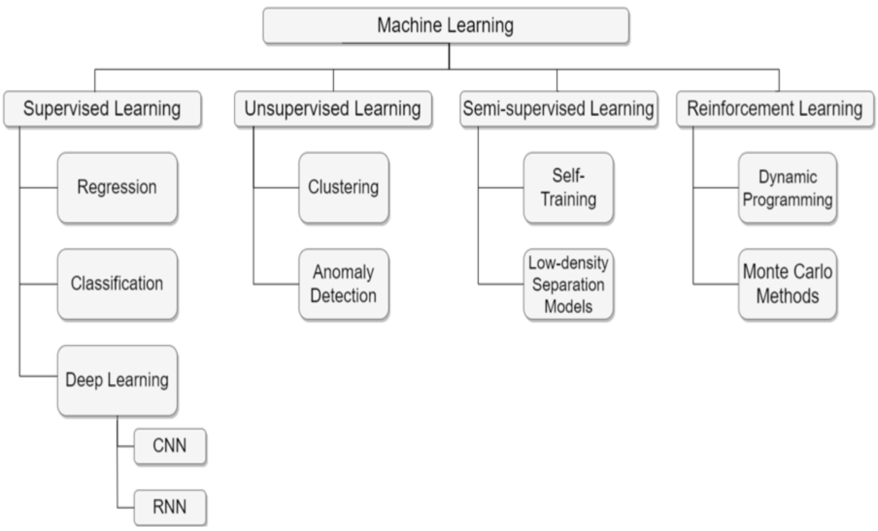

Figure 1. The main learning paradigms of Machine Learning (Source: made by the Authors)

METHODOLOGY

The aim of the research was to develop a machine learning model capable of identifying fake news within the target domain with high accuracy and reasonable computational requirements. Labeled data was available for training the models, so we applied supervised learning (SL) methodology in our research. Supervised learning is one of the fundamental paradigms of machine learning, where, during the learning process, the system learns the outputs associated with the inputs (labeled data), enabling it to make accurate predictions for new data (Goodfellow et al., 2016).

The two types of SL approaches are regression and classification, of which we focused on binary classification in the present study, the aim of which is to classify the phenomena under investigation into two discrete categories (true and false news). Our goal was to identify a model that could provide the best prediction while still having acceptable computational capacity requirements and runtime. The algorithms used ranged from simple, easy-to-understand models to architectures with high predictive potential but more complex and significantly higher computational requirements. The selected algorithms can all be integrated into the Azure Machine Learning environment, ensuring modularity, flexibility, and reproducibility. First, we tested two-class logistic regression suitable for solving binary classification problems (Kravets et al., 2024). The algorithm is based on a linear predictor and provides detailed probability values for class membership. This model works well for linear decisions, but it is not always able to recognize complex decision boundaries. In such cases, the use of decision trees is justified, which structure the model using hierarchical conditions, minimize the loss function, and are capable of strong prediction (Tanha et al., 2020).

The boosted tree we use can improve the model's performance iteratively by correcting the errors of previous runs. Since a single decision tree is prone to overfitting, we also used a decision forest, which trains multiple trees in parallel and provides more stable and robust results through a voting mechanism (Biau, 2012). Due to the relatively diverse dataset, we also examined the applicability of Support Vector Machine (SVM), which is one of the most commonly used classification methods due to its good generalization capabilities (Cervantes et al., 2020).

The approach is based on the principle of maximum marginal separation and is particularly effective for high-dimensional classification tasks. Finally, we also tested classification with neural networks, which, thanks to their layered architecture, are capable of modeling complex, nonlinear patterns, but require large data sets and significant computational resources. Figure 2 illustrates the methodology used during the research process.

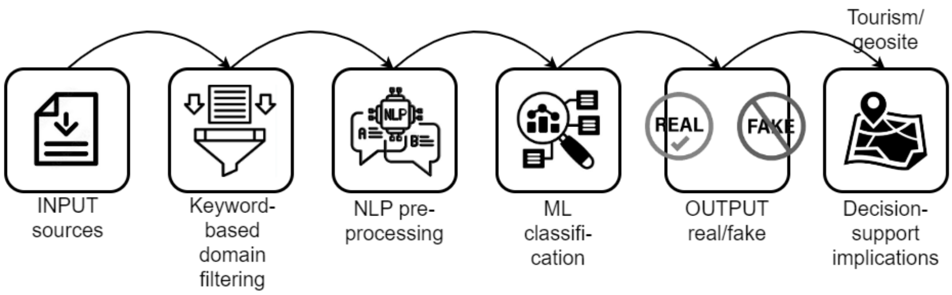

Figure 2. Overview of the methodology used in this study, including keyword extraction, data cleaning, N-gram text vectorization, model training, and evaluation (Source: authors)

Beyond academic use, the model could be applied in the future to predict the veracity of content appearing on social media platforms, thereby contributing to reducing the spread of misinformation. This chapter provides a detailed description of the data preparation process, the data cleaning procedures used, the methodology for dividing the data into training and test sets, and the machine learning algorithms used in the study. The models were built using the Azure Machine Learning cloud-based environment, which supports the creation of structures, the reproducibility of experiments, and the comparative evaluation of the effectiveness of different algorithms with its extensive technological toolkit.

Data set and data cleaning

The validation and training sets were provided by the ISOT Fake News detection dataset (Fake and Real News Data), which is the work of Subhadeep (2022) and a frequently used database in the field of fake news research. Although the True-False labeling of the data is accurate, the nature of the news used in the database and the unknown context of social media posts somewhat limit its usability, as does the fact that it covers a limited time interval. At the same time, the news items are global in nature, come from countless areas, and the dataset contains a particularly large amount of data (23,502 fake and 21,417 real news items), making it suitable for classifying news. The keywords were collected from the abstracts of the Geojournal of Tourism and Geosites (GTG) and organized into an Excel spreadsheet, from the very first issue (2008/Volume I) to the first issue of 2025. A total of 1003 articles out of 1374 were included in the table (the rest were not suitable for processing for various reasons, e.g., review articles, manuscript submissions, notes, etc.). Table 1 below shows the keywords collected in relation to the area under study and their frequency of occurrence.

Table 1. Most common keywords from Geojournal of Tourism and Geosites abstracts (2008/1 – 2025/1)

Rank	Keyword	Occurrence	Rank	Keyword	Occurrence
1	tourism	3379	**11**	geomorphological	644
2	tourist	1751	**12**	geomorphosite	61
3	hotel	335	**13**	beach	59
4	travel	280	**14**	traveller	58
5	hospitality	137	**15**	museum	57
6	geosite	87	**16**	gastronomic	57
7	accommodation	77	**17**	festival	55
8	restaurant	77	**18**	archaeological	55
9	tour	65	**19**	geopark	45
10	cave	64	**20**	resort	43

Our keyword corpus derived from GTG abstracts was only used to extract domain-specific keywords from the ISOT dataset. As a result of keyword filtering, 1,106 topic-relevant entries remained from the original dataset, which proved to be sufficient for training the models effectively. To validate the relevance of the keyword filtering, a random sample of 50 entries was manually inspected, revealing that 96% of the content was directly relevant to travel, hospitality, or geospatial context.

The individual records were already labeled as true or false, so the model evaluation process consisted of two phases: evaluating the results and comparing the predictions with the actual values.

Preprocessing

The first step in compiling the models was to clean up the data in order to obtain uniform, easy-to-manage and processable records. To do this, we had to correct the mislabeled entries and merge the news “Title” fields with the content. Empty fields, unreadable characters, and delimiters were removed. The resulting .csv file was still not handled correctly by Azure, which identified four classes instead of two due to hidden characters and the heterogeneous structure. Removing non-printable characters solved this problem. During preprocessing, stop words, special characters and multiple spaces were removed, lemmatization was performed, plurals were converted to singulars, numbers were filtered, and hyphenated words were separated. This was followed by lowercase conversion, British and American spelling harmonization, sentence recognition, deletion of unnecessary duplicate characters, and finally, tokenization.

The resulting cleaned dataset was divided into a 20% validation dataset and an 80% training dataset (so, a standard 80-20 split was used). This resulted in a .csv file containing a total of 2 columns and 1,106 rows. To generate text attributes, we used N-gram feature extraction, which has proven to be the most effective in our previous research. As its name suggests, an N-gram is a sequence of "n" elements (usually words or syllables) that the algorithm assigns a specific value to, thereby facilitating pattern recognition (Joshi et al., 2025). The goal is to identify common expressions and structures for detecting fake news. During the procedure, in accordance with our previous findings, the N-gram size was set to 3, the weighting function used a binary approach, the minimum word length was 3, and the maximum was 25 characters.

Training the models

To find the best solution, we tested five different machine learning algorithms, each with numerous different parameters, in order to produce the most accurate model. The algorithms used were as follows:

- Two-class logistic regression: During model training, different L2 regularization weights (strengths) were used, ranging from mild (0.01) to stronger penalties, which forced the model to simplify more. The higher the L2 value, the more the model penalizes excessive weights, resulting in a more general but less flexible model. The optimization tolerance value was tested in several steps from 1.00001 to 0.00000001. Here, the smaller the value, the more accurate the result, but at a certain point, the computational requirements become too high – the goal was to find the value where further refinement does not result in significantly more accurate results.
- Two-class Support Vector Machine (SVM): In the case of the model, the number of iterations was tested between 1 and 100 in three steps. This determined how many times the model could attempt to refine the decision boundary. The lambda parameter (0.00001; 0.0001; 0.01; 0.1) specified as the regularization parameter determined the model's willingness to compromise, i.e., it sought to strike a balance between the accuracy of the data points and the model's generalization ability.
- Two-class boosted decision tree: The complexity of the trees was set using the maximum number of leaves per tree parameter; here, it was run with 2, 8, 32, and 128 leaves to find out how many decision outputs could be generated in a tree and which setting would yield better performance. Overfitting was avoided by setting the minimum number of samples per leaf node parameter, where the number of elements varied between 1, 10, and 50. The learning rate (0.025, 0.05, 0.1, 0.2, and 0.4) determines how quickly the new tree can correct previous errors in the model. In general, the lower the learning rate, the more stable the result, but the slower the process.
- Two-class decision forest: In the model, the number of trees was varied using 1, 8, and 32 decision trees. Although a single tree does not constitute an ensemble method in the strict sense, it provides a meaningful baseline for performance comparison. Increasing the number of trees generally improves predictive accuracy and robustness due to the reduction of model variance; however, it also results in higher computational costs during both training and evaluation. The complexity (maximum depth) of each tree was varied between 1 and 64. This parameter determines the maximum number of questions that can be asked before a mandatory decision is made. Model training employed bootstrap aggregating (bagging) resampling, ensuring that each tree was trained on a different bootstrapped subset of the original dataset. This technique promotes model diversity and mitigates overfitting by decreasing sensitivity to noise in the training data.
- Two-class neural network: The network contained 100 hidden neurons to recognize and learn correlations (at an appropriate speed, avoiding the risk of overfitting). The learning rate was varied between 0.1 and 0.4, searching for the optimum between a slow and stable, and a fast but less reliable setting, in order to find out what step size is most effective for learning. Training was performed on the entire dataset in 20, 40, 80, and 160 iterations, respectively, searching for the balance that avoids both underfitting and overfitting. The training data was shuffled before each epoch.

RESULTS

Table 2 below shows the top-performing runs for each machine learning algorithm, i.e., the models with the best configurations, based on the metrics of the evaluation executed in Azure.

The configuration of hyperparameters strongly influences the efficiency of an ML algorithm. The process of finding the specific set of hyperparameter values that leads to the best estimated generalization performance is called hyperparameter optimization (HPO) (Bischl et al., 2022). Overall, it can be concluded that not all models achieved better results with HPO enabled and N-grams applied, as the predictive performance was not always positively affected.

Table 2. Prediction accuracy achieved using different models and components in Azure

Component	**Model1**	**Model2**	**Model3**	**Model4**	**Model5**
Machine Learning algorithm	Logistic regression	Support Vector Machine	**Boosted decision tree**	Decision forest	Neural network
Accuracy	0.955	0.937	**0.959**	0.864	0.95
Precision	0.97	0.949	**0.97**	0.831	0.963
Recall	0.956	0.949	**0.963**	0.978	0.956
F1 Score	0.963	0.949	**0.967**	0.899	0.959
AUC	0.993	0.977	**0.988**	0.967	0.993

The most effective algorithm among the various approaches tested was Boosted Decision Tree, resulting in an Accuracy of 0.959, a Precision of 0.97, and was also the best in F1 Score, that is 0.967. To evaluate the performance of the most effective model, we applied two widely used tools in the field of binary classification: the confusion matrix to evaluate the accuracy of classifications (True Positive, False Positive, True Negative, False Negative) (Yang & Berdine, 2017), and the Receiver Operating Characteristics (ROC) curve, which illustrates the relationship between the True Positive Rate (TPR, or sensitivity) and the False Positive Rate (FPR, or 1–specificity) across all possible classification thresholds (Gocoglu et al., 2024). A closer examination of the confusion matrix in Figure 3 reveals that there were only 5 False Negatives and 4 False Positives out of a total of 1,106 entries. These results indicate an exceptionally high level of model performance and demonstrate its outstanding predictive accuracy. This can also be seen from the ROC curve in Figure 4.

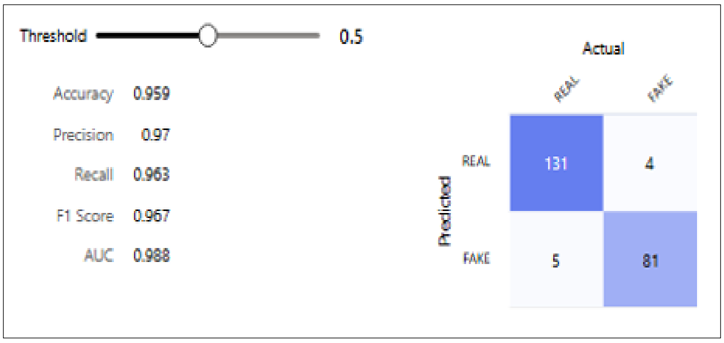

Figure 3. Metrics and the confusion matrix of the most effective model based on Boosted Decision Tree

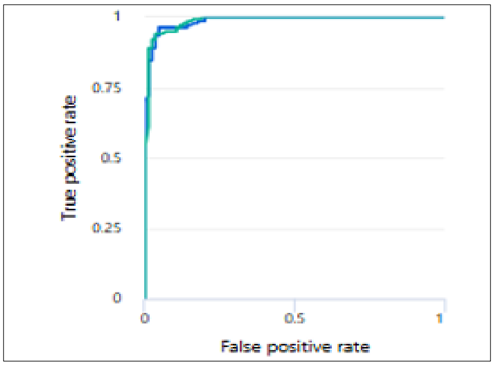

Figure 4. ROC Curve of the most effective model

All machine learning algorithms were also tested in a hyperparameter tuning variant (HPO enabled), which, interestingly, only led to better results in the case of decision forest-based models. In fact, in the case of boosted decision trees, the indicators even slightly deteriorated with the use of HPO. (For an explanation, see the Discussions section.) In the present research, the default "Single Parameter" setting proved to be a strong starting point, performing exceptionally well on the dataset in the case of the boosted decision tree algorithm, where the search range (maximum number of leaves per tree) used a setting close to the optimal one (20 leaves), while the HPO search range worked with 2, 8, 32, and 128 leaves, which resulted in models that were either too simple (2 and 8 leaves) or more complex than necessary (32 and 128 leaves). Similar HPO-anomaly results have been confirmed by other relevant articles, highlighting that in such cases, default settings (without HPO) often yield the best results (Mantovani et al., 2024), especially if the settings are well-suited to the dataset (Meaney et al., 2025). The algorithm, based on a neural network, which performs very well in most areas of text processing, only achieved moderate performance in this comparison. This result is more understandable, as numerous studies highlight that complex deep models perform worse than traditional base models in text classification tasks. The results are supported by the findings of Yan & Wu (2018), which indicate that complex models are often inappropriate when training data is limited.

Simpler models, such as logistic regression or tree-based models, often perform better when applied to small to medium-sized data sets. Si et al., 2017 also confirm in their study that the use of boosted decision trees in multi-class tasks with sparse document matrices leads to smaller models and better prediction times, with similar performance.

DISCUSSION

During the research, we tested five different artificial intelligence-based machine learning algorithms (logistic regression, support vector machine, boosted decision tree, decision forest, and neural network) with the aim of finding the most effective model for classifying fake news (True vs False) based on Accuracy, Precision, Recall, F1 score, and AUC. Based on the results, boosted decision trees and logistic regression showed the highest accuracy (0.959 and 0.955), and the best F1 score (0.967 and 0.963), while neural networks also approached these values (Accuracy 0.950), indicating that although they are more complex, more traditional models can also be highly effective on given databases (Cervantes et al., 2020). The decision forest proved to be the weakest (Accuracy = 0.864; F1 = 0.899).

The high AUC values (0.977–0.993) indicate that all models adequately separated the True and False categories. Contrary to expectations, in the case of boosted decision trees, the use of HPO worsened the results. At first glance, it seems somewhat surprising that an optimization process would produce results contrary to those expected.

The reason for this phenomenon can primarily be traced back to the fact that excessive optimization of validation errors arising from smaller data sets can lead to overfitting at the HPO level. As Schneider et al., 2025 also noted in their relevant study, due to over-tuning, the HPO setting considered to be the optimal may perform less well in critical situations than if we had simply used the default configuration. So, due to excessive searching, the model does not find a true global optimum, but rather random validation minima, which leads to further errors.

As we pointed out in the Results section, the performance of neural networks lagged slightly behind that of the best models, even though they are more complex. The main reason for this is that N-gram-based representation is a classic text processing technique that creates a very wide but sparse dataset – that is, there are relatively many columns, but most columns have a value of zero for a given text. Models such as logistic regression and boosted decision trees perform very well in handling this type of high-dimensional, sparse data (Abramovich & Grinshtein, 2019; Omar et al., 2024). Thus, in text classification, it is not surprising – as confirmed by the present study – that when using N-grams, these have proved to be the strongest models.

The main advantage of neural networks in Natural Language Processing (NLP) is usually seen when they learn the features from the raw text themselves e.g. using word embeddings, RNNs, or LSTM models (LeCun et al., 2015). However, to function truly effectively, they require large amounts of data to learn the more complex patterns that they are particularly good at recognizing. With relatively small datasets (the database consisted of 1006 records, 80% of which, approximately 805 training samples, were used for training), the neural network can easily overfit or fail to learn the patterns efficiently.

From a practical perspective, the results of this study also have significant operational implications for tourism managers and geosite administrators. The high-performing machine learning models demonstrated here could support real-time monitoring systems that automatically flag misinformation relevant to natural hazards, heritage interpretations, and destination safety—topics that frequently shape visitor behaviour. For instance, early detection of deceptive claims about volcanic threats, cave collapses, restricted access, or controversial archaeological findings would allow authorities to respond quickly with verified information, thereby stabilising visitor flows and preventing reputational damage.

The model could also assist DMOs in identifying coordinated misinformation campaigns targeting local tourism businesses or disseminating misleading sustainability claims. Such automated tools would help organisations prioritise communication efforts, enhance transparency, and protect geosites from both unnecessary alarmism and harmful over-promotion. By incorporating AI-based detection mechanisms into routine digital communication strategies, tourism stakeholders can strengthen resilience against misinformation and ensure that decision-making processes are informed by credible, timely, and domain-specific data.

Although the used dataset did not contain geolocation attributes, the keyword-filtered tourism-related fake news identified by the model frequently overlaps with typical geosite misinformation themes such as natural hazards, archaeological myths, or geomorphological exaggerations. This supports the assumption that automated detection tools could be adapted later for spatially explicit misinformation monitoring.

CONCLUSION

In summary, it can be concluded that this study has made a domain-specific contribution to the scientific discourse on the detection of fake news spreading on social media. The research examined the possibility of identifying fake news from a new perspective, in the context of news related to tourism and geosites.

This approach provided an opportunity to explore how information of uncertain veracity appears and spreads in a field where reliable communication of natural and cultural values is of paramount importance. During the study, the performance of five machine learning models – logistic regression, support vector machine, decision forest, boosted decision tree, and neural network – was compared in a Natural Language Processing-based binary classification task. The analysis revealed in detail the characteristics and differences in the predictive performance and generalizability of each algorithm, as well as the extent to which these models are able to reliably identify fake news in the narrower disciplinary context examined. The results confirm the findings of previous research, namely that in the case of relatively small, sparse, or unbalanced datasets, simpler, interpretable models – such as logistic regression and tree-based algorithms – are often more effective at recognizing relevant patterns than more complex neural networks.

The comparative analysis resulted in a model with high predictive accuracy, which may be suitable in practice for classifying the factual content of social media platforms related to tourism and geosites. Further development of the model could form the basis of an artificial intelligence-based system that is potentially capable of classifying tourism-related news appearing on different platforms, thereby promoting authentic online communication and more reliable information. This model shows particular promise for monitoring high-risk geosites, where misinformation regarding seismic activity or accessibility can lead to immediate safety hazards. Identifying fake news can directly contribute to more effective decision-making in the areas of regional development and professional practice, as it enables the rapid filtering of misinformation related to tourism, recognizes malicious or inauthentic reviews or fake posts by travel influencers, and helps filter out truly relevant information (Tham & Chen, 2022).

In this way, it supports the work of DMOs, regional decision-makers, and tourism professionals by enabling them to rely on reliable data and apply better marketing and crisis communication strategies. For instance, DMOs could incorporate automated misinformation detection into their communication workflows, enabling earlier intervention when harmful narratives emerge about hazards, accessibility, or heritage sites.

In addition, real, reliable information has a positive influence on tourists' decision-making and choice of location (Pagaldiviti, 2025). Automated fake news detection systems raise the reliability of regional management information to a higher level through real-time monitoring and help prevent economic or social damage caused by fake news. People are increasingly relying on social media for news content (Gwon et al., 2024). If they can read reliable news, it supports the sustainability of tourism by strengthening consumer confidence. An additional benefit is that by using validated, filtered information and automating decision preparation, professional decision-making can be faster, data-driven, and fact-based (Kovari, 2024), thus having a direct impact on tourism performance and the protection of geosites.

As a future research direction, the use of more advanced deep learning models, such as BERT and large language models (LLMs), offers promising opportunities for even more accurate recognition of the complex context of natural language. In addition to expanding the range of models used, the range of data can also be expanded – it may be worthwhile to integrate the original dataset with additional sources. Furthermore, the development of multilingual and geospatial databases can further increase the applicability of models in the changing environment of global tourism. The possible integration of such a multilayer data model could open up another new dimension in the analysis of destination-specific fake content.

LIMITATIONS

It is important to note that the trained models may have limitations that can affect their performance and generalizability to other datasets. The data used for training (Subhadeep, 2022) covers only a portion of the posts on social media platforms and a specific interval, which may lead to bias in the predictions.

In addition, the input dataset contained only English-language texts, so it does not reflect the complexity of international, multilingual social media. However, adapting the model to other languages would require further fine-tuning, as linguistic structures and n-gram patterns may differ. Furthermore, it does not focus specifically on geospatial attributes, which would be key in the context of geotourism for the analysis of false information free from errors.

This domain transfer limitation must be addressed in future research. For all these reasons, the detected misinformation does not necessarily cover the full picture of actual tourism misinformation, so predictions may be distorted, and the model may find it relatively difficult to classify social media posts that do not fall within the time horizon or context considered.

It must also be seen that text-based fake news detection is unable to handle visual or geolocation falsification, and does not provide protection against location proofing, which is becoming increasingly significant in the digital environment (Kim et al., 2023), so it is not possible to support prevention in this regard with the current approach.

Author Contributions: Conceptualization, R.N., S.A. and F.E.; methodology, R.N. and F.E.; software, S.A. and V.V.; validation, S.A. and V.V.; formal analysis, R.N., V.V. and F.E.; investigation, R.N. and F.E.; data curation, R.N. and F.E.; writing - original draft preparation, R.N. and F.E.; writing - review and editing, R.N. and V.V.; visualization, S.A. and F.E. supervision, R.N.; project administration, F.E. All authors have read and agreed to the published version of the manuscript.

Funding: Not applicable.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study may be obtained on request from the corresponding author.

Acknowledgements: The research undertaken was made possible by the equal scientific involvement of all the authors concerned.

Conflicts of Interest: The authors declare no conflict of interest.

REFERENCES

Abramovich, F., & Grinshtein, V. (2019). High-dimensional classification by sparse logistic regression. *IEEE Transactions on Information Theory, 65*(5), 3068–3079. https://doi.org/10.1109/TIT.2018.2884963

Akay, B. (2020). Examining the Rural Tourism Experiences of Tourists in Emerging Rural Tourism Destination: Burdur Province, Turkey. *Geojournal of Tourism and Geosites, 29*(2), 534–544. https://doi.org/10.30892/gtg.29212-487

Allcott, H., & Gentzkow, M. (2017). Social Media and Fake News in the 2016 Election. *Journal of Economic Perspectives,* 31(2), 211-236. https://doi.org/10.1257/jep.31.2.211

Barberi, F., Davis, M. S., Isaia, R., Nove, R., & Ricci, T. (2008). Volcanic risk perception in the Vesuvius population. *Journal of Volcanology and Geothermal Research, 172*(3-4), 244–258. https://doi.org/10.1016/j.jvolgeores.2007.12.011

Bastick, Z. (2021). Would you notice if fake news changed your behavior? An experiment on the unconscious effects on disinformation. *Computers in Human Behavior, 116*. https://doi.org/10.1016/j.chb.2020.106633

Biau, G. (2012). Analysis of a Random Forests Model. *Journal of Machine Learning Research, 13,* 1063–1095. https://dl.acm.org/doi/10.5555/2188385.2343682

Bischl, B., Binder, M., Lang, M., Pielok, T., Richter, J., Coors, S., Thomas, J., Ullmann, T., Becker, M., Boulesteix, A. L., Deng, D., & Lindauer, M. (2022). Hyperparameter optimization: Foundations, algorithms, best practices, and open challenges. *WIREs Data Mining and Knowledge Discovery, 13*(2), e1484. https://doi.org/10.1002/widm.1484

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning.* Springer https://link.springer.com/book/9780387310732

Bonadonna, C., Asgary, A., Romerio, F., Zulemyan, T., Frischknecht, C., Cristiani, C., Rosi, M., Gregg, C. E., Biass, S., Pistolesi, M., Menoni, S., & Ricciardi, A. (2022). Assessing the effectiveness and the economic impact of evacuation: the case of Vulcano Island, Italy. *Natural Hazards and Earth System Sciences, 22,* 1083–1108. https://doi.org/10.5194/nhess-22-1083-2022

Brantley, S. R., Kauahikaua, J., Babb, J. L., Orr, T. R., Patrick, M. R., Poland, M. P., Trusdell, F., & Oliveira, D. (2019). Communication strategy of the U.S. Geological Survey Hawaiian Volcano Observatory during the lava-flow crisis of 2014–2015, Kīlauea Volcano, Hawai'i. In: M.P. Poland (ed.), 2019, *Field Volcanology: A Tribute to the Distinguished Career of Don Swanson*. Geological Society of America, USA. https://doi.org/10.1130/2018.2538(16)

Cervantes, J., García-Lamont, F., Rodríguez-Mazahua, L., & Lopez-Chau, A. (2020). A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing, 408*, 189–215. https://doi.org/10.1016/j.neucom.2019.10.118

Cho, K., Van Merrienboer, B., Gulcehre, C., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing,* 1724–1734. https://doi.org/10.3115/v1/d14-1179

Cook, J., Echer, U. K., & Lewandowsky, S. (2015). Misinformation and How to Correct It. In *Emerging Trends in the Social and Behavioral Sciences, 1–17*. Wiley & Sons. https://doi.org/10.1002/9781118900772.etrds0222

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 1*. https://doi.org/10.18653/v1/N19-1423

Dina, N. Z., & Juniarta, N. (2022). Deriving Customers Preferences for Hotels from Unstructured Data. *Geojournal of Tourism and Geosites, 43*(3), 872–877. https://doi.org/10.30892/gtg.43305-899

Drápela, E. (2022). Geoheritage and overtourism: a case study from sandstone rock cities in the Czech Republic. *Geological Society, London, Special Publications, 530*, 257-275. https://doi.org/10.1144/SP530-2022-102

Fedeli, G. (2020). 'Fake news' meets tourism: a proposed research agenda. *Annals of Tourism Research*, *80*, 102684. https://doi.org/10.1016/j.annals.2019.02.002

Gelfert, A. (2018). Fake News: A Definition. *Informal Logic*, *38*(1), 84–117. https://doi.org/10.22329/il.v38i1.5068

Gocoglu, A., Demirel, N., & Bozdogan, H. (2024). A Novel Information Complexity Approach to Score Receiver Operating Characteristic (ROC) Curve Modeling. *Entropy, 26*(11), 988. https://doi.org/10.3390/e26110988

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning.* MIT Press.

Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R., & Schmidhuber, J. (2017). LSTM: A Search Space Odyssey. *IEEE Transactions on Neural Networks and Learning Systems*, *28*(10), 2222–2232. https://doi.org/10.1109/tnnls.2016.2582924

Gwon, N., Jeong, W., Kim, J. H., Oh, K. H., & Jun, J. K. (2024). Effects of Intervention Timing on Health-Related Fake News: Simulation Study. *JMIR Formative Research, 8*, e48284. https://doi.org/10.2196/48284

Hilberts, S., Evers, S. M. A. A., Govers, M., & Petelos, E. (2024). The impact of misinformation on social media in the context of natural disasters. *European Journal of Public Health*, *34*(3), ckae144.245. https://doi.org/10.1093/eurpub/ckae144.245

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation, 9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Hodgkin, A. L., & Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology, 117*(4), 500–544. https://doi.org/10.1113/jphysiol.1952.sp004764

Holtorf, C. (2005). Beyond crusades: How (not) to engage with alternative archaeologies. *World Archaeology,* 37(4), 544–551. https://doi.org/10.1080/00438240500395813

Idiongo, P. (2024). The Impact of Fake News on Public Trust in Traditional Media Outlets. *Journal of Communication*, *5*(3), 45-58. https://doi.org/10.47941/jcomm.1984

Joshi, A., Erukude, S. T., & Shamir, L. (2025). Explainable Identification of Similarities Between Entities for Discovery in Large Text. *Future Internet, 17*(4), 135. https://doi.org/10.3390/fi17040135

Karaş, Z. (2024). Effects of AI-Generated Misinformation and Disinformation on the Economy. *Dücze University Journal of Science & Technology*, *12*(4), 2349–2360. https://doi.org/10.29130/dubited.1537268

Kaszás, F., Supeková, S. C., & Keklak, R. (2025). Fake News in Tourism: A Systematic Literature Review. *Social Sciences*, *14*(8), 454. https://doi.org/10.3390/socsci14080454

Kim, C., Chang, S. Y., Lee, D., Kim, J., Park, K., & Kim, J. (2023). Reliable Detection of Location Spoofing and Variation Attacks. *IEEE Access, 11*, 10813–10825. https://doi.org/10.1109/ACCESS.2023.3241236

Kovari, A. (2024). AI for Decision Support: Balancing Accuracy, Transparency, and Trust Across Sectors. *Information, 15*(11), 725. https://doi.org/10.3390/info15110725

Kravets, P., Pasichnyk, V., & Prodaniuk, M. (2024). Mathematical Model of Logistic Regression for Binary Classification. Part 1. Regression Models of Data Generalization. *Vìsnik Nacìonal'nogo unìversitetu "L'vìvs'ka polìtehnìka". Serìâ Ìnformacìjnì sistemi ta merežì, 15*, 290. https://doi.org/10.23939/sisn2024.15.290

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Communications of the ACM, 60*(6), 84–90. https://doi.org/10.1145/306538

Lazer, D., Baum, M., Benkler, Y., Berinsky, A., Greenhill, K., Menczer, F., Metzger, M., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S., Sunstein, C., Thorson, E., Watts, D., & Zittrain, J. (2018). The Science of Fake News. *Science*, *359*(6380), 1094–1096. https://doi.org/10.1126/science.aao2998

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521, 436–444. https://doi.org/10.1038/nature14539

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE, 86*(11), 2278–2324. https://doi.org/10.1109/5.726791

Mantovani, R. G., Horvath, T., Rossi, A. L. D., Cerri, R., Barbon Jr., S., Vanschoren, J., & de Carvalho, A. (2024). Better Trees: An empirical study on hyperparameter tuning of classification decision trees. *Data Mining and Knowledge Discovery*. *38*(3), 1364–1416. https://doi.org/10.1007/s10618-024-01002-5

Meaney, C., Wang, X., Guan, J., & Stukel, T. A. (2025). Comparison of methods for tuning machine learning model hyper-parameters: with application to predicting high-need high-cost health care users. *BMC Medical Research Methodology, 25*, 134. https://doi.org/10.1186/s12874-025-02561-x

Omar, E. D., Mat, H., Karim, A. Z. A., Sanaudi, R., Ibrahim, F., Omar, A., Ismail, M. Z. H., Jayaraj, V. J., & Goh, B. L. (2024). Comparative Analysis of Logistic Regression, Gradient Boosted Trees, SVM, and Random Forest Algorithms for Prediction of Acute Kidney Injury Requiring Dialysis After Cardiac Surgery. *International Journal of Nephrology and Renovascular Disease, 17*, 197–204. https://doi.org/10.2147/IJNRD.S461028

Pagaldiviti, S. R. (2025). Exploring the Influence of Online Travel Reviews on Tourist Decision making and Destination Choices - A Comprehensive Review. *International Journal for Multidimensional Research Perspectives*, *1*(1), 17–24. https://www.chandigarhphilosophers.com/index.php/ijmrp/article/view/61

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, *65*(6), 386–408. https://doi.org/10.1037/h0042519

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence, 1*(5), 206–215. https://doi.org/10.1038/s42256-019-0048-x

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*, 533–536. https://doi.org/10.1038/323533a0

Schneider, L., Bischl, B., & Feurer, M. (2025). Overtuning in Hyperparameter Optimization. *arXiv:2506.19540*. https://doi.org/10.48550/arXiv.2506.19540

Si, S., Zhang, H., Keerthi, S. S., Mahajan, D., Dhillon, I. S., & Hsich, C. J. (2017). Gradient Boosted Decision Trees for High Dimensional Sparse Output. *Proceedings Machine Learning Research, 70*, 3182–3190.

Štrba, Ľ. (2015). Identification and evaluation of geosites along existing tourist trail as a primary step of geotourism development: case study from the Spiš region (Slovakia). *Geojournal of Tourism and Geosites, 16*(2), 127–141. https://gtg.webhost.uoradea.ro/PDF/GTG-2-2015/1_180_STRBA.pdf

Subhadeep, C. (2022). *Fake and Real News Data* [Data Set]. Kaggle. https://doi.org/10.34740/KAGGLE/DSV/3229106

Tanha, J., Abdi, Y., Samadi, N., Razzaghi, N., & Asadpour, M. (2020). Boosting methods for multi-class imbalanced data classification: an experimental review. *Journal of Big Data, 7*(70). https://doi.org/10.1186/s40537-020-00349-y

Tham, A., & Chen, S. H. (2022). Fake News and Tourism – Whose Responsibility Is It? *Journal of Responsible Tourism Management, 2*(1), 32–36. https://doi.org/10.47263/JRTM.02-01-04

Vasist, P. N., & Krishnan, S. (2023). Disinformation 'gatecrashes' tourism: An empirical study. *Annals of Tourism Research, 101*, 103575. https://doi.org/10.1016/j.annals.2023.103575

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., & Kaiser, L. (2017). Attention is All You Need. *31st Conference on Neural Information Processing Systems, 30*, 6000–6010.

Walters, R. M. (2018). How to Tell a Fake: Fighting Back Against Fake News on the Front Lines of Social Media. *Texas Review of Law & Politics*, *23*(1), 111–179. https://www.proquest.com/docview/2193092672/E88E0E12009B4F59PQ/4

Yan, Z., & Wu, Y. (2018). A Neural N-Gram Network for Text Classification. *Journal of Advanced Computational Intelligence and Intelligent Informatics, 22*(3), 380–386. https://doi.org/10.20965/jaciii.2018.p0380

Yang, S., & Berdine, G. (2017). The receiver operating characteristics (ROC) curve. *The Southwest Respiratory and Critical Care Chronicles*, *5*(19), 34-36. https://doi.org/10.12746/swrccc.v5i19.391

Article history: Received: 24.11.2025 Revised: 16.03.2026 Accepted: 17.06.2026 Available online: 30.06.2026