sustainability

MDPI

Article

Evaluation and Interpretation of Tourist Satisfaction for Local Korean Festivals Using Explainable AI

Hoonseong Oh [1] and Sangmin Lee [2,*]

[1] Korea Culture & Tourism Institute, 154 Geumnanghwaro, Gangseo-gu, Seoul 07511, Korea; hsoh@kcti.re.kr
[2] School of Information Convergence, College of Software and Convergence, Kwangwoon University, Nowon-gu, Seoul 01897, Korea
* Correspondence: smlee5679@gmail.com; Tel.: +82-2-940-5764

Abstract: In this paper, we propose using explainable artificial intelligence (XAI) techniques to predict and interpret the effects of local festival components on tourist satisfaction. We use data-driven analytics, including prediction, interpretation, and utilization phases, to help festivals establish a tourism strategy. Ultimately, this study aims to identify the most significant variables in local tourism strategy and to predict tourist satisfaction. To do so, we conducted an experimental study to compare the prediction accuracy of representative predictive algorithms. We then built a surrogate model based on a game theory-based framework, known as SHapley Additive exPlanations (SHAP), to understand the prediction results and to obtain insight into how tourist satisfaction with local festivals can be improved. Tourist data were collected from local festivals in South Korea over a period of 12 years. We conclude that the proposed predictive and interpretable strategy can identify the strengths and weaknesses of each local festival, allowing festival planners and administrators to enhance their tourist satisfaction rates by addressing the identified weaknesses.

Keywords: local festivals; explainable artificial intelligence; Shapley additive explanations; tourism strategy; tourist satisfaction; light gradient boosting machine learning

check for updates

Citation: Oh, H.; Lee, S. Evaluation and Interpretation of Tourist Satisfaction for Local Korean Festivals Using Explainable AI. *Sustainability* **2021**, *13*, 10901. https://doi.org/10.3390/su131910901

Academic Editors: João Guerreiro and Sandra Maria Correia Loureiro

Received: 12 August 2021
Accepted: 28 September 2021
Published: 30 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The existence of local festivals is closely related to regional suburban development outside of city centers [1]. Festival tourism, which utilizes local tourism resources to attract visitors to local festivals, is an important means of promoting the tourism industry and has become one of the most critical sources of tourism in many countries [2–5]. Local festivals also contribute to population growth, as they promote the establishment or supplementation of local infrastructure [6]. Furthermore, local festivals not only create jobs by revitalizing the local economy but also help regions preserve their history and culture and nurture their local identities [7,8].

Existing tourism research has focused on the planning of local festivals to improve tourist satisfaction, which, it argues, can be done through the provision of cultural experiences, food and drink, and other forms of entertainment [9,10]. A considerable amount of research has determined how towns and cities can attract more tourists through local festivals [11].

Many studies have focused on determining how festival tourist satisfaction levels can be improved. The key factors, which have been suggested by previous studies, are summarized Table 1. Kim et al. [12] identified significant positive and negative factors affecting local food festivals and clarified the relationship between tourist satisfaction and loyalty. Velikova et al. [13] presented case studies of tourist satisfaction by focusing on local wine festivals and found that certain product and service attributes cause greater satisfaction. Some studies have considered tourist travel motivations as the most dominant factor affecting satisfaction [14–17], as travel motivations play a significant role in encouraging travelers to research the offerings of local festivals and plan their travel routes in

advance. However, other factors may also affect the satisfaction of tourists toward local festivals. Albayrak et al. [18] noted that the classic model for satisfaction, which assumes a direct relationship between motivation and overall satisfaction, may differ from reality. Therefore, in their study, they examined the significant effect of the timing of motivation measurements on outcomes and showed that other post-experience factors can also be a determinant of overall satisfaction.

Table 1. A summary of key festival factors described in the literature.

Key Factors	References
Details on festival program and location	[5,11]
Convenience, travel information, employees, program content, souvenirs, and food	[7,8]
Experiences of cultural content, food and drink, entertainment	[9,10]
Products and services	[8,13]
Festival location and transportation, sufficient employees with volunteers, necessary festival information regarding publicity, food, and accommodation facilities	[12,15,16]
Travel motivation, emotional experience, personal benefits, and cultural communication	[14–18]

It is difficult to establish a universal strategy to make local festivals attractive because there are too many decision-making factors [19]. Moreover, it is impossible to consider regional diversity for all suburban regions [20]. Regarding tourist satisfaction, various studies have examined the effect of travel recommendations [21–26] and have provided content or user-based travel recommender systems to match and recommend appropriate routes or travel packages to travelers. These studies stated that travel recommender systems are helpful for eliciting positive customer feedback.

Several existing studies have focused their research into customer satisfaction by concentrating on holidays such as Christmas, Easter, or Ascension days [27,28]. To determine how to increase holiday tourism in specific sightseeing cities, such as Merano or Bethlehem, they conducted a quantitative experiment using an ad hoc survey. However, these studies examine only a few days over the course of one year, which makes it difficult to use their results to establish a macro-level strategy for local festivals more generally.

More recently, various data-driven studies have addressed the limitations of the existing research into the activation of local festivals. In particular, machine learning methods have become an increasingly popular means of obtaining better solutions to the problem of tourist satisfaction because they reflect realistic uncertainties. Such methods have widely been used to identify event, tourist, and market segmentations and have outperformed other methods in their ability to resolve real-world problems [29–31]. Nevertheless, these studies focus only on the use of marketing strategies to persuade specific tourist or customer groups. Therefore, to revitalize local festivals, it is necessary to recognize their strengths and weaknesses to identify the improvements that will promote continuous growth.

As for research gaps, our study differs from previous research on local festivals in the following ways: First, existing studies have not provided sufficient data on the activation of local festivals, nor have they provided accurate predictions for attendee satisfaction. Specifically, no study has used artificial intelligence models to consider the variables that affect the satisfaction of tourists toward local festivals. Therefore, an interpretation model is needed that distinguishes between positive and negative evaluation variables for each local festival, and an accurate prediction model is needed because poor prediction performance leads to incorrect interpretations.

To resolve this, we propose a data-driven approach using artificial intelligence techniques to accurately predict tourist satisfaction and to identify the variables affecting tourist satisfaction for local festivals. For over 12 years and using 20 survey questions, we gathered

data from tourists who have attended local festivals. The tourist satisfaction score was set as the dependent variable, with all other variables set as independent variables.

Figure 1 illustrates the novel framework used in this study. The seven variables presented in the dotted box are considered to be critical components that determine the success of a local festival. Here, we used these independent variables to predict a dependent variable: tourist satisfaction. In particular, the importance of the interpretability of AI has become more apparent in order to increase the accuracy of AI while enabling humans and AI to pursue joint work. As such, these understandable AI techniques play an important role in the analysis of local festivals in our study.

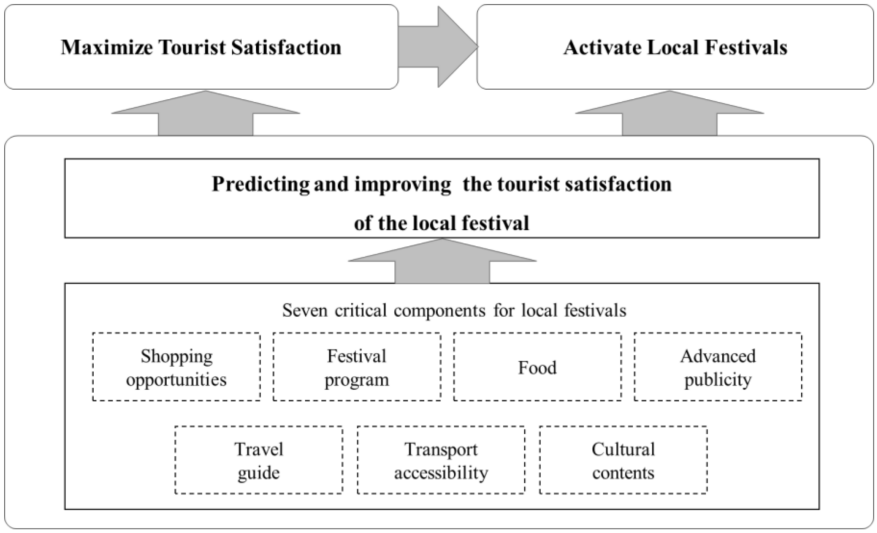


Figure 1. Proposed framework to maximize tourist satisfaction and to activate local festivals.

As individual indicators for tourist satisfaction, the following variables were investigated: shopping opportunities, the festival program, food, advanced publicity, travel guide, transport accessibility, and cultural content. When interpreting the significant effects between the independent and dependent variables, merely identifying the variables with significance for local festivals in general does not help us recognize the shortcomings of individual festivals. To resolve this problem, we proposed the use of Shapley Additive exPlanations (SHAP), a game theory-based framework, to identify the significant variables affecting each local festival sample. By understanding the weaknesses and strengths of an individual local festival, it is possible to identify which of the festival's marketing strategies should be strengthened and which strategies should be modified.

The main contributions of this study are as follows: First, we propose an AI-based tourism evaluation approach for local festivals. The proposed AI-based approach aims to indicate which festivals receive good results in terms of tourist satisfaction. We provide thorough explanations of the overall data-driven procedure, detailing the data preprocessing methods, including the handling of missing values, normalization issues, and the learning framework.

Second, to obtain an accurate prediction model for tourist satisfaction, representative regression models were used to determine the best predictive model for the given data set. A comparative evaluation revealed that the proposed deep learning model outperformed other models in terms of prediction accuracy. To the best of our knowledge, research into tourist satisfaction has yet to use explainable AI (XAI) to interpret results for individual local festivals.

Third, the experimental results were derived from a local festival data set gathered over 12 years. The experiments confirm that the conventional approach, which is operated by human experts, can be improved through the incorporation of AI-based approaches. In terms of the results of some of the interpretations, the experimental results are contrary to the understanding of local festival agencies and administrators.

The remainder of this paper is structured as follows Section 2 shows the overall analysis procedure, including data preprocessing and the results of the exploratory data analysis. Sections 3 and 4 present the theoretical descriptions and experimental results for the model predictions and explanations, respectively. Finally, Section 4 presents the concluding remarks.

2. Methods

2.1. Overview

This section outlines the proposed approach, as shown in Figure 2. First, we collected the data representing tourist satisfaction and specific evaluation metrics for local festivals in South Korea. The data sets were first validated for missing values and distribution shape. In particular, the data set was transformed using quantile transformation, which is a robust preprocessing schema used to reduce the impact of outliers [32]. We then conducted log transformations to achieve better prediction accuracy. Based on the preprocessed data set, the prediction models were built to regress tourist satisfaction. We used 10 more representative machine learning algorithms for a regression task and then adopted an XAI technique, SHAP, to allow us to decompose the prediction results. Based on the decomposed SHAP value for a variable, we then indicated the feature's importance for predictions.

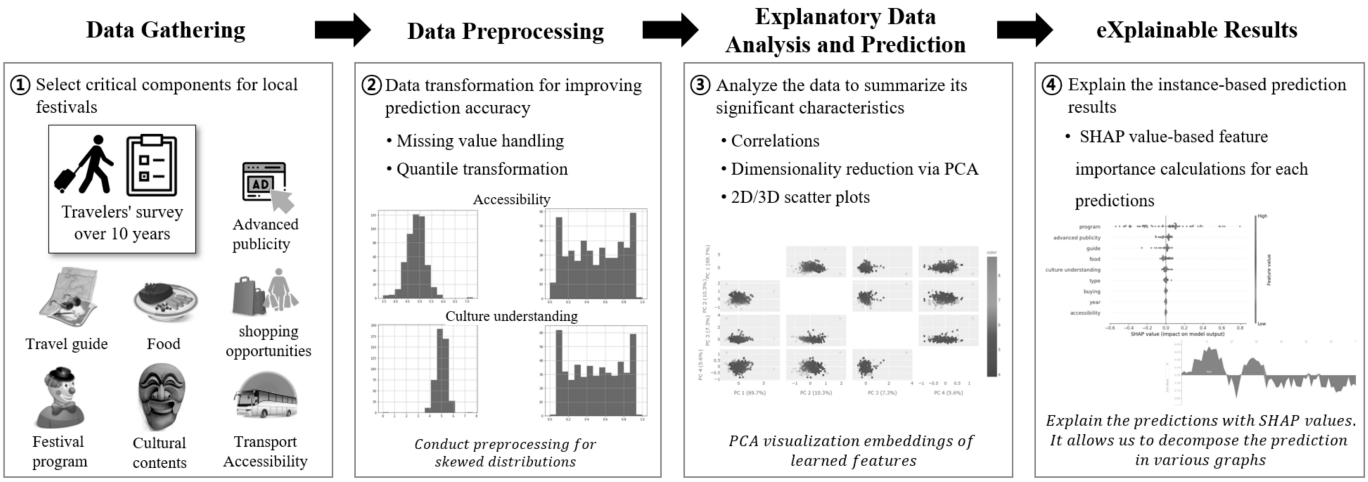


Figure 2. Procedure of the proposed approach with predictive and explainable methods.

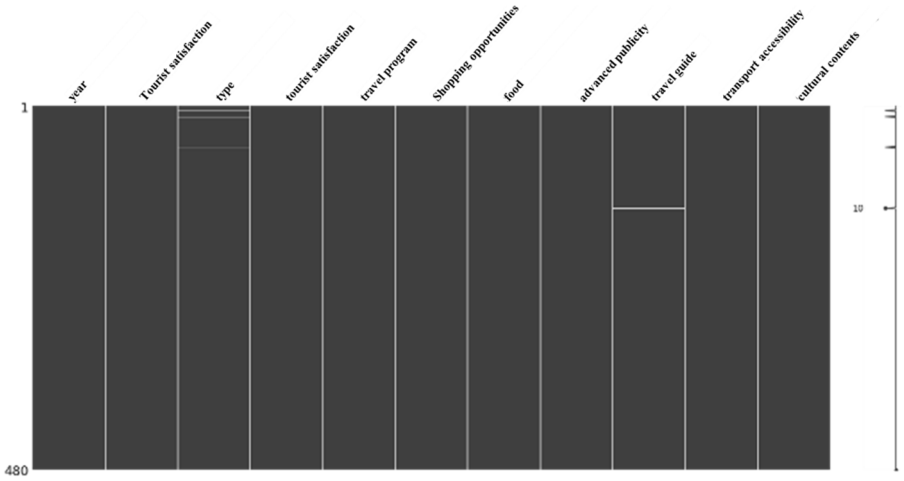
2.2. Data Description and Transformation

We gathered survey data from tourists at local festivals in South Korea over a span of 12 years. The Korean government has conducted surveys on local festivals from 1995 in order to improve their attraction for foreign tourists. The Ministry of Culture and Tourism selected the superior local festivals and provided financial support. http://www.mcst.go.kr/kor/s_notice/press/pressView.jsp?pSeq=17724 (accessed on 11 August 2021).

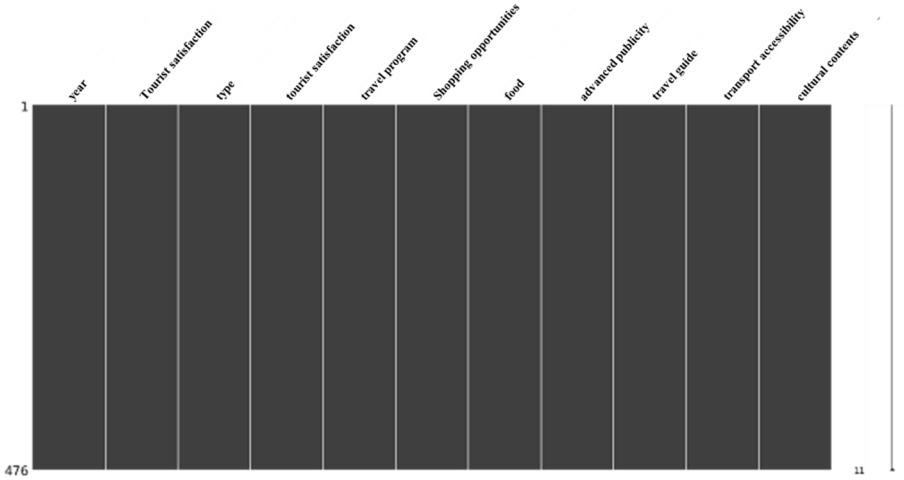
There were 476 total observations. After removing insignificant variables by adopting qualitative and quantitative methods, the final data sets included were the year, festival ID, festival type, the festival program, shopping opportunities, food, advanced publicity, travel guide, and cultural content. We used these data sets to achieve two goals: (1) to predict the tourist satisfaction toward different festivals and (2) to build an explainable model to identify the strengths and weaknesses of each festival. Based on the interpretations of the corresponding tourist satisfaction rate, we indicated the strengths and weaknesses of each festival.

First, we checked the proportions of the missing values to investigate the completeness of the data sets. Figure 3a illustrates the original data, including missing values, by providing missing data visualizations and by utilizing a quick summary of the data completeness. At a glance, we can see that there were missing values in a few observations

related to two variables: type and guide. We have two reasons for simply deleting these observations, as shown in Figure 3b. The proportion of observations with missing values is trivial, and we did not adopt the order information between observations from the data because the order information is insignificant.



(**a**) Original data including missing values



(**b**) Transformed data with missing values removed

Figure 3. Procedure to handle missing values. Because there were few observations with missing values and because autoregressive patterns were not considered, we simply deleted the observations with missing values.

Second, to improve prediction accuracy, we checked the data transformation phase to handle data distribution by considering the quantile and log transformations. After several experiments, we selected the quantile transformation for a normalization procedure. As shown in Figure 4, the less skewed and sparse distributions for each variable result from the quantile transformations. Finally, we used min-max scaling, which realizes equal scaling for independent variables. After conducting transformations for all of the independent variables, we obtained both the accurate performance and fast convergence speed of the prediction models because the transformed data set reduced the sparse area in the data space [33].

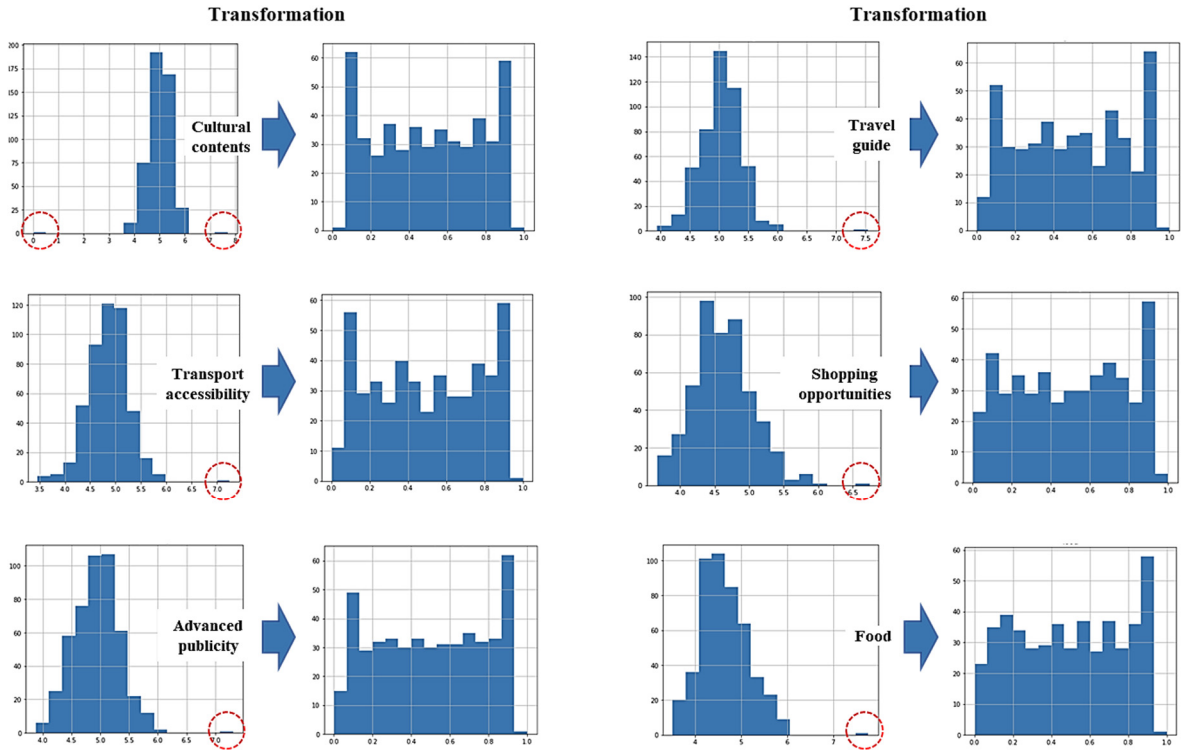


Figure 4. Histograms of three significant variables—cultural content, transport accessibility, and advanced publicity—before and after quantile transformation. It can be seen that the distribution of sparse regions is greatly reduced.

2.3. Exploratory Data Analysis

Here, we present the exploratory data analysis used to understand the data distribution and simple but significant data patterns. Figure 5 presents a scatter matrix showing associations between independent variables. The length of the rows and columns of the matrix represents the number of variables, and each cell plot in the matrix displays the scatter plot of the variables X_i and X_j. Regarding year, it is difficult to indicate significant changes in time for all variables. Regarding festival types, the types of festivals held each year are similar, and no significant difference exists among them. As for the pairwise relations among the seven significant variables, the cell plots of the red dotted rectangle in Figure 5 illustrate that positive linear relations are observed in general.

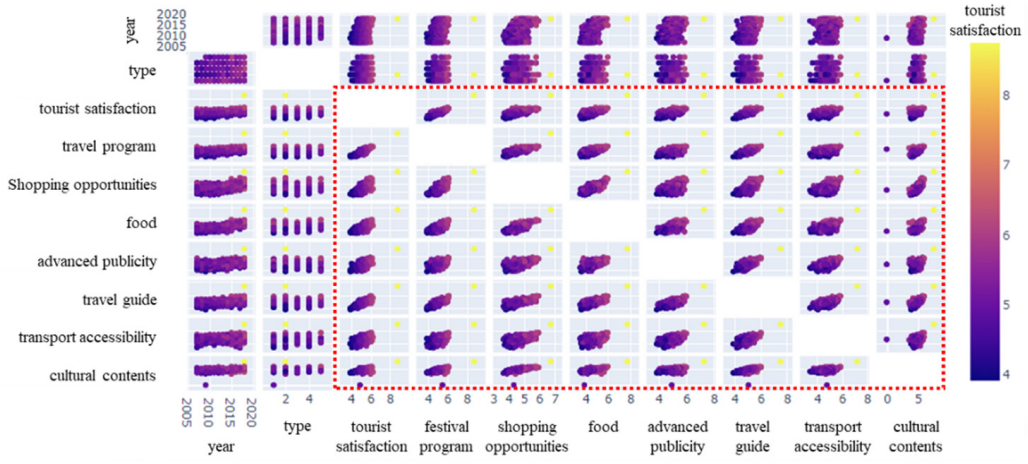


Figure 5. Data preparation and exploratory data analysis.

As shown in Figure 6, we intuitively identified the highly correlated variable pairs, including {food, shopping opportunities}, {travel guide, festival program}, {shopping opportunities, festival program}, {food, festival program}, and {shopping opportunities, cultural content}. The results of the correlation matrix informed us that each variable is closely tied to improvements in tourist satisfaction toward local festivals. We found that the variable combination with the highest linear correlation among three different variables (including dependent variables) was {festival program, food, and tourist satisfaction}. Note that only linear correlations are visualized in this plot. Furthermore, Figure 7a shows a three-dimensional scatter plot with a grid-patterned hyperplane. As described, it shows how much more important the festival program is to leveraging tourist satisfaction than food, but food is also important. However, Figure 7b shows that on the plot with travel guide, food, and tourist satisfaction, it is difficult to infer clear linear correlations between the two independent variables (cultural content and transport accessibility) and the dependent variable (tourist satisfaction).

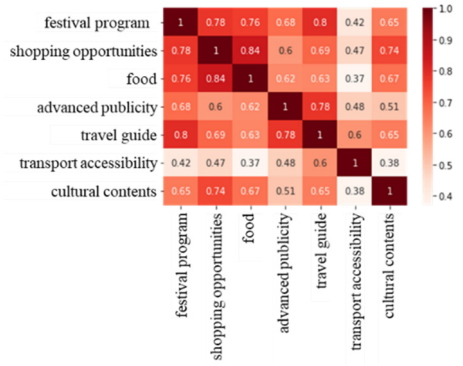
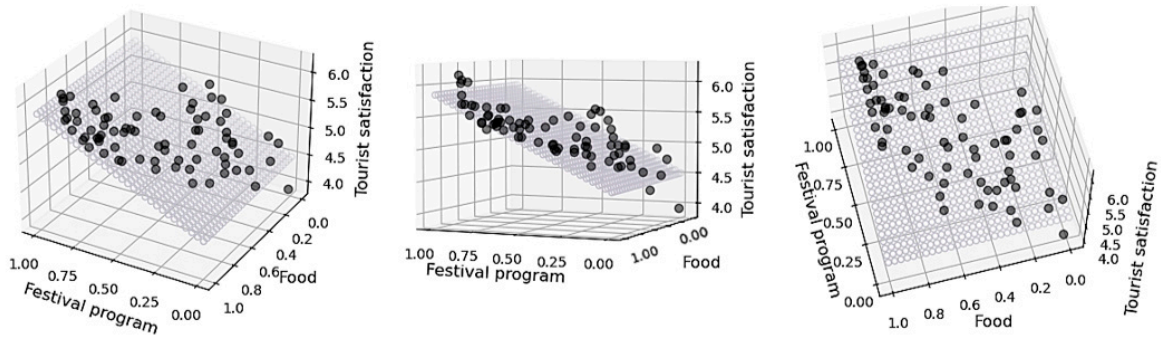
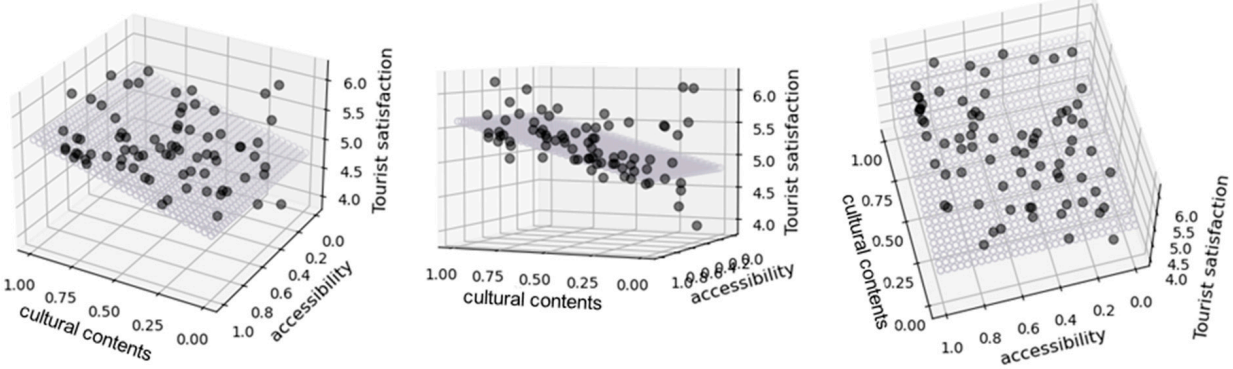


Figure 6. Results of correlation matrix with heatmap plotting.



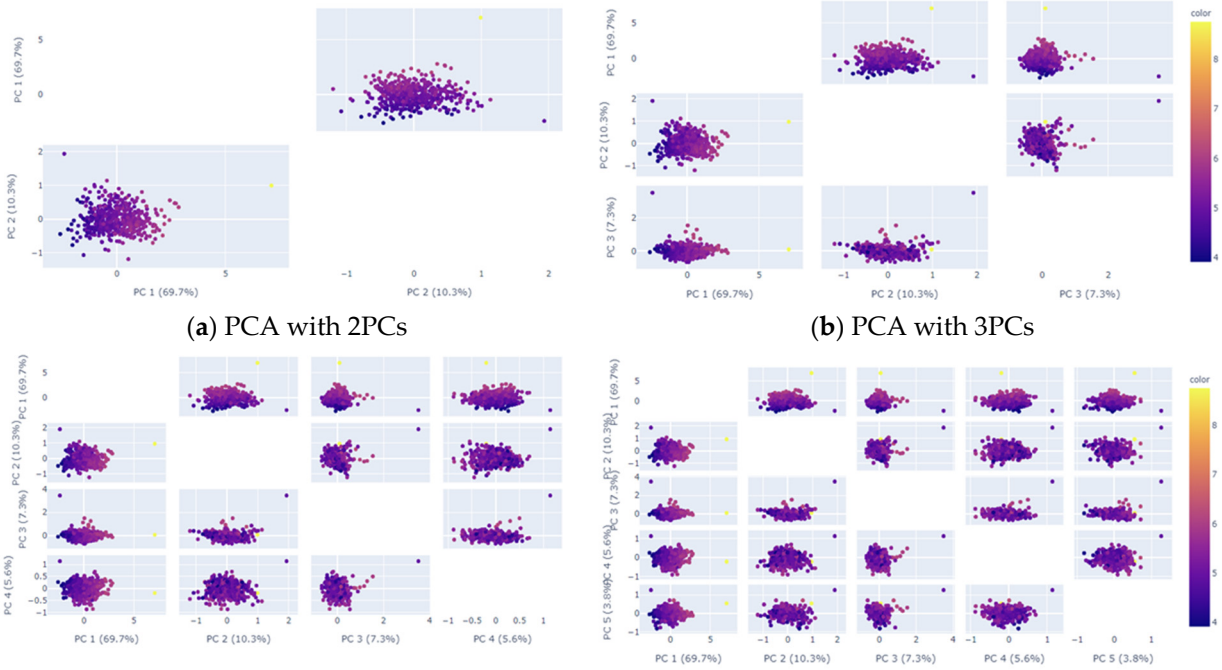
(**a**) Three-dimensional plots for festival program, food, and tourist satisfaction.



(**b**) Three-dimensional plots for cultural content, transport accessibility, and tourist satisfaction.

Figure 7. Three-dimensional plots for two different combinations of independent and dependent variables.

Next, we explored whether the independent variables have sufficient explanatory power in our data sets. Here, we used a dimensionality reduction technique to conduct a multivariate data analysis. Figure 8 shows the results of the dimensionality reduction with a principal component analysis (PCA). Each figure presents a plot that was drawn while changing the number of principal components (PCs). Figure 8a indicates that the first two PCs account for over 80% of the total variance in the original data sets. Figure 8b–d show the results for three PCs, four PCs, and five PCs, respectively. Since the PCA assumes independence between PCs, it is natural that the scatter plot matrix does not appear to be correlated. Note that the purple-colored dots denote a higher tourist satisfaction rate, while the yellow-colored dots denote a lower rate. For each PC, we observed that the PC is linearly correlated with a dependent variable (i.e., tourist satisfaction), indicating that the independent variables can be used to predict the dependent variable in our data sets. Table 2 summarizes basic statistics for each variable.



(**a**) PCA with 2PCs (**b**) PCA with 3PCs

(**c**) PCA with 4PCs (**d**) PCA with 5PCs

Figure 8. Results of the dimensionality reduction using principal component analysis (PCA).

Table 2. Basic statistics for each variable. Values are displayed to two decimal places.

Variable	Count	Mean	Std	Min	25%	50%	75%	Max
Festival program	476	5.09	0.46	3.88	4.76	5.075	5.39	8.75
Shopping opportunities		4.63	0.43	3.67	4.35	4.60	4.90	6.75
Food		4.65	0.53	3.52	4.28	4.59	5.00	7.76
Advanced publicity		4.94	0.40	3.88	4.65	4.97	5.19	7.31
Travel guide		5.04	0.35	3.94	4.84	5.05	5.24	7.55
Transport accessibility		4.86	0.40	3.46	4.60	4.87	5.12	7.25
Cultural content		5.01	0.49	0	4.75	5.03	5.33	7.70

3. Prediction

This section outlines how the regression models were built to compare the performance of the representative machine learning algorithms. Figure 9 shows the data structure used to build the prediction models, which includes seven independent variables (festival program, shopping opportunities, food, advanced publicity, travel guide, transport accessibility, and cultural content) and two additional variables (year and festival type). Among these, the variable "year" was used to verify whether tourist satisfaction differed over time.

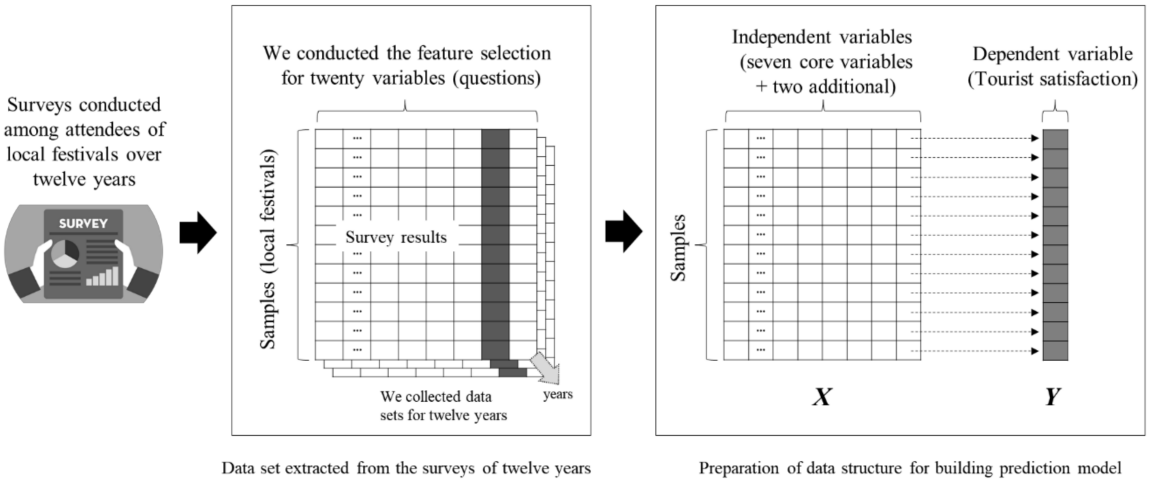


Figure 9. Data structure used to build the prediction models.

To select the best model, we considered 17 representative machine learning algorithms: (1) linear regression models, including lasso [34], ridge [35], elastic net [36], and passive aggressive regressors [37]; (2) *k*-nearest neighbor [38]; (3) decision tree regressor; (4) support vector regressors (SVR), including SVR with linear or polynomial kernels and nu-SVR [39]; (5) bagging methods, such as bagging regressors and random forest regressors [40]; (6) boosting methods, such as an AdaBoost regressor, gradient boosting machines, and an XGB regressor [41]; and (7) deep neural networks [42].

(1) The lasso, ridge, and elastic net linear regressors all identify the fitting function, which minimizes the prediction error with different regularization terms and shrinkage roles of parameter variance. The cost functions for these shrinkage methods are calculated as follows:

$$J(w) \text{ of lasso regressor} = \sum_{i=1}^{n} (y_i - wx_i)^2 + \lambda \sum_{j=1}^{p} |w_j|, \quad (1)$$

$$J(w) \text{ of ridge regressor} = \sum_{i=1}^{n} (y_i - wx_i)^2 + \lambda \sum_{j=1}^{p} w_j^2, \quad (2)$$

$$J(w) \text{ of elastic net regressor} = \sum_{i=1}^{n} (y_i - wx_i)^2 + \lambda \sum_{j=1}^{p} |w_j| + \lambda \sum_{j=1}^{p} w_j^2, \quad (3)$$

where λ denotes a weight hyperparameter, p denotes the number of variables, and w is a parameter to be learned. Although our data set has few independent variables, the experiments were performed using three methods to evaluate the relative performance accuracy. The passive aggressive regressor is a first-order online learning method that updates the weight, w, to optimize the following equation:

$$w_{t+1} : \min_{w} \frac{1}{2} \|w - w_t\|^2, \quad (4)$$

where w_t denotes a parameter to be learned at time t. We aggressively update w_t when the loss is nonzero as $w_{t+1} : w_t - \varrho_t \frac{\partial l_t(w_t)}{\partial w_t}$, where ϱ_t denotes $\frac{l_t(w_t)}{\|x_t\|^2}$, which is the learning rate at time t, and x_t denotes the sample.

(2) The k-nearest neighbor algorithm is a nonparametric method that uses k-nearest training samples in the feature space. The k-nearest neighbor regressor predicts the dependent variable by using the average values of its k-nearest neighbors.

(3) The decision tree algorithm is also a well-known nonparametric method with a tree-like structure. We used classification and regression tree (CART) techniques, which recursively divide data into sets of rectangular regions and model the distribution of the dependent variables in order to make predictions [43].

(4) Support vector machines (SVM) are a supervised learning method that efficiently handle high-dimensional data. SVRs embed the independent variables onto a high-dimensional feature space to build a linear regressor [44]. The objective function of the SVR is defined as follows:

$$\min_{w}(\frac{1}{2}\|w\|^2 + c\sum_{i=1}^{n}|\xi_i|)s.t.|y_i - w_i x_i| \leq \varepsilon + |\xi_i|, \quad (5)$$

where ξ_i denotes the deviation from the support vector margin for the concept of slack variables when ε is the threshold for lower error sensitivity in the training data set.

(5) Bagging methods build an ensemble of multiple classifiers by manipulating the training data with weak learners. Among these, the random forest is a representative ensemble algorithm that constructs multiple decision trees to avoid overfitting. As for regression, the random forest makes predictions by averaging the predicted values of its individual decision trees [40]. The random forest technique is well-known before being robust against noise and overfitting problems.

(6) Boosting methods produce a predictive model by combining weak learners produced in an iterative fashion [41]. Among these, the light gradient boosting machine (LGBM) learning method is best able to show highly accurate performance in various fields. LGBM has two benefits: (1) it achieves higher accuracy than other boosting approaches, such as eXtreme Gradient Boosting or AdaBoost, by enabling more complex leaf-oriented split trees, and (2) it is faster in the training phase and offers high efficiency in terms of gradient descent [45].

(7) Finally, artificial neural networks have recently received increased attention because of the impressive accuracy of their predictions. These networks build a cascade of several layers for linear and nonlinear processing to perform representation learning [42].

To obtain reliable hyperparameter settings, a 10-fold cross validation with grid search was used to minimize the mean squared error (MSE) for each model with the given data set. We measured three accurate performance metrics: R^2, adjusted R^2, and MSE. R^2 is a statistical measure that shows the proportions of the variance of the predicted values for the variance of the actual value in the dependent variable. Adjusted R^2 is a modified measure of R^2 that is created by adjusting for the number of independent variables in the trained model. The adjusted R^2 is used to correct for overestimation. Finally, MSE is calculated as follows:

$$\text{MSE} = \frac{1}{n}\sum_{j=1}^{n}\left|y_j - \hat{y}_j\right|^2. \quad (6)$$

Note that the reason we built so many predictive methods was to ensure that the predictors could determine tourist satisfaction with a high level of accuracy. The accuracy of the predictors must be confirmed because their performance directly connects to the next research step (i.e., XAI) to identify the true strengths and weaknesses of each local festival. The less accurate the predictor performance is, the more likely we are to arrive at incorrect characteristics for the local festivals.

Table 3 presents the adjusted R^2, R^2, and MSE results for the 17 regressors used in this study. Overall, the light gradient boosting regressor outperforms in three metrics. Therefore, based on these experiments, we selected the light gradient boosting regressor as our

predictive model for tourist satisfaction. In the next section, we present the decomposition of the prediction for each festival in a step-by-step manner.

Table 3. Prediction results of 17 predictors in the test phase (for R^2, higher is better; for MSE, lower is better).

Category	Algorithms	*Adj.* R^2	R^2	*MSE*
Boosting	Light Gradient Boosting Regressor	0.85	0.87	0.0169
Boosting	Gradient Boosting Regressor	0.84	0.85	0.0225
Boosting	eXtreme Gradient Boosting Regressor	0.83	0.85	0.0225
Neural net	Artificial Neural Networks	0.82	0.84	0.0241
Bagging	Random Forest Regressor	0.81	0.84	0.0256
Bagging	Bagging Regressor	0.8	0.83	0.0289
Boosting	AdaBoost Regressor	0.77	0.79	0.0324
Support vector machines (SVM)	Nu Support Vector Regressor (SVR)	0.76	0.79	0.0324
SVM	SVR (Polynomial kernel)	0.76	0.79	0.0324
kNN	*k*-Nearest Neighbors Regressor	0.72	0.75	0.04
SVM	SVR (Linear kernel)	0.72	0.75	0.04
Linear	Elastic Net Linear Regression	0.71	0.74	0.04
Linear	Lasso Linear Regression	0.71	0.74	0.04
Linear	Ridge Linear Regression	0.71	0.74	0.04
Boosting	Stochastic Gradient Descent Regressor	0.69	0.73	0.0441
Decision tree	Decision Tree Regressor	0.68	0.72	0.0441
Linear	Passive Aggressive Regressor	0.43	0.49	0.0841

In addition, we presented the critical hyperparameter setting for the algorithms. Regarding bagging and boosting-based predictors, we set the number of the weak classifiers to 260–300, used the bootstrap procedure, and explored the adequate max depth of the tree-based classifiers from 3–20. For the decision tree, we considered the tree height (from 4–20) and the performance metric as GINI impurities. Regarding kNN, we set the number of neighbors to 4 and used Manhattan distance to measure the similarity between the observations, As for linear models such as lasso, ridge, and elastic net, we identified the optimal value of the hyperparameter as $\lambda_0 = 0.13$ (lasso), $\lambda_1 = 0.02$ (ridge), $\lambda_0 = 0.0015$, $\lambda_1 = 0.00001$ (elastic net). In SVM, the value of the hyperparameters were: C = 10, gamma = 0.0001, and kernel = "radial basis function". For neural networks, the model contained one input layer, three hidden layers, and one output layer. Regarding the hidden layers, the dropout (probabilities = 0.1) and rectified linear units were used to prevent overfitting. For the other variables, the default value was used. Overall, the random state of each algorithm was set to 2021 for reproducibility.

4. Explainable Artificial Intelligence for Predictions of Tourist Satisfaction

4.1. Shapley Additive Explanations

Here, we explain the variable significance for each observation regarding our predictions. Before describing how we used XAI in our study, we present the basics of SHAP, which is one of the most popular XAI frameworks. Based on game theory, SHAP describes the performance of a predictive model. To determine a model's explanation capability, SHAP uses an additive feature attribution technique, defining the output model as a linear addition of the contributions of independent variables $X = \{x_1, x_2, \ldots, x_p\}$, where p is the number of variables. Here, we define the predictive model as $f(\cdot)$ and the explainable model as $g(\cdot)$; this can be formulated as follows:

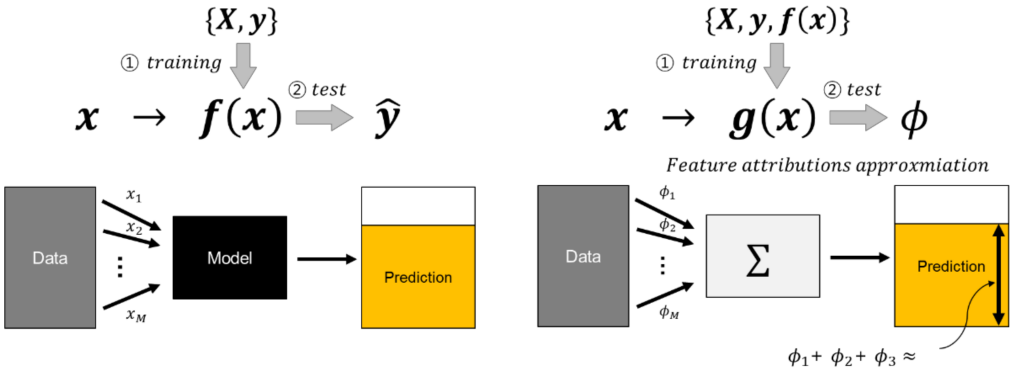
$$f(x) = \hat{y} = g(z') = \phi_0 + \sum_{i=1}^{p} \phi_i z'_i, \tag{7}$$

where z' denotes the transformed independent variables as $z' \in \{0,1\}^p$ for all Shapley values ϕ. As shown, $g(\cdot)$ is a linear function that can be obtained by summing ϕ_i. To confirm absences, ϕ_0 is the constant value when all of the independent variables are missing. To obtain $g(\cdot)$, we used the equation below:

$$\phi_i(f,x) = \sum_{z' \subseteq x'} \frac{|z'|!(p-|z'|-1)!}{p!}[f_x(z') - f_x(z' \backslash i), \tag{8}$$

where $|z'|$ is the number of nonzero variables, and $z' \backslash i$ denotes $z'_i = 0$. Only $g(\cdot)$ can be obtained by this formula.

Using the SHAP framework, we constructed two building steps: building a prediction model and building an explainable model based on the given data set $\{x, y\}$. Figure 10 illustrates the differences between the two steps.



(**a**) Building a prediction model (**b**) Building an explainable model.

Figure 10. Data structure for building prediction models.

First, we built an accurate prediction model $f(x)$ by minimizing the sum of the squared residuals (SSR) function between the observed y and predicted $\hat{y}$. We then built the explainable model by interpreting how $f(x)$ predicts $\hat{y}$. SHAP, a model agnostic method, then allowed us to decompose the prediction results and indicate which variables have a relatively significant impact score, ϕ_i, for the ith variable in the predictions. SHAP uses the Shapley value ϕ, which denotes the mean of the marginal contributions across all permutations of the variables in the predictions.

4.2. Results

Figure 11 provides a summary plot that presents the explainability of the overall SHAP values. The independent variables are ordered according to their predictive contributions, and the colors in the plot illustrate the Shapley value of each independent variable for each observation. Figure 11 shows that the festival program is the most important indicator for the success of a local festival. Put another way, tourist satisfaction with the festival program has a considerable influence on their overall satisfaction with the festival. Lower program scores lead to lower tourist satisfaction, and higher program scores lead to higher tourist satisfaction. Advanced publicity and travel guide are the next most important variables for tourist satisfaction. In particular, we found that tourists underestimate a festival if the travel guide scores low. However, the influences of advanced publicity and travel guide are quite limited, as shown by their respective SHAP values. In contrast, a positive evaluation of the food offered at a festival has a positive influence on tourist satisfaction. In terms of cultural understanding (cultural content), the variable's effect on overall satisfaction is not simply linear, as the figure also presents an inverse relation. The results of the remaining

independent variables are insignificant, and the results of the year variable indicate that no significant change in satisfaction occurred over the 12-year study period.

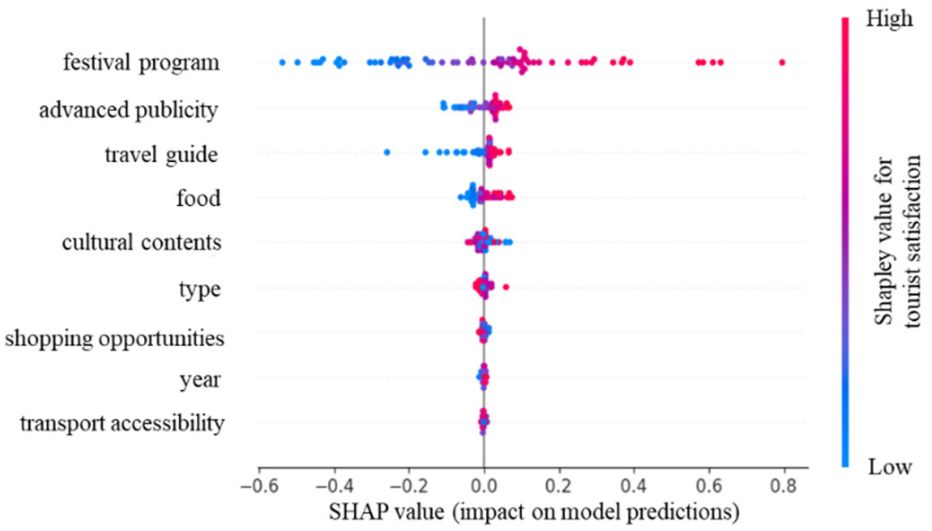


Figure 11. SHAP summary plot for predictors of tourist satisfaction.

We next investigated which variables had a major influence on the cases where the tourist satisfaction score is high and low. To do so, we first divided the observations into two groups based on tourist satisfaction: overestimated and underestimated. Tourist satisfaction was then predicted using the predictive model $f(\cdot)$, and we measured the SHAP values for each variable using the explainable model $g(\cdot)$. Figures 12–15 present the festival cases in which the predictions of tourist satisfaction were underestimated. As seen here, tourists largely identified the festival's program and advanced publicity as the main reasons for a festival's low score. As such, these charts illustrate the weaknesses of these festivals, providing festival planners with insight into which areas should be improved.



Figure 12. SHAP force plot for an underestimated case predicted at a value of 4.69 with identified weaknesses.

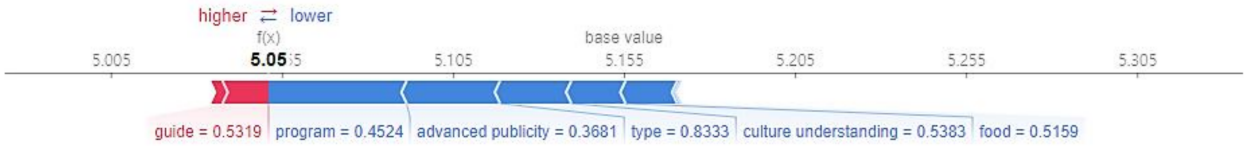


Figure 13. SHAP force plot for an underestimated case predicted at a value of 5.05 with identified weaknesses. (Cultural understanding = cultural content).

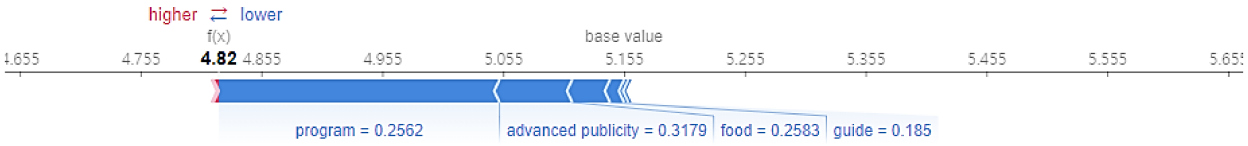


Figure 14. SHAP force plot for an underestimated case predicted at a value of 4.82 with identified weaknesses.

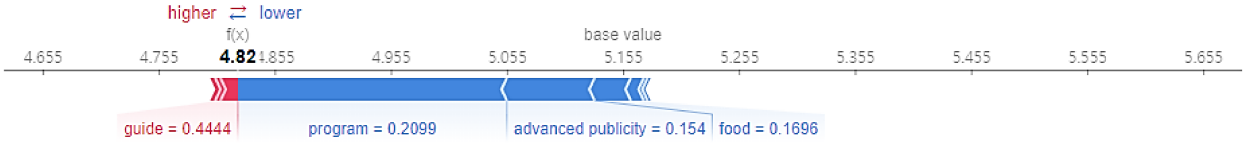


Figure 15. SHAP force plot for an underestimated case predicted at a value of 4.82 with identified weaknesses.

In contrast, Figures 16–19 present the festival cases in which the predictions of tourist satisfaction were overestimated, thereby providing insight into the strengths of each festival. Finally, Figure 20 shows the interaction plots for the single effect of the shopping opportunities and food variables, which are positively correlated to tourist satisfaction; however, the festival program continues to have a large influence on tourist satisfaction.

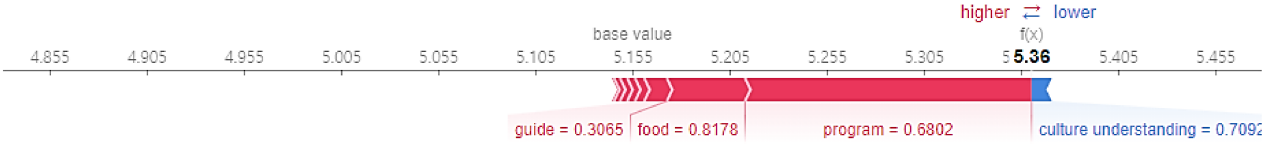


Figure 16. SHAP force plot for an overestimated case predicted at a value of 5.36 with identified strengths.

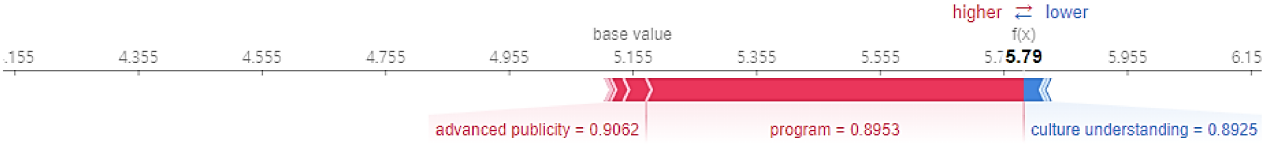


Figure 17. SHAP force plot for an overestimated case predicted at a value of 5.79 with identified strengths.

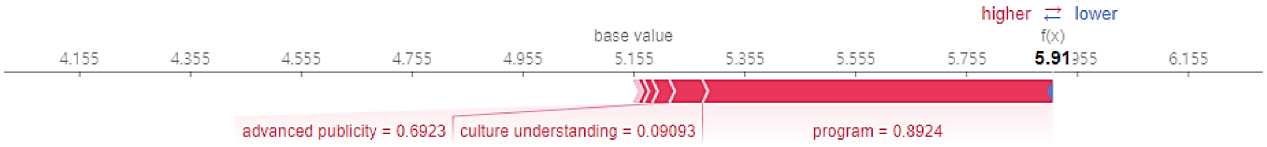
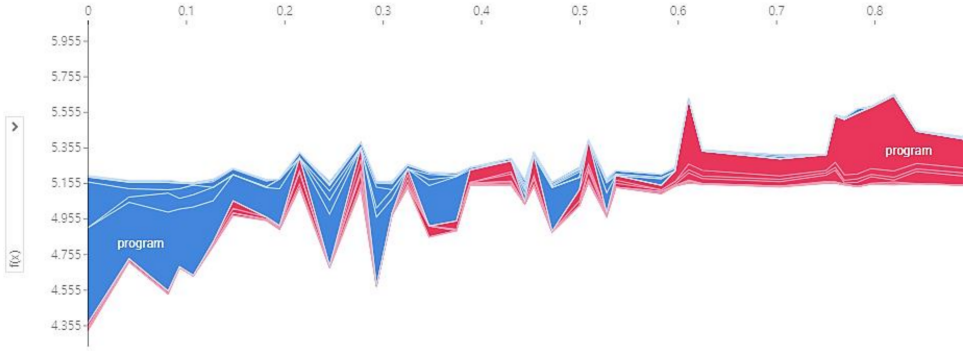


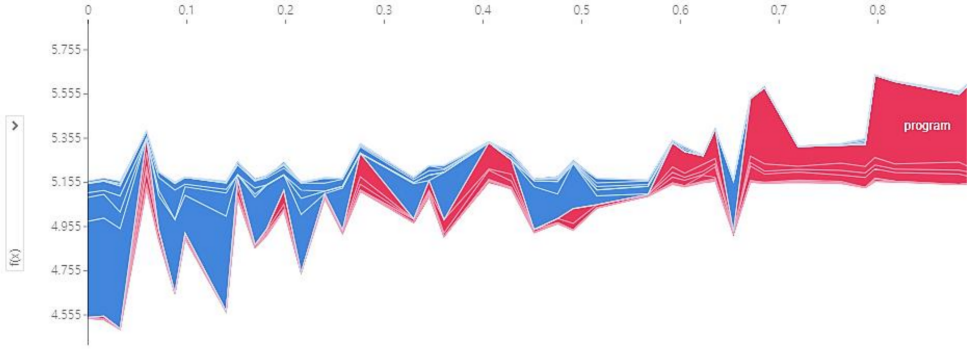
Figure 18. SHAP force plot for an overestimated case predicted at a value of 5.91 with identified strengths.



Figure 19. SHAP force plot for an overestimated case predicted at a value of 5.35 with identified strengths.



(**a**) Prediction changes according to the effect of shopping opportunities. (Program = festival program).



(**b**) Prediction changes according to the effect of food.

Figure 20. Visualizing the feature responsibility for a change in the prediction results.

5. Conclusions

This study aims to predict and explain tourist satisfaction for local festivals by identifying the significant variables to enable festivals to establish an adequate tourism strategy. We built various machine learning models and compared their predictive performance to obtain both the performance and explanation accuracy of predictive models. Subsequently, we reviewed the explanations of predictive results and presented the strength and weakness characteristics of each local festival. The proposed approach is a practical solution to minimize the uncertainty of revitalizing tourism at local festivals by identifying important variables of local festivals and by drawing a deeper understanding of their success points. The experimental results of the XAI demonstrate that the prediction and explanation results offer valuable insights for identifying the problems of local festivals and their potential solutions.

The main contributions of our study are twofold. First, we proposed machine learning-based festival estimation models including both predictive and explainable models. The proposed methods are not only helpful to identify the key success factors for each local Korean festival but also explain what factors should be improved to capture the attention of more tourists. Experimental results based on real data collected over 12 years demonstrated the applicability and effectiveness of our approach. Therefore, the proposed approach could be useful to promote the attraction strategy for tourists, resulting in leveraging the success of local festivals.

There are some limitations to our study. First, we did not consider time series patterns such as increases of preferences for specific types of local festivals and the increase or decrease of the number of foreign tourists. The stationary conditions for all of the variables were assumed to build uncomplex models for relationships between the significant variables for local festivals and tourist satisfaction. Second, our approach is only designed to estimate and explain tourist satisfaction and does not consider exogenous variables such as complex economic effects within the local cities, the effect of an increase or decrease

of the number of foreign and native festival visitors, and the support from local the local administration. Finally, recent local festivals have been greatly affected by the impact of COVID-19, experiencing problems such as (1) fewer tourists, (2) program restrictions, and (3) budget cuts [46]. We should identify the critical effects of a serious social pandemic such as COVID-19 on local festivals.

Regarding future research, we have two plans: (1) a methodological approach and (2) a sustainability approach. First, we plan to discover the deep causal relationships between the significant variables of local festivals over time. We extend the predictive approach to addressing the complex relations between many other variables to improve the applicability and the prediction robustness of the proposed method. Further, future studies may extend this XAI-based approach to model other prediction-based tourism research. Second, our study can be extended to resolving potential problems in the area od sustainability. For overtourism especially, the proposed method can be adopted to diagnose the causes of overtourism. We can establish a strategy to disperse tourists from overcrowded areas by identifying the critical causes for the adjacent less-popular areas. Moreover, we can use the proposed approach to suggest the optimal use of the budget to revitalize tourism. Finally, we can establish an integrated strategy for dispersing tourists by combining the survey results of adjacent cities, especially in cases where a tourism imbalance problem exists. Using the multi-task method, our approach can be helpful to combine multiple survey results and to represent global optimal solutions for tourism strategies.

Author Contributions: Conceptualization, H.O.; methodology, S.L. and H.O.; software, S.L. and H.O.; data curation, H.O.; writing, H.O., S.L. Both authors have read and agreed to the published version of the manuscript.

Funding: This research was conducted through a research grant from Kwangwoon University in 2021. This research was supported by the MIST (Ministry of Science and ICT), under the National Program for Excellence in SW (2017-0-00096), supervised by the IITP (Institute for Information & communications Technology Planning & Evaluation.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Not applicable.

Acknowledgments: This research was conducted through a research grant from Kwangwoon University in 2021. This research was supported by the MIST (Ministry of Science and ICT), under the National Program for Excellence in SW (2017-0-00096), supervised by the IITP (Institute for Information & communications Technology Planning & Evaluation.

Conflicts of Interest: All authors declare no conflict of interest in the present study.

References

1. Stankova, M.; Vassenska, I. Raising Cultural Awareness of Local Traditions Through Festival Tourism. *Tour. Manag. Stud.* **2015**, *11*, 120–127.
2. Liu, C.R.; Lin, W.R.; Wang, Y.C.; Chen, S.P. Sustainability Indicators for Festival Tourism: A Multi-Stakeholder Perspective. *J. Qual. Assur. Hosp. Tour.* **2019**, *20*, 296–316. [CrossRef]
3. Shim, C.; Santos, C.A. Tourism, Place and Placelessness in the Phenomenological Experience of Shopping Malls in Seoul. *Tour. Manag.* **2014**, *45*, 106–114. [CrossRef]
4. Wickens, E. The Sacred and the Profane. *Ann. Tour. Res.* **2002**, *29*, 834–851. [CrossRef]
5. O'Sullivan, D.; Jackson, M.J. Festival Tourism: A Contributor to Sustainable Local Economic Development? *J. Sustain. Tour.* **2002**, *10*, 325–342. [CrossRef]
6. Getz, D.; Page, S.J. Progress and Prospects for Event Tourism Research. *Tour. Manag.* **2016**, *52*, 593–631. [CrossRef]
7. Lee, J.S.; Lee, C.K.; Choi, Y. Examining the Role of Emotional and Functional Values in Festival Evaluation. *J. Travel Res.* **2011**, *50*, 685–696. [CrossRef]
8. Long, P.T.; Perdue, R.R. The Economic Impact of Rural Festivals and Special Events: Assessing the Spatial Distribution of Expenditures. *J. Travel Res.* **1990**, *28*, 10–14. [CrossRef]

9. Rodríguez-Campo, L.; Braña-Rey, F.; Alén-González, E.; Antonio Fraiz-Brea, J. The Liminality in Popular Festivals: Identity, Belonging and Hedonism as Values of Tourist Satisfaction. *Tour. Geogr.* **2020**, *22*, 229–249. [CrossRef]
10. Báez-Montenegro, A.; Devesa-Fernández, M. Motivation, Satisfaction and Loyalty in the Case of a Film Festival: Differences Between Local and Non-Local Participants. *J. Cult. Econ.* **2017**, *41*, 173–195. [CrossRef]
11. Van Aalst, I.; Van Melik, R. City festivals and urban development: Does place matter? *Europ. Urban Reg. Studies* **2012**, *19*, 195–206. [CrossRef]
12. Kim, Y.G.; Suh, B.W.; Eves, A. The Relationships Between Food-Related Personality Traits, Satisfaction, and Loyalty Among Visitors Attending Food Events and Festivals. *Int. J. Hosp. Manag.* **2010**, *29*, 216–226. [CrossRef]
13. Velikova, N.; Slevitch, L.; Mathe-Soulek, K. Application of Kano Model to Identification of Wine Festival Satisfaction Drivers. *Int. J. Contemp. Hosp. Manag.* **2017**, *29*, 2708–2726. [CrossRef]
14. Sohn, H.K.; Lee, T.J.; Yoon, Y.S. Relationship Between Perceived Risk, Evaluation, Satisfaction, and Behavioral Intention: A Case of Local-Festival Visitors. *J. Travel Tour. Mark.* **2016**, *33*, 28–45. [CrossRef]
15. Yoo, I.Y.; Lee, T.J.; Lee, C.K. Effect of Health and Wellness Values on Festival Visit Motivation. *Asia Pac. J. Tour. Res.* **2015**, *20*, 152–170. [CrossRef]
16. Yoon, Y.S.; Lee, J.S.; Lee, C.K. Measuring Festival Quality and Value Affecting Visitors' Satisfaction and Loyalty Using a Structural Approach. *Int. J. Hosp. Manag.* **2010**, *29*, 335–342. [CrossRef]
17. Uysal, M.; Gahan, L.; Martin, B. An Examination of Event Motivations: A Case Study. *Festiv. Manag. Event Tour.* **1993**, *1*, 5–10.
18. Albayrak, T.; Caber, M. Examining the Relationship Between Tourist Motivation and Satisfaction by Two Competing Methods. *Tour. Manag.* **2018**, *69*, 201–213. [CrossRef]
19. Vu, H.M.; Ngo, V.M. Strategy Development from Triangulated Viewpoints for a Fast Growing Destination Toward Sustainable Tourism Development—A Case of Phu Quoc Islands in Vietnam. *JoTS* **2019**, *10*, 117–140. [CrossRef]
20. Truong, T.L.H.; Lenglet, F.; Mothe, C. Destination Distinctiveness: Concept, Measurement, and Impact on Tourist Satisfaction. *J. Destin. Mark. Manag.* **2018**, *8*, 214–231. [CrossRef]
21. Ravi, L.; Vairavasundaram, S.; Palani, S.; Devarajan, M. Location-Based Personalized Recommender System in the Internet of Cultural Things. *J. Intell. Fuzzy Syst.* **2019**, *36*, 4141–4152. [CrossRef]
22. Lim, K.H.; Chan, J.; Leckie, C.; Karunasekera, S. Personalized Trip Recommendation for Tourists Based on User Interests, Points of Interest Visit Durations and Visit Recency. *Knowl. Inf. Syst.* **2018**, *54*, 375–406. [CrossRef]
23. Leal, F.; Malheiro, B.; Burguillo, J.C. Context-Aware Tourism Technologies. *Knowl. Eng. Rev.* **2018**, *33*, e13. [CrossRef]
24. Cenamor, I.; de la Rosa, T.; Núñez, S.; Borrajo, D. Planning for Tourism Routes Using Social Networks. *Expert Syst. Appl.* **2017**, *69*, 1–9. [CrossRef]
25. Sylejmani, K.; Dorn, J.; Musliu, N. Planning the Trip Itinerary for Tourist Groups. *Inf. Technol. Tour.* **2017**, *17*, 275–314. [CrossRef]
26. Tarus, J.K.; Niu, Z.; Yousif, A. A Hybrid Knowledge-Based Recommender System for E-Learning Based on Ontology and Sequential Pattern Mining. *Future Gener. Comput. Syst.* **2017**, *72*, 37–48. [CrossRef]
27. Brida, J.G.; Disegna, M.; Scuderi, R. Segmenting Visitors of Cultural Events: The Case of Christmas Market. *Expert Syst. Appl.* **2014**, *41*, 4542–4553. [CrossRef]
28. Brida, J.G.; Disegna, M.; Osti, L. Segmenting Visitors of Cultural Events by Motivation: A Sequential Non-Linear Clustering Analysis of Italian Christmas Market Visitors. *Expert Syst. Appl.* **2012**, *39*, 11349–11356. [CrossRef]
29. Tkaczynski, A.; Rundle-Thiele, S.R. Event Segmentation: A Review and Research Agenda. *Tour. Manag.* **2011**, *32*, 426–434. [CrossRef]
30. Kruger, M.; Saayman, M.; Ellis, S. Segmentation by Genres: The Case of the Aardklop National Arts Festival. *Int. J. Tour. Res.* **2011**, *13*, 511–526. [CrossRef]
31. Tuma, M.N.; Decker, R.; Scholz, S.W. A Survey of the Challenges and Pifalls of Cluster Analysis Application in Market Segmentation. *Int. J. Mark. Res.* **2011**, *53*, 391–414. [CrossRef]
32. Quantile Transformer, Transform Features from Dataset in Scikit-Learn. Available online: https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.QuantileTransformer.html (accessed on 29 July 2021).
33. Lee, S.; Lim, D.E.; Kang, Y.; Kim, H.J. Clustered Multi-Task Sequence-to-Sequence Learning for Autonomous Vehicle Repositioning. *IEEE Access* **2021**, *9*, 14504–14515.
34. Zou, H. The Adaptive Lasso and Its Oracle Properties. *J. Am. Stat. Assoc.* **2006**, *101*, 1418–1429. [CrossRef]
35. Hoerl, A.E.; Kennard, R.W. Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics* **1970**, *12*, 55–67. [CrossRef]
36. Zou, H.; Hastie, T. Regularization and Variable Selection via the Elastic Net. *J. R. Stat. Soc. B* **2005**, *67*, 301–320. [CrossRef]
37. Lu, J.; Zhao, P.; Hoi, S.C. Online sparse passive aggressive learning with kernels. In Proceedings of the 2016 SIAM International Conference on Data Mining, Miami, FL, USA, 5–7 May 2016; pp. 675–683.
38. Peterson, L.E. K-nearest neighbor. *Scholarpedia* **2009**, *4*, 1883. [CrossRef]
39. Basak, J. A Least Square Kernel Machine with Box Constraints. In *JPRR 19th International Conference on Pattern Recognition, Tampa, FL, USA, 8–11 December 2008*; IEEE: New York, NY, USA, 2008; Volume 5, pp. 38–51.
40. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
41. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, 1189–1232.

42. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef] [PubMed]
43. Breiman, L. Bagging predictors. *Mach. Learn.* **1996**, *24*, 123–140. [CrossRef]
44. Awad, M.; Khanna, R. Support vector regression. In *Efficient Learning Machines*; Apress: Berkeley, CA, USA, 2015; pp. 67–80.
45. Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T.Y. Lightgbm: A Highly Efficient Gradient Boosting Decision Tree. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 3146–3154.
46. Roman, M.; Niedziółka, A.; Krasnodębski, A. respondents' involvement in tourist activities at the Time of the COVID-19 Pandemic. *Sustainability* **2020**, *12*, 9610. [CrossRef]