



Pantano, E., Priporas, C-V., Stylos, N., & Dennis, C. (2019). Facilitating tourists' decision making through open data analyses: A novel recommender system. *Tourism Management Perspectives*, 31, 323-331. <https://doi.org/10.1016/j.tmp.2019.06.003>

Peer reviewed version

License (if available):  
CC BY-NC-ND

Link to published version (if available):  
[10.1016/j.tmp.2019.06.003](https://doi.org/10.1016/j.tmp.2019.06.003)

[Link to publication record in Explore Bristol Research](#)  
PDF-document

This is the accepted author manuscript (AAM). The final published version (version of record) is available online via Elsevier at <https://doi.org/10.1016/j.tmp.2019.06.003> . Please refer to any applicable terms of use of the publisher.

## University of Bristol - Explore Bristol Research

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:  
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

## **Facilitating tourists' decision making through open data analyses: A novel recommender system**

### **Abstract**

A number of studies have recently been published reporting researchers' efforts to create new, more efficient recommender systems to support tourists' decision making. This current research operationalizes a recommender system by filtering user-generated data that is abundantly available online, based on individuals' evaluation criteria, to produce a dataset for analysis. Drawing upon an array of predictive models, this research proposes a new recommender system able to facilitate the tourist decision making process through successful managing of open data. It further presents a rating estimation method using ratings that pertain to online users-specified criteria (profile). The model is able to predict consumers' ratings of a certain product with high reliability starting from open data on their profiles.

**Keywords:** Recommender system; tourist decision making process; consumer information processing; classifier systems; open data analysis

## **1. Introduction**

The latest advances in information technology are changing the practice of marketing (Quinn et al., 2016). Specifically, continuous advances in digital technology, via evolutionary hardware capabilities (e.g. powerful lightweight personal computers, broadband networks), sophisticated software applications (e.g. social network platforms) and the rapid diffusion of smart technologies in the vast range of global population, have dramatically affected consumers' behaviours (Maity & Dass, 2014; Pantano & Priporas, 2016).

Many researchers have acknowledged consumers' tendencies to search for innovative experiences and technologies (Arts, Frambach, & Bijmolt, 2011; Lowe & Alpert, 2015; Pantano & Viassone, 2014). As a consequence, innovation and technology adoption theories have been exploited in marketing research to predict consumers' usage of particular technology (Al-Qeisi, Dennis, Alamanos, & Jayawardhena, 2014; Davis, 1989; Dennis, Jayawardhena, Merrilees, & Wright, 2009; Ng, 2016; Venkatesh, Thong, & Xu, 2012), experience (Blazquez, 2014; Dennis, Brakus, Gupta, & Alamanos, 2014; Verhoef et al., 2009), and purchase behaviour (Alnawas & Aburub, 2016; Pantano & Priporas, 2016). In particular, this technology may influence consumers' behaviours by providing digital tools for searching, comparing and buying products (Kang, Mun, & Johnson, 2015). Hence, an insightful and interactive digital space created by consumers (user-generated content) is freely provided, offering consumers recommendations and prompting them to purchase particular products, services, and brands.

This novel technological enrichment provides innovative pervasive spaces for supporting consumers' access to information and dissemination of consumers' knowledge (including experience, opinions, reviews, etc.) via social networking services (SNSs) (Balaji, Khong, & Chong, 2016). Therefore, internet channels are shifting towards online networks characterized by users' interactive knowledge generation and sharing. These digital spaces offer substantial open data that can be really beneficial to the public and consumers, in

particular. However, natural limitations in the human brain's ability to process large volumes of information packages, make the selection stage in the consumer proposition acquisition process a very challenging task that can lead to information overload, non-optimized resource management (e.g. wasted time, money and/or effort), misguided decision making processes, physical exhaustion and distress or to a combination of these repercussions (Colace et al., 2015; Willemsen & Johnson, 2011). To address these issues in the tourism industry, recommender systems analyse, delineate and mirror tourists' characteristics to create meaningful associations of human profiles and respective needs by leveraging the power of open databases (Oliveira et al., 2017; Pantano et al. 2017). In fact, it has been reported that conduct of open data analyses improve matching of available product propositions to tourists' needs (Okazaki et al., 2015; Zhang et al., 2016). This approach that can potentially generate useful knowledge for both consumers' and companies' decision-making in general, could also lead to happy tourists and improved business performance (Liu & Shih, 2005; Nguyen & Cao, 2015).

The aim of this study is to explore the effectiveness of a novel recommender system algorithm for successfully predicting tourists' preferred choices in their decision making. To this end, this research is grounded on information gain theory data mining (Liao et al., 2012; Sudheep et al., 2011) and proposes [an advanced](#) recommender system based on a rating estimation method through ratings pertaining to users-specified criteria (profile).

Theoretically, the paper investigates the growing phenomenon of technology as facilitator of consumer information processing and decision making, by offering an application in the tourism field, i.e. tourism destinations. From a practical viewpoint, it demonstrates how marketing professionals can exploit open data analysis to influence tourists' decision-making processes. [Thus, the study makes also a methodological contribution which may be of interest to the researchers involved in marketing and management decision-making \(quantitative](#)

analysis) and to tourism scholars who are interested in enhancing their arsenal of computational tools while researching tourists' behaviour.

The remainder of the paper is organised in three main sections. First, it focuses on the consumer information and decision-making process, followed by a review of actual recommender systems employed for supporting tourists. This part also introduces predictive models, also known as Support Vector Machines (SVM) (Suykens, & Vandewalle, 1999) as employed in open data analysis and novel recommender systems development. Second, a case study example of the tourism sector is provided to illustrate the usefulness of the proposed recommender system. Finally, implications are discussed to illustrate the significance of the findings for academics and practitioners, respectively.

## **2. Theoretical background**

### *2.1. Consumer information processing and decision making*

Consumer generated content (CGC) and social networks are the building blocks of social commerce (Amblee & Bui, 2011). During the last few years, social networks have been dramatically changing marketing. On one hand social networks have become one of the leading information sources for consumers (Davis & Khazanchi, 2008), and on the other, they are fast-integrating e-commerce services to sell products directly to internet users (Senecal & Nantel, 2004). In this context, electronic word-of mouth (e-WOM) facilitates consumers' desires to share opinions and experiences in order to help others in their purchases, as well as for personal prestige (Park & Kim, 2008). E-WOM makes the online exchange of opinions an easy and cost-effective process. It overcomes the obstacles encountered in traditional word-of-mouth through scalability, speed of diffusion, multi-way asynchronous information exchange, and measurability of the format and quantity of information released (Cheung & Thadani, 2012; Park & Kim, 2008).

Information theory, which is mainly based on mathematics, statistics and information engineering provides the theoretical underpinnings of classificatory systems and machine learning (MacKay, 2003). Furthermore, machine learning supports decision analysis and predictive modelling through decision trees and data mining, with main aim to better reach and represent choices and decisions (Witten et al., 2016).

Abundant information available on the Internet increases consumers' awareness of product specifications and availability of alternatives yet amplifies the difficulty of making a choice out of a proposition set (Bonhard & Sasse, 2006). Thus, consumers try to find ways to filter internet resources to decrease information overload and at the same time make appropriate choices (Chen, 2008; Li & Du, 2011). This process is a challenging one though, because internet users exchange opinions with others that may not belong in their personal social network of contacts, thus raising concerns about strangers' credibility (Cheung & Thadani, 2012). Hence, source credibility reflects the expertise and trustworthiness of the message source and the extent to which this source could be perceived as believable and competent (Chu & Kamal, 2008; Davis & Khazanchi, 2008). Recommendations coming from trusted sources and friends in particular, greatly contribute to loosening social e-shopping-related constraints (Harris & Dennis, 2011). Moreover, it is suggested that the number of customer reviews of a product plays a key role in motivating consumers to engage with e-shopping via retail portals and social media.

Human cognitive capacity is finite, and individuals can process only part of the plethora of information transmitted by media and other sources. Consumers therefore can retain and respond to only a subset of the total data received, according to individual criteria (Johnson et al., 2003). Hence, recommendations included in product reviews are said to considerably facilitate the selection process (Park *et al.*, 2012; Zhang *et al.*, 2010). Influential communications no longer derive primarily from advertising but from a variety of consumer-

focused media activities (Lawlor, Dunne, & Rowley, 2016). Online recommendation sources range from reviews published in online portals to customised recommendations created by electronic decision-making aids, namely recommender systems. In fact, recommender systems are so influential to consumers that researchers have suggested they should also be extensively utilized to validate experts' opinions, thus supporting opinion formers, and opinion leaders in their recommendations to consumers (Adomavicius & Tuzhilin, 2001; García-Crespo et al., 2011).

Current progress in digital information technologies is offering a wide range of systems able to support consumers in information processing, thus influencing their purchase choices, for example, holiday planning. The tourism industry is widely adopting *intelligent* digital recommender systems to assist the large number of online platforms in tourism industries (e.g. Tripadvisor, Trivago, booking.com). This takes place in order to provide more pertinent and centred information, and, eventually, enriched tourism experiences, by means of web-mining, context-aware systems, and autonomous agents (Gretzel, 2011).

## 2.2. *Recommender systems*

Recommenders as a stimulating marketing technology offer important value to both consumers and firms. Recommenders assist consumers in learning about products/services through large choice sets, whilst at the same time, benefit firms by converting browsers to buyers, promoting cross-selling and increasing loyalty by providing a custom browsing experience (Lee & Hosanagar, 2019). Recommender systems have been developed for various industries to tackle the problem of information overload (Sun, Guo, & Zhu, 2019). A recommender system consists of personalized information-filtering technology, able to filter a set of items based on consumers' preferences, and thus predict a possible preference (Ghazanfar, 2015; Pantano et al., 2017). The effectiveness of recommender systems derives from the ability to learn

consumers' favourites by analysing their past behaviour responses. In other words, the system learns what consumers prefer starting from their previous choices (machine learning) and identifies future preferences (Borras, Moreno, & Valls, 2014; Noguera, Barranco, Segura, & Martinez, 2012; Park, Kim, Choi, & Kim, 2012). Indeed, these systems can be exploited to provide additional suggestions about places to visit or products of interest dynamically, by gathering real time data on the *state of the consumers* (e.g. location) and adapting accordingly (context-awareness) (Borras et al., 2014). These features characterize intelligent and autonomous agents able to (i) analyse the behaviour of a user, (ii) learn automatically his/her preferences, and (iii) provide advice (recommendations) according to the consumer's profile characteristics (Borras et al., 2014).

Recommender systems are classified by their applications of collaborative, context-aware and hybrid methods (Ghazanfar, 2015). Collaborative systems consider simultaneously the given consumer interest profile with the profiles of other consumers with similar interests (Yang, Cheng, & Dia, 2008). The main limit of collaborative recommender systems concerns their need for more information on users to make recommendations, thus their efficiency is limited in the cases of new users or new items (He, Parra, & Verbert, 2016). Context aware systems adapt the possible recommendation to users' current states (contextual situation), by requiring the proper match between user preferences and contextual factors (Baltrunas Ludwig, Peer, & Ricci, 2012). They often use obtrusive methods to collect contextual factors on consumers' actual states (He et al., 2016). A further evolution of these recommender systems is based on the emotion-aware concept, which exploits emotions as contextual factors by collecting data from consumers shared online via tweets, posts, etc. (Narducci, 2015). Finally, hybrid systems combine the collaborative and content aware filtering approaches (Shinde & Kulkarni, 2012).

Recommendations based on consumer reviews and users' profiles are more influential than promotional content created by marketers, due to the higher credibility of the former compared to the latter (Cheung & Thadani, 2012; Harris & Dennis, 2011). Scholars and practitioners face a challenge to predict information that meets users' requirements, representing a critical issue for data extraction techniques. Going one step further, recommender systems have the ability to analyse the preferences of online users. Based on the profile characteristics of existing and potential users they may suggest appropriate choices to potential users (Felfernig et al., 2013). Thus, electronic recommender systems contribute to enhancing the trustworthiness of recommendations (Bobadilla, Ortega, Hernando, & Gutiérrez, 2013). However, the integration of location information, which is being used in some recommender systems, in conjunction with real-time user-centric data drawn from mobile smart devices, offers opportunities for advanced recommendation capabilities (Knijnenburg et al., 2012).

### *2.3. Predictive Models*

Predictive modelling is a process using data mining techniques to identify a causal relationship between a “dependent” variable (target) and “independent” variables, based on the formulation of a statistical model. Thus, the aim of the predictive model is to predict the future values of the dependent variables, drawing upon the past values.

When numerical datasets are available, literature suggests several established predictive models; for instance, neural networks (NN) (Horikawa et al., 1992), k-nearest neighbour (k-NN) (Weinberger et al., 2009), support vector machines (SVM) (Chapelle et al., 1999), and so on. The selection of the best predictive method is considered a challenge for scholars and practitioners, requiring deep analysis of the actual context, the analysis of the nature of the data (i.e. strings, numbers, etc.), and the computational cost in terms of capability of performing the data and time

for executing the task (for instance in some cases the huge volume of data cannot be analysed by a traditional computer and requires higher-performances machines).

In particular, Support Vector Machines analyse data and identify clear patterns, which can be effectively used for classification analyses. Starting from a certain set of examples for training purposes (labelled by a Boolean value 1/0, in/out, true/false, etc.), an SVM training algorithm renders a model that supports true/false as outcomes. In particular, an SVM can be considered as a mathematical representation of the examples of the set as single points in space, plotted in order to make the Boolean examples (i.e. true and false) divided into two as distant as possible areas in the space. Other instances are further added into that space, while the SVM predicts the value as true or false according to the specific area in the space they fall. Additionally, SVMs may perform a non-linear classification efficiently using the kernel trick, thus mapping their inputs into high-dimensional feature spaces.

Formally, a support vector machine builds a hyper-plane, which can be used for classification, regression, or other tasks. Intuitively, a good separation is achieved by the hyper-plane that has the largest distance to the nearest training-data point of the true and false classes.

The first step consists of the definition of  $D$ , as the set of data used to train the system:

$$D = \{(\vec{x}_i, y_i) \mid \vec{x}_i \in \mathbb{R}^n, y_i \in \{\text{true}, \text{false}\}\}; \quad (1)$$

Since any hyper-plane can be written as the set of points  $\vec{x}_i$  satisfying the equation:

$$\vec{w} \cdot \vec{x} - b = 0; \quad (2)$$

SVM estimates the maximum-margin hyper-plane able to divide the space in two separate regions distinguishing the points having  $y_i=\text{true}$  from those having  $y_i=\text{false}$  (when the training data are linearly separable).

The hyper-planes can be further defined as:

$$\begin{aligned}\vec{w} \cdot \vec{x} - b &= 1 \\ \vec{w} \cdot \vec{x} - b &= -1,\end{aligned}\tag{3} \& (4)$$

The distance between the two emerging hyper-planes is expressed as  $2/\|\vec{w}\|$  (margin); thus, minimizing  $\|\vec{w}\|$ , the system maximizes the distance. Two conditions have to be further set to avoid obtaining any point into the margin (in other words, to clearly separate the two regions in the space):

$$\begin{aligned}\vec{w} \cdot \vec{x} - b &\geq 1 \text{ for } \vec{x} \text{ belonging to the "true" class} \\ \vec{w} \cdot \vec{x} - b &\leq -1 \text{ for } \vec{x} \text{ belonging to the "false" class,}\end{aligned}$$

if  $y_i = \text{true} = +1$  and  $y_i = \text{false} = -1$ , the condition to satisfy is:

$$y_i (\vec{w} \cdot \vec{x} - b) \geq 1,\tag{5}$$

The above-mentioned model is an example of SVM working for Boolean classes, but it can be extended to easily predict more than two classes of values, in other words this example predicts a Boolean variable, but it can be improved to predict more values. To achieve this goal, the new SVM (multiclass SVM) is reduced into multiple binary classification problems (Duan & Keerthi, 2005). In particular, this process is based on two main methods: (i) the development of a machine that distinguishes between the value of one variable and all the other (one-versus-all), in this case the classifier with the highest output function assigns the class to all the new instances; and (ii) the development of a machine that distinguishes between pair of classes (one-versus-one), in this case, each classifier assigns the instance to one of the two categories, then the vote for the assigned class is increased by one vote, and finally the categories with the most votes determines the instance classification.

*Mathematica<sup>TM</sup>* is a powerful software package that has the ability to generate the classification function and predict the value of the dependent variable by automatically choosing the most appropriate methods through an internal algorithm according to the expected outcomes. In particular, this function allows the evaluation of a huge variety of data sets, including numerical, textual, sound, image and a combination of all (Wolfram, 2015). In this case *Mathematica<sup>TM</sup>* serves as the computational platform for data analysis and prediction purposes. In the next section, the methodological steps for this study are illustrated on a step-by-step basis.

### **3. Methodology**

#### *3.1. Case study of new recommender systems for the tourism sector*

As aforementioned, digital recommender systems can analyse open data in order to select the information that might better fit consumers' preferences (through the prediction model), thus recommending products with a higher likelihood to fit into consumers' interests. A new recommender system based on open data analysis is illustrated in Figure 1.

#### **[Figure 1 Here]**

In this new recommender system, data from different sources are processed and converted into a format well-matched with the predictor system. Data are also cleaned, normalized and optimized in advance to make the learning and prediction processes faster and even more accurate. Once the learning phase completes and a fresh dataset from new users is acquired, the system carries on with suggesting a fitting holiday destination via its predictive algorithm. The prediction is based on the principle that users with similar needs, interests and requirements would probably choose similar products, thus the prediction will be more accurate if initial data used for learning are also accurate.

In the tourism destinations area, a relevant application could be based on the selection of open data from TripAdvisor to predict the extent to which a tourist will like a certain destination. In our case, a famous tourism attraction was selected, i.e. Disneyland Park in Paris. This was originally opened as ‘Euro Disney’ theme park in Paris on April 12, 1992. It is part of the wider Disneyland complex in Paris that also includes the Walt Disney Studios Park and seven Disney owned hotels. The whole Disney complex extends over approximately 5,510 acres area. Some 320 million individuals have visited Disneyland Park since opening, with an annual attendance of more than 10 million visitors (DisneyNews, 2017). [A theme park is a useful context for a case study. In particular, Disneyland theme parks are known world-wide and attract visitors of a wide range of ages, with most of them being fairly well-educated \(Geissler & Rucks, 2011; Toyoda, 2014\).](#)

### *3.2. Data analysis*

In order to test our recommender systems’ ability to identify a trend in visitor’s evaluations of their lived experiences in Disneyland Park (Paris), two datasets were formed: one set of 263 Trip Advisor users expressing a totally positive (5 stars, excellent) and another one consisting of 263 different users providing totally negative evaluations (0 stars, terrible) of the chosen attraction, based on their online profiles. The user profiles offer information on 18 topics (Figure 2 shows the 18 categories; each user needs to choose at least three to create a TripAdvisor profile).

**[Figure 2 Here]**

For convenience, we chose to represent these elements in binary mode: a value equal to 1 is an indication of interest to the specific topic, 0 otherwise. Similarly, a value equal to 0 represents a negative evaluation, whereas 1 a positive one (as suggested by Pantano, Priporas, & Stylos, 2017).

In mathematical terms, we consider  $I$  as the set of instances consisting an 18-bit string (which means that the cardinality of the set is 18, as 18 is the number of possible topics identifying

a certain tourist on TripAdvisor, in other words  $n=2^{18}$ ). Thus, our predictive system aims at identifying the  $2^{18}$  values of the target; Figure 3 represents the table of rules (where the blank shape represents a value of “0” and the black one the value of “1”).

**[Figure 3 Here]**

TripAdvisor allows users to rate each attraction with stars (from 0 to 5). For convenience, we can give to each attraction the value 1 if tourists assigned 5 stars and 0 if they gave 0 stars, if considering  $S = \{0,1\}$  as the two possible values of the tourist attraction, the function  $f$  assigning to each data a value of 0 or 1 will be:

$$f: I \rightarrow S \quad (6)$$

$$(x_1, x_2, \dots, x_{18}) \in I \rightarrow s \in S \quad (7)$$

where  $x_i, \forall i=1, \dots, 18$ .

Thus, the task of the system is to identify the function describing these relationships, in other words the goal of the predictive system is to define the  $2^{18}$  target values to predict the values of the element of  $I$  based on the characteristics of the individual TripAdvisor users).

To implement this theoretical framework in the Disneyland case, *Mathematica*<sup>TM</sup> software read the dataset, and then built the set of rules for the training set consisting of 250 items of review data, linking the input data with the expected results. Finally, it built for  $x$  times the prediction function, and compared the findings applying the result emerging from the prediction function with the target value of all the data in the sample.

The experiment ran based on the development of 200 classifying functions. In other words, the system used 200 classifying functions to validate the SVM model by determining the successful cases for the classifying function as proposed by Mathematica. The reliability of the classifier machine can be further evaluated through the sum of percentage values of the cases in which the system identified properly the value of 1 and 0 were determined. The proportion of successful cases

in which the classifier machine identified properly the value of 1 is 0.69, while the proportion of success of 0 is 0.58, which lead to a total value of 1.268 (Figure 4).

**[Figure 4 Here]**

Figure 5 graphically shows the extent to which our classifier machine predicts properly the value (prevision) of the known data (target) (Figure 5) by considering 20 random results, where the blank shape represents the value of “0” and the black one the value of “1”.

**[Figure 5 Here]**

#### **4. Research findings**

Utilizing equations/functions 1 to 7 we reach the graph appearing in Figure 4. The values exceed 1.12 in only two cases, while the best prediction appears on the 184<sup>th</sup> case, as the proportion of successful identification of 1 is 0.77 and the successful identification of 0 is 0.44, for a global value of 1.214. The reliability of the estimated values equals 0.707. To achieve this result, *Mathematica* adopted the Random Forest method. A Random Forest is a particular approach that generates different Decision Trees by randomly cutting out sub-samples of the data observations and sub-sections of the data variables, and then allowing the emerging decision tree models to converge on the better solution (in this case, the system predicts from Breiman-Cutler ensemble of decision trees). Recent approaches to decision trees (as random forests) allow for multi-class classification, as well as ordinal multi-class classification, prediction (Cardoso & Costa, 2007; Frank & Hall, 2001).

The subsequent evaluation of the confusion matrix (error matrix) allows visualizing the performance of the emerging supervised learning machine, where each row represents the instances in the predicted class, and each column denotes the instances in the actual class (Figure 6). From this matrix, the reliability value is 0.607.

**[Figure 6 Here]**

The proposed recommender algorithm is thus able to predict consumers' preferences for a certain product (in this case potential visitation of a tourism attraction/destination), based upon the open data freely accessible via SNSs like TripAdvisor. The results show that the proposed method has a high predictive power and increases the accuracy of recommendations, and thus might usefully support tourists decision-making process. While literature studies on recommender systems are mainly based on the proposal of systems with a higher prediction accuracy, our system would be oriented to provide more precise solutions starting from new input variables (i.e. consumers' facial expression and related emotion) that requires a limited time of performance. In this way, our results reinforce the value of recommender systems in terms of value and time complexity reduction, as solicited by recent researches (Bag et al., 2019).

Hence, the appropriate destinations can be recommended to those tourists with those sets of interests.

## **5. Discussion and conclusion**

Technology and its applications in informatics consist an area of rapid developments during the last two decades (Chen et al., 2015). It has been widely recognized that the emergence of social media platforms, which was matched by the expansion of the smartphones and tablets market (Fotiadis & Stylos, 2017), has greatly influenced consumers' priorities with regards to receiving, storing, processing and evaluating information (Pantano & Priporas, 2016). Consumers' needs for efficient data management and effective individual decision making resulted to the introduction of software applications for recommendations (Gavalas, Konstantopoulos, Mastakas, & Pantzio, 2016).

The introduction of recommender systems in web commerce sites is promoted by the combination of remarkable increases in computers' calculating power, volume decrease of computer devices and new software capabilities, which creates new opportunities for delivering

and customizing information on demand (i.e. cloud computing) (Hashem et al., 2015). Recommender systems facilitate consumers in high-quality decision-making buying processes in e-commerce settings, since the options are more tailored to their needs and preferences, less time consuming, require less effort and thus, potentially result in higher user satisfaction (Farokhi, Vahid, Nilashi, & Ibrahim, 2016; Scholz, Dorner, Schryen, & Benlian, 2017). The interest on behalf of academics and practitioners in this area continues unabated due to growing user demands, despite the numerous research and development advances (Khusro, Ali, & Ullah, 2016).

Although past studies emphasize the need for supporting consumers' decision-making utilising recommender systems (Barragáns-Martínez, Costa-Montenegro, & Juncal-Martínez, 2015), researchers have argued that personalized online applications in various consumer markets are scarce (Wang, 2015; Yeh & Cheng, 2015). The present paper proposes a new recommender algorithm for tourists by exploiting open data. It responds to the call of Gretzel (2011), who highlights the need for more efficient recommender systems in delivering highly customized recommendations to online platform users with several tourism-related choices (i.e. from the choice of a restaurant, to the choice of a gift, to a book, to an activity for the weekend with friends, etc.). In particular, this paper presents an estimation method that is based on ratings pertaining to users-specified criteria (profile) and, ultimately, suggesting a product with a major consumer adoption likelihood.

In summary, this proposed novel recommender system is able to: 1) extract tourist profiles via TripAdvisor (including the 18 above mentioned topics) to produce a dataset that can be processed by the SVM; 2) collect knowledge concerning tourism destinations via TripAdvisor; 3) convert heterogeneous data, such as images, texts, and videos that can be processed in order to run via the predictive system; 4) be properly trained; and 5) create a representation of users' data as points in space; hence, every point in space will correspond to

the most appropriate classification, i.e. in this case a location. All in all, the proposed computational framework predicts the tourism destination that would potentially best fit system users' requirements as entered in the algorithm. Thus, this paper builds upon the work of Pantano and colleagues (2017) by clearly showing the applicability of the proposed predictive model in delivering highly efficient and reliable recommender systems within a tourism destinations and theme parks context.

This study contributes to the literature on recommender systems by providing additional findings based on open data analysis concerning the facilitation of consumer decision making processes. It offers further evidence about recommender systems as Memon et al. (2015), who highlighted the need for high-quality recommender systems to render optimized recommendations to tourists across a wide spectrum of applications. The recommender system algorithm proposed in this study shows a high predictability of tourists' choices, thus demonstrating its high effectiveness in providing suggestions. Therefore, it further enhances the researchers' toolbox with an effective alternative to other recommender algorithms provided in previous research studies (Baltrunas et al., 2012; Chu et al., 2016; Moreno et al., 2013; Nilashi, Bin Ibrahim, Ithnin, & Sarmin, 2015). This study also responds to the call of Pantano et al. (2017) for testing the applicability of the Random Forest decision trees technique. The findings not only provide support for the implementation of the particular computational approach, but also demonstrate for a specific tourism attraction the high degree of visitors' preferences predictability based on the recommendations produced by other visitors to that tourism attraction. On the practical side, organisations/destinations can acquire useful information on their products/services, can better reach their clientele and can design products/services that meet their clientele's needs and wants. Also, the proposed system adds value in the form of recommendations, as it avoids information overflow and can thus increase users' satisfaction which to a large extent depends on the first interactions with the system and

at the same time is an important factor for a recommender systems' success and acceptance (Su & Khoshgoftaar, 2009).

Notwithstanding that the current study introduced a functional real-life recommender system using a novel approach, there are some unavoidable limitations. The study used a relatively small data set, whereas larger data sets might offer a better validation of the proposed recommender system. Also, this study focused on a specific tourism setting, so future works could further test the applicability of the recommender system to other types of tourism destinations. In any case, the authors believe that, given the focus of this paper that is to propose a new recommender system, the limitations mentioned do not substantially undermine the value of this study.

The methodological approach followed in this paper responds to the call of Williams, Dwivedi, Lal, and Schwarz (2009) 'for authors [of technology adoption studies] to make greater use of the theoretical and methodological variety available to them' (Williams et al., 2009, p.9). Indeed, the current study offers both quantitative researchers working in decision sciences and, more specifically, scholars investigating tourism destination selection an additional valuable tool for predictive analysis and decision-making process evaluation. As the paper seeks to open new avenues in decision-making, it has tried to balance the technical aspect of the tool proposed with the case study context to cover both the wider audience of computational analysts, as well as the more specific research group of tourism scholars, respectively.

In sum, the present work provides a new recommender system for consumers built on open data analytics, which to the authors' knowledge, is novel in the tourism marketing literature. In particular, it presents a rating estimation method using ratings that pertain to users-specified criteria (profile) to suggest a product with a major likelihood of consumers purchase. User behavior is complex and dynamic, since consumers' preferences can change over time

(Ansari, Essegai, & Kohli, 2000, Mobasher, Cooley, & Srivastava, 2000; Scholz et al., 2015).

It would therefore be interesting for forthcoming studies to research the effect on quality of recommendations when tourists recommend the same product/service many times in applications where they are allowed to do so more than once.

## References

- Adomavicius, G., & Tuzhilin, A. (2001). Using data mining methods to build customer profiles. *Computer*, 34(2), 74-82.
- Al-Hassan, M., Lu, H. & Lu, J. (2015). A semantic enhanced hybrid recommendation approach: a case study of e-Government tourism service recommendation system. *Decision Support Systems*, 72, 97-109.
- Al-Quesi, K., Dennis, C., Alamanos, E., & Jayawardhena, C. (2014). Website design quality and usage behaviour: unified theory of acceptance and use of technology. *Journal of Business Research*, 67(11), 2282-2290.
- Alnawas, I., & Aburub, F. (2016). The effect of benefits generated from interacting with branded mobile apps on consumer satisfaction and purchase intentions. *Journal of Retailing and Consumer Services*, 31, 313-322.
- Amblee, N., & Bui, T. (2011). Harnessing the influence of social proof in online shopping: The effect of electronic word of mouth on sales of digital microproducts. *International Journal of Electronic Commerce*, 16(2), 91-114.
- Ansari, A., Essegai, S., & Kohli, R. (2000). Internet recommendation systems. *Journal of Marketing Research*, 37(3), 363-375.
- Arts, J.W.C., Frambach R.T., & Bijmolt T.H.A. (2011). Generalizations on consumer innovation adoption: a meta-analysis on drivers of intention and behaviour. *International Journal of Marketing Research*, 28(2), 134-144.

- Bag, S., Kumar, S., Awasthi, A., & Tiwari, M. K. (2019). A noise correction-based approach to support a recommender system in a highly sparse rating environment. *Decision Support Systems*, 118, 46-57.
- Balaji, M. S., Khong, K. W., & Chong, A. Y. L. (2016). Determinants of negative word-of-mouth communication using social networking sites. *Information & Management*, 53(3), 528-540.
- Baltrunas, L., Ludwig, B., Peer, S. & Ricci, F. (2012). Context relevance assessment and exploitation in mobile recommender systems. *Personalized Ubiquitous Computing*, 16, 507-526.
- Barragáns-Martínez, B., Costa-Montenegro, E. & Juncal-Martínez, J. (2015). Developing a recommender system in a consumer electronic device. *Expert Systems with Applications*, 42(9), 4216-4228.
- Blazquez, M. (2014). Fashion shopping in multichannel retail: the role of technology in enhancing the customer experience. *International Journal of Electronic Commerce*, 18(4), 97-116.
- Bobadilla, J., Ortega, F., Hernando, A., & Gutiérrez, A. (2013). Recommender systems survey. *Knowledge-Based Systems*, 46, 109-132.
- Bonhard, P., & Sasse, M. A. (2006). Knowing me, knowing you'—Using profiles and social networking to improve recommender systems. *BT Technology Journal*, 24(3), 84-98.
- Borras, J., Moreno, A., & Valls, A. (2014). Intelligent tourism recommender systems: a survey. *Expert Systems with Applications*, 41(15), 7370-7389.
- Boser, B. E., Guyon, I. M., and Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, ACM, 144-152.

- Cardoso, J. S., & Costa, J. F. (2007). Learning to classify ordinal data: The data replication method. *Journal of Machine Learning Research*, 1393-1429.
- Chapelle, O., Haffner, P., & Vapnik, V. N. (1999). Support vector machines for histogram-based image classification. *IEEE transactions on Neural Networks*, 10(5), 1055-1064.
- Chen, Y. F. (2008). Herd behavior in purchasing books online. *Computers in Human Behavior*, 24(5), 1977-1992.
- Chen, M., Zhang, Y., Hu, L., Taleb, T., & Sheng, Z. (2015). Cloud-based wireless network: virtualized, reconfigurable, smart wireless network to enable 5G technologies. *Mobile Networks and Applications*, 20(6), 704-712.
- Chu, S. C., & Kamal, S. (2008). The effect of perceived blogger credibility and argument quality on message elaboration and brand attitudes: An exploratory study. *Journal of Interactive Advertising*, 8(2), 26-37.
- Chu, Y., Wang, H., Zheng, L., Wang, Z., & Tan, K. L. (2016). TRSO: A tourism recommender system based on ontology. *International Conference on Knowledge Science, engineering and management*, Springer International Publishing, 567-579.
- Colace, F., De Santo, M., Greco, L., Moscato, V., & Picariello, A. (2015). A collaborative user-centered framework for recommending items in Online Social Networks. *Computers in Human Behavior*, 51, 694-704.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340.
- Dennis, C., Brakus, J. J., Gupta, S. and Alamanos, E. (2014). The effect of digital signage on shoppers' behaviour: the role of the evoked experience. *Journal of Business Research*, 67(11), 2250-2257.
- Dennis, C, Jayawardhena, C., Merrilees, W., & Wright, L. T. (2009). e-Consumer behaviour. *European Journal of Marketing*, 43(9/10), 1121-1139.

- DisneyNews (2017). 25 Interesting Disneyland Paris Statistics and Fun Facts (June 2017),  
Retreived from: <https://disneynews.us/disneyland-paris-statistics-fun-facts/>, on  
October, 14, 2017.
- Duan, K. B., & Keerthi, S. S. (2005). Which is the best multiclass SVM method? An empirical  
study. Multiple classifier systems. *Lecture Notes on Computer Science*, 3541, 278–285.
- Farokhi, N., Vahid, M., Nilashi, M., & Ibrahim, O. (2016). A multi-criteria recommender  
system for tourism using fuzzy approach. *Journal of Soft Computing and Decision  
Support Systems*, 3(4), 19-29.
- Felfernig, A., Jeran, M., Ninaus, G., Reinfrank, F., Reiterer, S. (2013). Toward the next  
generation of recommender systems: applications and research challenges. *Multimedia  
services in intelligent environments*, Springer International Publishing, 81-98.
- Fotiadis, A. K., & Stylos, N. (2017). The effects of online social networking on retail consumer  
dynamics in the attractions industry: The case of ‘E-da’theme park, Taiwan.  
*Technological Forecasting and Social Change*, 124, 283-294.
- Frank, E., & Hall, M. (2001). A simple approach to ordinal classification. In De Raedt, L., &  
Flach, P. (eds), *Machine Learning: ECML 2001*. Berlin: Srpinge (pp. 145-156).
- Freedman, D. A. (2005). *Statistical models: Theory and practice*. New York: Cambridge  
University Press.
- García-Crespo, Á., López-Cuadrado, J. L., Colomo-Palacios, R., González-Carrasco, I., &  
Ruiz-Mezcua, B. (2011). Sem-Fit: A semantic based expert system to provide  
recommendations in the tourism domain. *Expert systems with applications*, 38(10),  
13310-13319.
- Gavalas, D., Konstantopoulos, C., Mastakas, K., & Pantziou, G. (2014). Mobile recommender  
systems in tourism. *Journal of Network and Computer Applications*, 39, 319-333.

- Geissler, G. L., & Rucks, C. T. (2011). The overall theme park experience: A visitor satisfaction tracking study. *Journal of Vacation Marketing*, 17(2), 127-138.
- Ghazanfar, M. A. (2015). Experimenting switching hybrid recommender systems. *Intelligent Data Analysis*, 19, 845-877.
- Gretzel, U., (2011). Intelligent systems in tourism: a social science perspective. *Annals of Tourism Research*, 38(3), 757-779.
- Harris, L., & Dennis, C. (2011). Engaging customers on Facebook: challenges for e-retailers. *Journal of Consumer Behaviour*, 10(6), 338-346.
- Hashem, I.A.T., Yaqoob, I., Anuar, N.B., Mokhtar, S., Gani, A., Khan, S. U. (2015). The rise of “big data” on cloud computing: Review and open research issues. *Information Systems*, 47, 98-115.
- He, C., Parra, D., & Verbert, K. (2016). Interactive recommender systems: a survey of the state of the art and future research challenges and opportunities. *Expert Systems with Applications*, 56, 9-27.
- Horikawa, S. I., Furuhashi, T., & Uchikawa, Y. (1992). On fuzzy modeling using fuzzy neural networks with the back-propagation algorithm. *IEEE transactions on Neural Networks*, 3(5), 801-806.
- Johnson, E.J., Bellman, S., & Lohse, G. L. (2003). Cognitive lock-in and the power law of practice., *Journal of Marketing*, 67(2), 62-75.
- Kang, J.-Y. M., Mun, J. M., & Johnson, K.K.P. (2015). In-store mobile usage: downloading and usage intention toward mobile location-based retail apps. *Computers in Human Behaviour*, 46, 210-217.
- Knijnenburg, B. P., Willemsen, M. C., Gantner, Z., Soncu, H., & Newell, C. (2012). Explaining the user experience of recommender systems. *User Modeling and User-Adapted Interaction*, 22(4-5), 441-504.

- Khusro, S., Ali, Z., Ullah, I. (2016). Recommender systems: Issues, challenges, and research opportunities. *Information Science and Applications (ICISA) 2016*, 1179-1189.
- Lawlor, M., Dunne, Á., Rowley, J. (2016). Young consumers' brand communications literacy in a social networking site context. *European Journal of Marketing*, 50(11), 2018-2040.
- Lee, D., & Hosanagar, K. (2019). How do recommender systems affect sales diversity? a cross-category investigation via randomized field experiment. *Information Systems Research*, <https://doi.org/10.1287/isre.2018.0800>
- Liao, S. H., Chu, P. H., & Hsiao, P. Y. (2012). Data mining techniques and applications—A decade review from 2000 to 2011. *Expert systems with applications*, 39(12), 11303-11311.
- Liu, D. R., & Shih, Y. Y. (2005). Integrating AHP and data mining for product recommendation based on customer lifetime value. *Information & Management*, 42(3), 387-400.
- Lowe, B., & Alpert, F. (2015). Forecasting consumer perception of innovativeness. *Technovation*, 45-46, 1-14.
- MacKay, D. J. (2003). *Information theory, inference and learning algorithms*. Cambridge University Press.
- Maity, M., & Dass, M. (2014). Consumer decision-making across modern and traditional channels: e-commerce, m-commerce, in-store. *Decision Support Systems*, 61, 34-46.
- Memon, I., Chen, L., Majid, A., Lv, M., Hussain, I., & Chen, G. (2015). Travel recommendation using geo-tagged photos in social media for tourist. *Wireless Personal Communications*, 80(4), 1347-1362.
- Mobasher, B., Cooley, R., & Srivastava, J. (2000). Automatic personalization based on web usage mining. *Communications of the ACM*, 43(8), 142-151.
- Moreno, A., Valls, A., Isern, D., Marin, L., & Borràs, J. (2013). Sigtur/e-destination: ontology-based personalized recommendation of tourism and leisure activities. *Engineering*

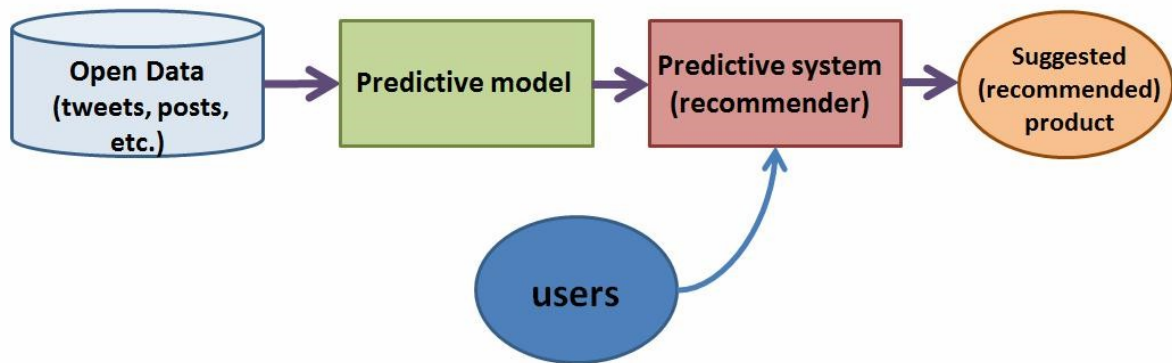
*Applications of Artificial Intelligence*, 26(1), 633-651.

- Narducci, F., De Gemmis, M., & Lops, P. (2015). A general architecture for an emotion-aware content-based recommender system. Proceedings of the 3<sup>rd</sup> Workshop on Emotions and Personality in Personalized Systems (EMPIRE), ACM, 3-6.
- Ng, M. (2016). Factors influencing the consumer adoption of Facebook: a two country study of youth markets. *Computers in Human Behavior*, 54, 491-500.
- Nilashi, M., Bin Ibrahim, O., Ithnin, N. and Sarmin, N. H. (2015). A multi-criteria collaborative filtering recommender system for the tourism domain using Expectation Maximization (EM) and PCA–ANFIS2. *Electronic Commerce Research and Applications*, 14(6), 542-562.
- Noguera, J. M., Barranco, M., Segura, R. and Martinez, L. (2012). A mobile 3D-GIS hybrid recommender system for tourism. *Information Sciences*, 215, 37-52.
- Okazaki, S., Díaz-Martín, A. M., Rozano, M., & Menéndez-Benito, H. D. (2015). Using Twitter to engage with customers: A data mining approach. *Internet Research*, 25(3), 416-434.
- Oliveira, J., Delgado, C., & Assaife, A. C. (2017). A recommendation approach for consuming linked open data. *Expert Systems with Applications*, 72, 407-420.
- Pantano, E., & Viassone, M. (2014). Demand pull and technology push perspective in technology-based innovations for the points of sale: The retailers evaluation. *Journal of Retailing and Consumer Services*, 21(1), 43-47.
- Pantano, E., and Di Blasi, G. (2015), “Big data analysis solutions for supporting tourist’s decision making”, in K. Andriotis (Ed.), *Proceedings of the international conference on tourism (ICOT2015)* in London, UK, 24-27 June, International Association for Tourism Policy, pp.440-452.

- Pantano, E., & Priporas, C.V. (2016). The effect of mobile retailing on consumers' purchasing experiences: A dynamic perspective. *Computers in Human Behavior*, 61, 548-555.
- Pantano, E., Priporas, C. V., & Stylos, N. (2017). ‘You will like it!’ using open data to predict tourists' response to a tourist attraction. *Tourism Management*, 60, 430-438.
- Park, D. H., & Kim, S. (2009). The effects of consumer knowledge on message processing of electronic word-of-mouth via online consumer reviews. *Electronic Commerce Research and Applications*, 7(4), 399-410.
- Park, D. H., Kim, H. K., Choi, I. Y. and Kim, J. K. (2012). A literature review and classification of recommender systems research. *Expert Systems with Applications*, 39, 10059-10072.
- Quinn, L., Dibb, S., Simkin, L., Canhoto, A. and Analogbei, M. (2016). Troubled waters: the transformation of marketing in a digital world. *European Journal of Marketing*, 50(12), 2103-2133.
- Scholz, M., Dorner, V., Schryen, G., & Benlian, A. (2017). A configuration-based recommender system for supporting e-commerce decisions. *European Journal of Operational Research*, 259(1), 205-215.
- Scholz, M., Dorner, V., Franz, M., & Hinz, O. (2015). Measuring consumers' willingness to pay with utility-based recommendation systems. *Decision Support Systems*, 72, 60-71.
- Senecal, S., & Nantel, J. (2004). The influence of online product recommendations on consumers' online choices. *Journal of Retailing*, 80(2), 159-169.
- Shinde, S. K., & Kulkarni, U. (2012). Hybrid personalized recommender system using centering-bunching based clustering algorithm. *Expert Systems with Applications*, 39, 1381-1387.
- Su, X., and Khoshgoftaar, T.M. (2009), “A survey of collaborative filtering techniques”, *Advances in Artificial Intelligence*, 4, doi:10.1155/2009/421425

- Sudheep, E. M., Idikkula, S. M., & Alexander, J. (2011). Design and performance analysis of data mining techniques based on decision trees and naive bayes classifier for employment chance prediction. *Journal of Convergence Information Technology*, 6(5), 89.
- Sun, L., Guo, J., & Zhu, Y. (2019). Applying uncertainty theory into the restaurant recommender system based on sentiment analysis of online Chinese reviews. *World Wide Web*, 22, 83-100.
- Suykens, J. A., & Vandewalle, J. (1999). Least squares support vector machine classifiers. *Neural processing letters*, 9(3), 293-300.
- Toyoda, Y. (2014). Recontextualizing Disney: Tokyo Disney resort as a kingdom of dreams and magic. *Social Science Japan Journal*, 17(2), 207-226.
- Venkatesh, V., Thong J. Y. L., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Quarterly*, 36(1), 157-178.
- Verhoef, P. C., Lemon, K. N., Parasuraman, A., Roggeveen, A., Tsiros, M., & Schlesinger, L. A. (2009). "Customer experience creation: determinants, dynamics and management strategies. *Journal of Retailing*, 85(1), 31-41.
- Wang, C. H. (2015). A market-oriented approach to accomplish product positioning and product recommendation for smart phones and wearable devices. *International Journal of Production Research*, 53(8), 2542-2553.
- Ward, J., & Peppard, J. (2016). *The strategic management of information systems: Building a digital strategy*. (4<sup>th</sup> ed.). Hoboken: John Wiley & Sons.
- Weinberger, K. Q., & Saul, L. K. (2009). Distance metric learning for large margin nearest neighbor classification. *Journal of Machine Learning Research*, 10(Feb), 207-244.

- Williams, M., Dwivedi, Y., Lal, B., & Schwarz, A. (2009). Contemporary trends and issues in IT adoption and diffusion research. *Journal of Information Technology*, 24(1), 1-10.
- Willemsen, M. C., & Johnson, E. J. (2011). Visiting the decision factory: Observing cognition with MouselabWEB and other information acquisition methods. *A handbook of process tracing methods for decision research: A critical review and user's guide*, 21-42.
- Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann.
- Yang W.-S., Cheng H.-C., & Dia J.-B. (2008). A location-aware recommender system for mobile shopping environments. *Expert Systems with Applications*, 34, 437-445.
- Yeh, D. Y., & Cheng, C. H. (2015). Recommendation system for popular tourist attractions in Taiwan using Delphi panel and repertory grid techniques. *Tourism Management*, 46, 164-176.
- Zhang, D., Zhou, L., Kehoe, J. L., & Kilic, I. Y. (2016). What online reviewer behaviors really matter? Effects of verbal and nonverbal behaviors on detection of fake online reviews. *Journal of Management Information Systems*, 33(2), 456-481.
- Zhang, Z. K., Zhou, T., & Zhang, Y. C. (2010). Personalized recommendation via integrated diffusion on user–item–tag tripartite graphs. *Physica A: Statistical Mechanics and its Applications*, 389(1), 179-186.



**Figure 1:** An overview of the predictive (recommender) system for selecting the product that best match users' requirements. (Source: The current authors)

## What kind of traveller are you?

Select 3 or more tags to include on your member profile

The interface features a progress bar at the top with three numbered steps: 1, 2, and 3. Below the progress bar, there are 18 tags arranged in four rows. The tags are: Shopping Fanatic, Foodie, Urban Explorer, Eco-tourist, Backpacker, Beach Goer, Nature Lover, Peace and Quiet Seeker, 60+ Traveller, Vegetarian, Nightlife Seeker, History Buff, Like a Local, Luxury Traveller, Trendsetter, Art and Architecture Lover, Family Holiday Maker, Thrifty Traveller, and Thrill Seeker. At the bottom, there are two buttons: a green 'Done' button and a blue 'Cancel' button.

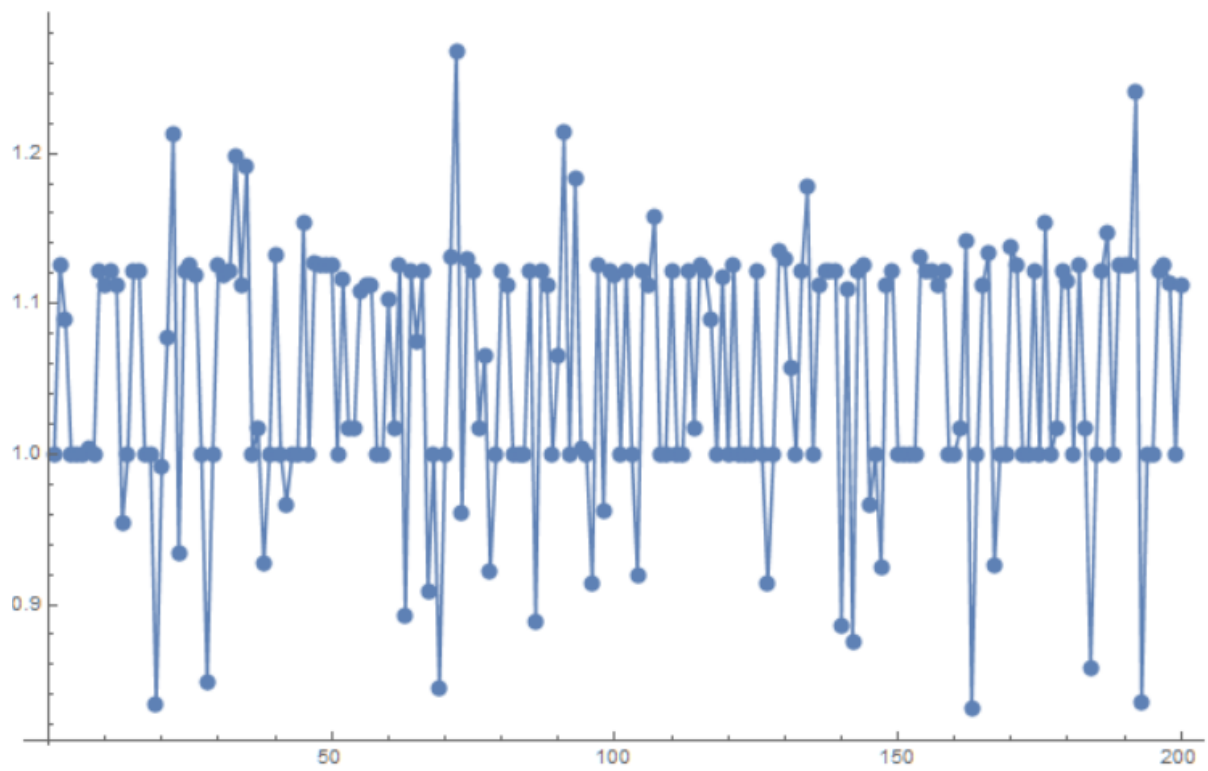
Shopping Fanatic	Foodie	Urban Explorer	Eco-tourist	Backpacker
Beach Goer	Nature Lover	Peace and Quiet Seeker	60+ Traveller	Vegetarian
Nightlife Seeker	History Buff	Like a Local	Luxury Traveller	Trendsetter
Art and Architecture Lover	Family Holiday Maker	Thrifty Traveller	Thrill Seeker	

Done Cancel

**Figure 2:** The 18 possible characteristics defining a tourist to be chosen to create a profile on TripAdvisor.

Data	Target
□□□□□□□□□□□□□□□□	?
□□□□□□□□□□□□□□■	?
□□□□□□□□□□□□□■□	?
□□□□□□□□□□□□■□■	?
□□□□□□□□□□□■□□□	?
□□□□□□□□□□■□□■	?
□□□□□□□□□■□■□□	?
□□□□□□□□■□■□■	?
□□□□□□□■□■□■□	?
□□□□□□■□■□■□■	?
□□□□□■□■□■□□□	?
□□□□■□■□■□□■	?
□□□■□■□■□□■□	?
□□■□■□■□■□□■	?
□□■□■□■□■□■□	?
□■□■□■□■□■□□	?
□■□■□■□■□■□■	?
■□■□■□■□■□□□	?
■□■□■□■□■□□■	?
■□■□■□■□■□■□	?
■□■□■□■□■□■□■	?
...	?

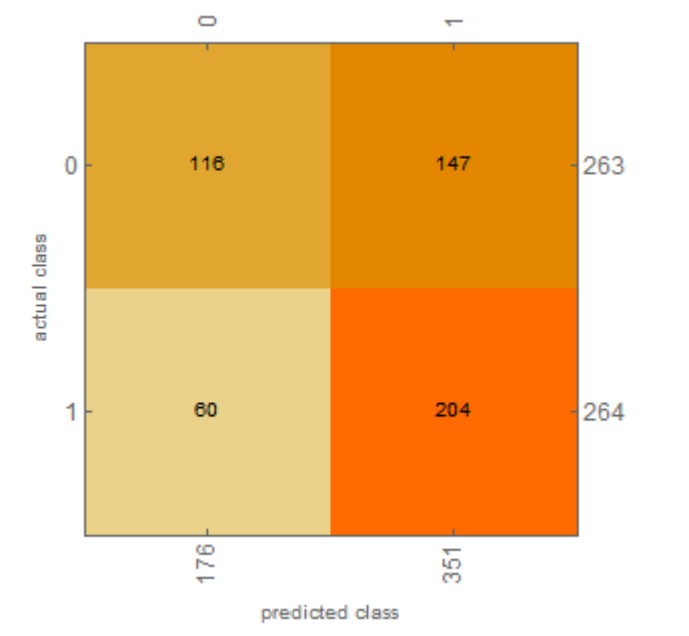
**Figure 3:** Table of rules for the classifier function (aiming at identifying the  $2^{18}$  target values).



**Figure 4:** Experiment results based on building 200 classifying functions.

Data	Target	Prevision
□□□■□□■□□□□□□□□□□■	□	■
■■□□■□□□□□□□□□□□■□	■	■
□□■□□□□□□□□□□□■□■□	□	■
□□□□□□□□■□□■□□□■□□	□	■
■■□□□□■□□□□□□□□□■□□	■	□
■□□■□□□□■□□□□□■□■□□	■	■
■□□□■□■□□□□□■□□□□□	□	□
□□■□□□□□■□□□□□■□□	□	□
□□□■□□□□□□□■□■□□□□	■	■
□□□□□■□□□□□□□□□■□□	■	■
■□□■□□□□□□□□□□□□■	■	■
■□□■□■□□■□□□□■□■□■□	■	■
□□■□□□□□□□□□□□□□■	■	■
□□□■□■□□□□□□□□□□■□	□	■
■□□■□■□□□□□□□□■□□□■	■	□
□□■□■□□□□□□□□□□□■	■	■
■□■□□□■□□□□□□□□■□□	■	■
□□□□□□□□□■□□■□■□□□	□	□
■□□■□■□■□□□□■□□□□□■	■	■
■□□■□□□□□□□□□□□□■□■	□	■

**Figure 5:** Table of rules for 20 random results.



**Figure 6:** Confusion matrix.