

Tourist services management through clients scoring using a bio-inspired agent architecture [★]

Raúl Moreno, Alberto Fernández-Isabel, Víctor Aceña,
Isaac Martín de Diego, and Javier M. Moguerza

Data Science Lab. at Rey Juan Carlos University
C/ Tulipán, s/n, Móstoles, 28933, Madrid, Spain
r.morenoi@alumnos.urjc.es, alberto.fernandez.isabel@urjc.es,
victor.acena@urjc.es, isaac.martin@urjc.es, javier.moguerza@urjc.es

Abstract. Tourism has become an economic engine for several countries during the last decades. Each time more and more individuals consider visiting other places during their vacation period. These places cover very different options, from the typical beach and mountain tourism to the less common urban and cultural trips. These travels hoard multiple transport means and facilities in the destination place that have to be correctly managed. Therefore, the rise of automatic systems to address the related operations and processes is a crucial issue nowadays. These systems are usually focused on the final users (the tourists) and make recommendations about their available possibilities. However, it is not easy to find a multi-purpose recommendation system covering all the needs from the perspective of the travel agency. In this paper, a complete framework called *Pharaoh* able to make recommendations to customers covering the final user perspective, and to provide support to the travel agents, is presented. This assistance filters the best travel, accommodation, and activity options according to the desires of the customers. This novel functionality allows selecting the customer with the best propensity to book a tourist service. This workload is distributed using a bio-inspired Multi-Agent System (MAS). Moreover, *Pharaoh* considers the feedback from clients after the completion of the tourist opportunity to improve future recommendations. Several experiments in real environments have been addressed to show the viability of the proposal. It can be concluded that the system enhances the quality of the service provided by the travel agency and its profits.

Keywords: Tourism, Recommendation system, Propensity assessment, Intelligent agents, Smart assistant framework

1. Introduction

Today, tourism is one of the world's most profitable industries [10]. Each year several tourists travel to visit famous cities and mythic architectures or to meet new cultures and ways of living. Moreover, other tourists only look for revelry or calm and relaxation during vacation time, where wonderful landscapes and beaches light up the stage [42].

[★] It is an advanced release of a previously published conference paper: Moreno, R., Viajes, M., Fernández-Isabel, A., de Diego, I. M., Moguerza, J. M., Lancho, C., & Cuesta, M. (2022, September). Automatic detection of potential customers by opinion mining and intelligent agents. In 2022 17th Conference on Computer Science and Intelligence Systems (FedCSIS) (pp. 93-101). IEEE

1 All these travel events produce the use of several means of transport like airplanes,
2 trains, buses, and cars [12]. These events generate millions of bookings in hotels, apart-
3 ments, and other similar buildings specifically dedicated to hosting tourists.

4 All these issues must be managed by travel agencies, retrieving the information from
5 destination places, and the preferences of the different tourists [34]. Therefore, the ap-
6 pearance of systems focused on providing support to the final user (the tourist and the
7 travel agent) that simplifies the booking process is necessary.

8 Furthermore, travel agents also need support to provide customers with the most in-
9 teresting options to capture their interest. Therefore, systems able to manage this task
10 automatically, making recommendations of the most suitable places and trips, or select-
11 ing the best transport schedules according to the client's needs are also basic [46].

12 Nevertheless, these systems have their limitations. The main bottleneck consists of
13 the difficulty of elaborating budgets by the travel agency employee. It is a very time-
14 consuming task since the configuration of a trip increases its complexity when multiple
15 destinations are selected. Another relevant problem is the typical inability to provide feed-
16 back about the tourist services offered to clients. This point is essential for travel agencies
17 because they need to know if the work achieved was adequate and if customers were sat-
18 isfied after finishing the contracted tourist event. In this sense, the recommendations that
19 filter the initial preferences of the clients are the key point. These usually mark the proper
20 evolution of the tourist services, as they must capture the interest of the client in the first
21 instance.

22 The COVID-19 outbreak is another point to consider in the tourism area. The pan-
23 demic situation all over the world has caused governments to implement several restric-
24 tions to control the spreading of the coronavirus. This has had a drastic impact on the mo-
25 bility of the population, reducing the number of travels, and, as a consequence, tourism
26 has suffered significant economic losses [6]. However, with the advances in vaccination
27 programs and the relaxation of the measures against the contagion of the virus, the rise of
28 tourism is going to become a reality very soon.

29 For these reasons, it is mandatory the development of a complete framework that
30 can cushion this new growing demand. This system must be able to address the weak
31 points detected, managing and encompassing all the possible issues produced since a
32 client shows interest in a tourist event, providing support to the travel agent in the selection
33 of the possibilities, and finally obtaining the feedback of the client when the tourist service
34 concludes.

35 Notice that the strength of a framework of this type lies in the management of tourist
36 packages where multiple destinations and services are contracted. Typical examples of
37 these trips are, among others, honeymoons and cruises.

38 In this regard, this paper presents the *Pharaoh* framework. It is an automatic analytical
39 Enterprise Resource Planning (ERP) and Customer Relationship Management (CRM)
40 system for Travel Agencies that includes several graphical assistants to provide complete
41 support to travel agents and tourists. Its novelty resides in the next points:

- 42 – It covers all the processes related to tourist services sales, focusing mainly on the
43 interest of the client and the specific bargains that fit better for the travel agency.
44 Thus, travel agents and clients are guided synchronously by the opinion produced by
45 the system at each step of the sale.

- 1 – It generates budgets in a fast way and with quotations in real time. This fact facilitates
2 the decision-making tasks of the travel agencies and the client can be informed at the
3 moment about the best individual available options or combinations in specific tourist
4 packages.
- 5 – It can learn from the experience and can prioritize those clients who are prone to com-
6 mit the book of tourist services. Notice that this ability is one of the most differential
7 strengths of the proposal regarding other existing approaches. An Machine Learning
8 (ML) model to estimate the probability of the sale has been included in the system to
9 achieve that point.
- 10 – Intelligent agents following a Multi-Agent System (MAS) have been included to es-
11 tablish communication and knowledge interchange. This MAS has been built accord-
12 ing to the rules of a bio-inspired anthill, efficiently promoting the distribution of the
13 workload [35]. This is relevant because companies usually need real-time interac-
14 tions to satisfy the requests made by customers and maximize the benefits through
15 saving computer resources (e.g., possible expenses in specific cloud computing ar-
16 chitectures).

17 A set of experiments has been achieved to show the viability of the proposal. The
18 module in charge of selecting the best clients has been tested independently to evaluate
19 its potential. Then, the complete system and its functionalities have been put into the
20 spotlight. Promising results have been obtained from these tests.

21 The remainder of the paper is organized as follows. Section 2 introduces the founda-
22 tions and relevant literature. Section 3 details the architecture and the different features
23 of *Pharaoh*. Section 4 presents a set of experiments to illustrate the performance of the
24 system. Finally, Section 5 concludes and provides further guidelines.

25 **2. Background**

26 This section introduces the foundations of the *Pharaoh* framework, covering the different
27 perspectives included in it together with the existing state of the art in the tourism domain.
28 First, an overview related to the knowledge management systems is achieved. There, it is
29 detailed how this kind of system works, and similar approaches that provide interesting
30 features are addressed. Second, recommendation systems are introduced explaining how
31 they are used in real environments. The different perspectives of these systems are also
32 detailed in this point. Finally, previous approaches oriented to managing tourism-related
33 processes by travel agencies and customers are presented.

34 **2.1. Recommendation systems**

35 Recommendation systems are mainly focused on filtering the information they store. This
36 filter is flexible and adaptable to the requirements and interests of the users [1].

37 Delving into the filtering process, these systems produce a profile for each one of
38 the users. This profile contains the different parameters that are used to measure (and
39 consequently compare) the interests of the user.

40 The parameters that produce the user's profile can be provided to the system dur-
41 ing the registration process of the user, or they can be included and modified during the

1 interaction of the user with the system. This latter is the more relevant, as it allows the
2 system adaptation to possible fluctuating opinions and interests of the users. Moreover,
3 it also eases the improvement of the system, making it more accurate in the recommen-
4 dations each time, as it can know if the previous recommendations were satisfied for the
5 user. A well-known instance of this approach is the recommendation systems based on
6 reinforcement learning [24].

7 Recommendation systems usually use three main perspectives to achieve the recom-
8 mendation process: the selected topics, the relevance of the content, and the associated
9 reputation. Notice that these perspectives are not excluded and they could be found mixed
10 in some recommendation systems.

11 Systems that consider the topic recommendation perspective focus on the most com-
12 mon elements the users usually select. Thus, the software learns from the user according
13 to the preferences [30]. For instance, if a user shows interest in beach destinations on a
14 travel website, the recommendations should be related to other beach spots they might find
15 appealing. This method aims to offer visitors new yet potentially attractive destinations,
16 based on the preferences of users with similar profiles.

17 Systems for the relevance of content perspective are focused on the importance or
18 popularity of the element to make recommendations. This importance is measured ac-
19 cording to the number of users interacting with the product. Thus, these systems usually
20 generate popularity rankings to promote the easy consultation of the products [25]. In
21 tourism recommendation systems, if a point of interest attracts many visitors and stands
22 out for its popularity at a given time, the system highlights it, presenting it among the top
23 suggestions to the user.

24 A system based on reputation provides recommendations according to the obtained
25 feedback [32]. In hotel competitiveness, where good service is crucial, positive feedback
26 boosts the hotel's importance, improving its reputation. In contrast, negative feedback
27 lowers its value, affecting its place in the market. They are specifically designed to reduce
28 the uncertainty for the users and facilitate trust between entities [7]. A well-known in-
29 stance is shared online where several users sell different products. Those systems usually
30 support products and users with a better reputation.

31 In the case of the *Pharaoh* framework, it is a recommendation system that includes
32 the three perspectives. Thus, it can organize different travels according to the topic (e.g.,
33 business, relaxation, or party among others), according to the relevance (e.g., well-known
34 travel companies can be filtered), and also according to the reputation thanks to the feed-
35 back the users can provide when the travel finishes. Moreover, the system can also make
36 recommendations to the travel agent regarding the possible customers. Therefore, this fact
37 could be classified into the relevance-based recommendation perspective.

38 2.2. Intelligent agents to distribute the workload

39 Intelligent agents are software abstractions able to establish communication channels (di-
40 rect or indirect), sharing a common environment. There, they interact with other agents
41 and with the environment extracting and providing different pieces of information. These
42 particularities allow the agents to simulate complex behaviors from the real world [18].

43 Regarding the features of the agents, they are proactive, autonomous, and independent.
44 Agents base their knowledge on predefined rules to tackle the raised problems. Their













Concept	Meaning	Icon
Agent	An active element with explicit goals that is able to initiate some actions involving other elements of the simulation.	
Role	A specific collection of tasks performed by agents when pursuing a goal or offering some service to the other members of the society. It has as a result a specific behavior.	
Environment Application	An element of the environment. Agents can act on the environment using its actions and perceive information through its events.	
Goal	An objective of a role/agent. Roles/agents try to satisfy their goals executing tasks. A goal is achieved or fails if some elements (i.e. frame facts and events) are present or absent in the agent groups or the environment.	
Task	A capability of a role/agent. To execute a task, certain elements (i.e. frame facts and events) must be available. The execution produces/consumes some elements as result.	
Frame Fact	An element produced by a task, and therefore by the roles/agents.	
Mental State	Part of the internal state of a role/agent. It groups goals, frame facts and events, and specify conditions on them.	
Belief	Part of the mental state. It contains the knowledge (specific, role-based or general) that an agent possesses. This knowledge can be used to interact with its surrounded environment.	
Conversation	Communication between two or more agents to exchange information. One of these agents plays the role of initiator of the conversation, while the others are the receptors.	
Plan	Representation of the means by which a goal can be satisfied. It is usually structured to provide a deterministic meaning to the operations.	
Mental State Manager	Part of the internal state of a role/agent. It provides for operations to create, destroy and modify mental entities.	
Mental State Processor	Part of the internal state of a role/agent. It determines how a mental state evolves, described in terms of rules or planning.	

Fig. 1. Main concepts of INGENIAS for developing MASs.

1 ability to interact can be exploited to solve complex issues or to distribute the workload
 2 with certain coordination.

3 The life cycle of intelligent agents consists of satisfying a collection of goals following
 4 the knowledge available (from a predefined set of rules or dynamically acquired from the
 5 environment or other agents), and through several associated concepts (see Fig. 1). These
 6 goals can be independent, or they can be organized hierarchically. This latter promotes
 7 the idea of having sets of sub-goals that accomplish other goals at higher levels when they
 8 are satisfied. Goals are associated with a task (or a set of them) that is executed by the
 9 agents. Both goals and tasks are elements of the mental state of agents. This mental state
 10 plays the role of a mind, being the place where the set of rules and knowledge are stored.
 11 Therefore, it supports the execution of the tasks and the satisfaction of the goals.

1 On the other hand, agents can take advantage of their ability to interact with the envi-
2 ronment to solve complex problems having partial or reduced knowledge about a problem.
3 The organization in MASs opens the collaboration, the competition, and also the nego-
4 tiation. Agent-Based Modeling (ABM) [43] and Agent-Oriented Software Engineering
5 (AOSE) [17] are standards in the domain that provide the elements and entities to address
6 this issue. Well-known approaches that use MASs to solve complex problems or simulate
7 real environments are road traffic simulations [13], distributed decision support systems
8 [20], bio-inspired ML-based systems [9], and computer games [44].

9 MASs are usually designed using specific artifacts and entities to address the develop-
10 ment phases of complex systems. It allows the production relationships and interactions
11 between agents that are graphically represented and later transformed to source code auto-
12 matically. INGENIAS, GAIA, Prometheus, and Tropos are well-known agent modelling
13 methodologies in the domain [40].

14 Agent platforms are usually the standard solution to implement a previously modelled
15 MAS. These platforms are commonly organized into source code libraries of a specific
16 programming language. There, they include features to ease the distribution of the agents
17 and manage their communication channels. Highlighted approaches in this area are JADE
18 (Java) [8] and MESA (Python)[26].

19 For the development of *Pharaoh*, the selected agent methodology has been INGE-
20 NIAS, adapting the agent model to the MESA framework for implementation purposes.
21 The agent model is a bio-inspired distribution model based on the organization of ant
22 colonies to generate a MAS distributed in several cumuli of agents (anthills) working
23 together or independently according to the workload of the system.

24 **2.3. Automatic managing of tourism**

25 Automatic tourism management is one of the basic tasks in the area due to the large
26 number of different processes to consider related to available offers, journey configura-
27 tions, marketing opportunities, and customer indications and preferences [27]. In addition,
28 travel agencies are evolving to manage tourist packages and customer preferences through
29 the Internet. This means that personal relationships and face-to-face meetings have been
30 changed for interactions with web pages that provide support and guidance.

31 The tourism management process is based on two main perspectives: the client and
32 the seller (i.e., the final destination or the travel agency). The first one is the most widely
33 addressed by the different studies and the developed systems. This issue is related to the
34 implementation of new technologies and e-tourism (i.e., marketing and offers of tourism
35 services on the Internet), which means that the tourist provider has direct contact with the
36 customer [21].

37 Regarding the approaches that automatize the different tourism-related processes,
38 most of the systems are recommendation systems for both perspectives.

39 In the case of the perspective of the travel agencies, they usually face clients who do
40 not know where to travel and what their preferences are. Therefore, systems able to extract
41 and process information from clients could provide support to the decision-making or to
42 select the optimal vendor from the company for that customer [23]. The information can
43 be textual content [28] or generated through the creation of profiles [2].

44 In the case of the perspective of customers, there are several possibilities to address.
45 For instance, systems able to elaborate specific trip configurations for clients with spe-

1 cial needs are very useful [36]. However, other approaches are more general. Examples
 2 of these general approaches are those that are focused on the selection of the best route
 3 according to the current weather or traffic conditions [41], or the recommendation of the
 4 best tourist trips for individuals or groups [22]. These latter are like virtual trip planner
 5 designers [3] that are usually based on the tastes of the clients and their former trips
 6 already completed [16]. For destination selection, similar systems have also been consid-
 7 ered. They process the opinions of the tourists to find the most popular places for them
 8 [48] and also produce rankings of interest [4]. Finally, the issue of finding a suitable res-
 9 idence to live in during the journey according to the specific features of the client has
 10 also been considered. These features are mainly the needs of the clients [14], and their
 11 economic capacity [38].

12 The *Pharaoh* framework is mainly focused on the travel agency perspective providing
 13 specific elements to interact with the clients. In the first case, the vendor obtains real-time
 14 recommendations according to the indications made by the clients and their interests.
 15 Moreover, the system presents a novel component able to inform the vendor about the
 16 clients who are more prone to buy a tourist activity. This eases the work and increases
 17 the profit of the agency. In the second case, clients can introduce their particularities and
 18 interests. The system uses these to provide filtered information to the vendor.

19 3. The Pharaoh framework

20 The main objective of the *Pharaoh* framework is to support customers and travel agents
 21 (i.e., the final users in the tourism domain). Therefore, the system is a ERP and CRM that
 22 covers both perspectives related to tourist service management.

23 From the tourist perspective, *Pharaoh* provides different tourist services according to
 24 specific preferences and features (i.e., customer searches and filters the available options).
 25 From the travel agent's perspective, the system makes available the most relevant tourist
 26 services according to the preferences provided by the customer, the previously configured
 27 ratings, and the feedback provided by former customers about previous experiences with
 28 the tourist service of interest. Moreover, *Pharaoh* includes a novel functionality for easing
 29 the selection of the optimal customers (i.e., those prone to complete the booking process
 30 of a tourist service). This functionality allows travel agents to establish priorities and
 31 different actions according to each customer.

32 The system uses several information sources to obtain the desired knowledge like
 33 available hotels, flights, cars, etc. It also includes a visualization tool where users (cus-
 34 tomers and travel agents) can interact. This tool is composed of different graphical assis-
 35 tants to support the different tasks implemented by *Pharaoh*.

36 The following sections detail the architecture of the *Pharaoh* framework considering
 37 the different modules and emphasizing the processes that cover its functionalities.

38 3.1. General architecture and modules

39 The architecture of the *Pharaoh* framework consists of five different modules: the *Informa-*
 40 *tion gatherer*, the *Tourist service manager*, the *Customer manager*, the *Travel agent*
 41 *manager* and the *Maintenance manager*. These modules are completed with two knowl-
 42 edge bases to persist the information: the *Suppliers knowledge* and the *Tourist services*

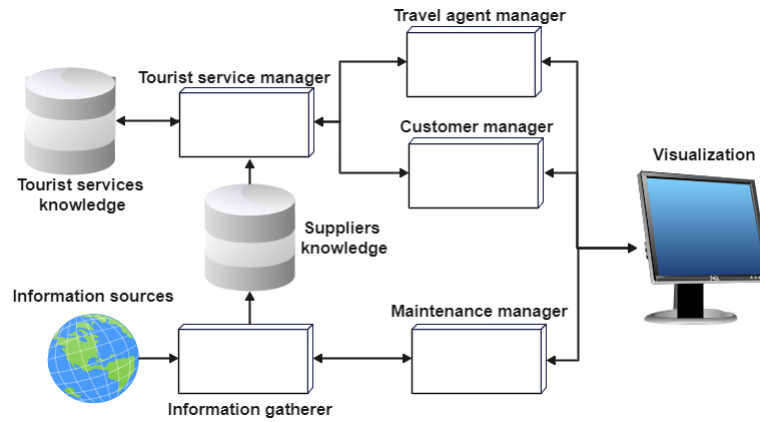


Fig. 2. Overview of the architecture of the *Pharaoh* framework.

1 *knowledge* repositories (see Fig. 2). These modules are capable of carrying out the three
 2 main processes: sale of a tourist service, customer feedback, and information update.

3 The *Maintenance manager* module serves as a central point for the maintenance of
 4 information and common parameters used by the rest of the modules of *Pharaoh*. Typical
 5 examples of maintenance of the information are related to the modifications of the tourist
 6 resource (e.g., name changes of hotels, new appearance of companies in the market, or
 7 permanent close events). It is also necessary to have the mapping of the same tourist re-
 8 source from different providers to facilitate the system to generate services with different
 9 rates. It uses the *Information gatherer* module to achieve these actions (see Fig. 3).

10 The *Information gatherer* module collects the information from the virtual tourism
 11 market. This market is based on the exchange of information coming from web sources
 12 between suppliers and consumers through XML and JSON technologies. The module re-
 13 trieves two types of information: static and dynamic. The first is not frequently modified,
 14 like the name, description, features, and relevant pictures. The second one fluctuates sev-
 15 eral times daily. Typical instances of this type of information are the occupancy level
 16 of a hotel or the rates. This information is gathered through web scraping techniques to
 17 automate the process and stored in the *Suppliers knowledge* repository. It is important to
 18 indicate that some sensible information is not stored due to contracts with suppliers, being
 19 only passively consumed by this module.

20 The *Suppliers knowledge* repository has as the main objective to manage the acquired
 21 knowledge corresponding to suppliers. The visual interface of *Pharaoh* provides specific
 22 graphical assistants to guide the user in managing this task. The system uses this stored
 23 information to produce automatic recommendations.

24 The *Tourist service manager* module is in charge of managing tourist opportunities
 25 and creating personalized offers to make recommendations (see Fig. 4). These recom-
 26 mendations are based on the calculations focused on providing the best tourism-related
 27 resources for an operation ordered by interest. Thus, it is defined the type of each tourism
 28 resource grouped by category. In the transport category, flights, collective public transport
 29 (e.g., trains or buses), and renting of private vehicles are included. In the accommoda-
 30 tion category, hotels, apartments, and similar places are considered. Finally, in the extras

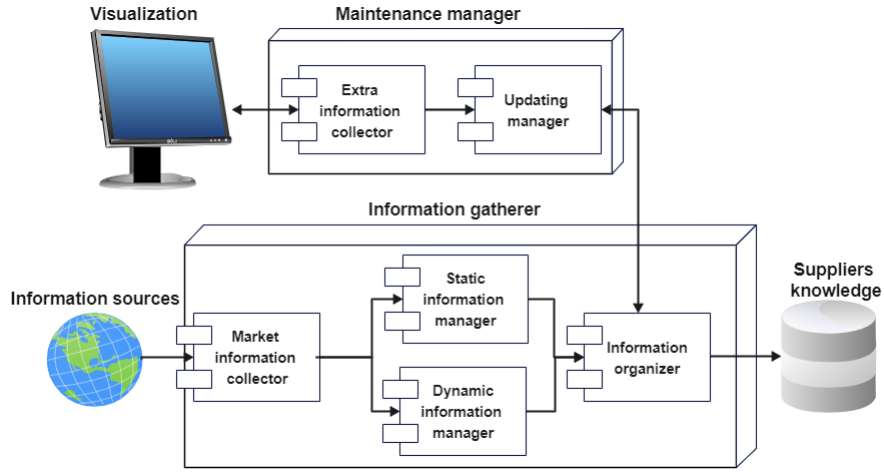


Fig. 3. Except of the *Maintenance manager* and *Information gatherer* modules.

1 category, tourist activities (e.g., tourist visits or reserved circuits in the destination) and
 2 insurance are classified here.

3 The expression that provides the best tourist resources to configure a complete tourist
 4 service (TS) is described as follows:

$$TS = \max_{k \in \{1, \dots, |T| \cdot |A| \cdot |E|\}} TS_k = \sum_{C \in \{T, A, E\}} \theta_C \cdot Score^C \quad (1)$$

5 where T is for transport, A is for accommodation, and E is for extras. In addition, the
 6 θ values are weights from 0 to 1 to indicate relevance. Note that the sum of these weights
 7 must be 1. Moreover, $Score^C$ is the average value of the values $Score_j^C$ corresponding to
 8 the selected tourist resources of the tourist service. Thus, $Score_j^C$ is defined as follows:

$$Score_j^C = w_{u,j}^C \cdot f_{u,j}^C + w_{e,j}^C \cdot f_{e,j}^C \quad (2)$$

9 where $j \in \{1, \dots, |C|\}$, where $|C|$ is the number of elements in category C . The
 10 values denoted with w are the weights of the partial contributions to the global score
 11 subsequently defined.

12 The element $f_{u,j}^C$ is about the preferred variables, and it is expressed as follows:

$$f_{u,j}^C = \frac{1}{n} \sum_i^n preferred_{j,i}^C \quad (3)$$

13 where $preferred_{j,i}^C$ is a value $\in [0, 1]$ indicating the feedback value provided by cus-
 14 tomers based on their previous experience. Thus, $f_{u,j}^C$ is the average value of the opinions
 15 of the customers.

16 In the case of the element $f_{e,j}^T$, it depends on the considered category (T, A, E):

$$f_{e,j}^T = w_{quality}^T \cdot quality_j^T + w_{stopovers}^T \cdot stopovers_j^T + w_{time}^T \cdot time_j^T + w_{price}^T \cdot price_j^T \quad (4)$$

$$f_{e,j}^A = w_{quality}^A \cdot quality_j^A + w_{location}^A \cdot location_j^A + w_{services}^A \cdot services_j^A + w_{price}^A \cdot price_j^A \quad (5)$$

$$f_{e,j}^E = w_{quality}^E \cdot quality_j^E + w_{flexibility}^E \cdot flexibility_j^E + w_{services}^E \cdot services_j^E + w_{price}^E \cdot price_j^E \quad (6)$$

1 At each category, it is considered a set of measures with values $\in [0, 1]$. With the
 2 Transport Category: *quality* means the quality of the transport, *stopovers* means the num-
 3 ber of possible stopovers, *time* means the estimated time spent to reach the destination,
 4 and *price* means the expensiveness. With the Accommodation Category: *quality* means
 5 the quality of the accommodation, *location* means the location concerning the distance
 6 to the tourist areas and the safety of the area, *services* means the level of the provided
 7 services, and *price* means the expensiveness. Finally, with the Extras Category: *quality*
 8 means the quality of the extra, *flexibility* means the flexibility in hours and hiring, *ser-*
 9 *vices* means the level of the provided services, and *price* means the expensiveness. Note
 10 that these measures are pondered by weights w whose sum is 1 for each category.

11 With the measures from each one of the categories, the system intends to answer the
 12 following key questions for a trip:

- 13 – *HOW*: by defining the quality or level of the resource.
- 14 – *WHEN*: by classifying the time needed for reaching the destination.
- 15 – *WHERE*: by classifying the place and its surroundings.
- 16 – *WHAT*: by classifying the acceptability of the extra resource in terms of what is of-
 17 fered.
- 18 – *WHY*: that defines the set of extras and safety offered by the resource.

19 The module is responsible for filtering customers according to their propensity to
 20 complete the purchase. A propensity purchase estimation algorithm is in charge of estab-
 21 lishing this classification. This algorithm uses classic scoring processing techniques and
 22 two different ML methods in a two-layer architecture to categorize the clients into five
 23 groups according to their booking probability. More details about this issue are presented
 24 later.

25 The module comprises three components: the *tourist services organizer*, the *tourist*
 26 *service filter* and the *customer feedback gatherer*. The first contains a MAS to distribute
 27 the workload and achieve the decision-making task parcelled out. The second is in charge
 28 of filtering activities according to the requests made by the customers, and the third man-
 29 ages the feedback provided by these latter.

30 Regarding the MAS, it achieves tasks related to the ML methods using a bio-inspired
 31 structure based on an anthill with a queen (one per anthill), soldiers and several workers
 32 agents. The queen agent distributes the different tasks between the soldiers and builds the
 33 final result. The soldiers organize the requests and make use of the workers to tackle the
 34 problem. This agent architecture is relevant for applying possible load-balancing politics.
 35 These politics consist of creating new soldiers and workers, or even new complete anthills
 36 to increment the capability of the system. More details about the MAS are provided in the
 37 next sections.

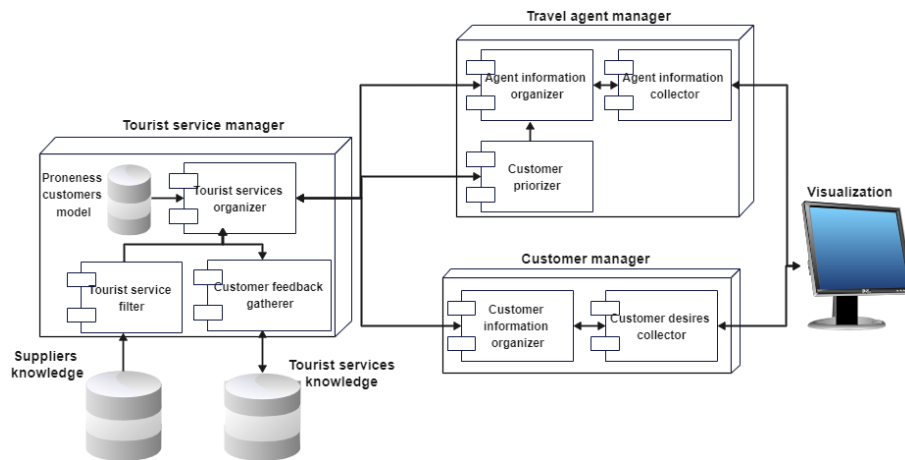


Fig. 4. Except of the *Tourist service manager*, *Travel agent manager*, and *Customer manager* modules.

1 The *Tourist services knowledge* repository stores the information related to contracted
 2 tourist services for subsequent modifications or consultations made by the travel agents.
 3 This information refers to the management decisions and configurations made by the
 4 *Tourist service manager* module. Thus, it serves as a guide in the processes of composition
 5 or self-generation of tourist services. The feedback returned from the customers after the
 6 tourist service concluded is also considered.

7 The *Customer manager* module has as the main functionality to provide support to
 8 the customer to perform the necessary operations to create a tourist activity in real-time.
 9 This module communicates with the *Tourist services knowledge* and *Suppliers knowledge*
 10 repositories through the *Tourist service manager* module to make compositions of differ-
 11 ent elements to include in the complete tourist service. The created service is stored for
 12 later verification and approval by the travel agent.

13 The *Travel agent manager* module has as a purpose to serve as the internal central
 14 management point for the operations related to the tourist service creation. It is also in
 15 charge of the review of these activities created by customers through the *Customer man-
 16 ager* module. It is directly connected to the graphical interface, being able to take in-
 17 formation about the requests made by the users. These requests are used to consult the
 18 *Tourist services knowledge* and the *Suppliers Knowledge* repositories through the *Tourist
 19 service manager* module and to obtain interesting tourist services for the user. This mod-
 20 ule can produce a tourist activity from scratch by selecting different possibilities. Note
 21 that the main difference for this module with respect to the previous one resides in the
 22 configuration of privileges of the end-user (i.e., a travel agent has more privileges than a
 23 customer).

24 Regarding the visual interface, it provides different configurations according to the
 25 two perspectives: the customer and the travel agent. For the first case, the visualization
 26 supports the creation of travel services (i.e., a complete set of tourist activities). Graphical
 27 assistants try to alleviate the complexity of the travel service design by reducing customer
 28 decision-making through a finite set of guided options. When the customer approves the

1 travel service, it is sent to the travel agency for its management: approval or modification.
 2 A series of exchanges take place between the customer and the travel agency until the
 3 customer confirms the tourist service firmly and in turn is approved. Once the services
 4 have been confirmed and the full payment has been verified, a process of sending the nec-
 5 essary documentation is executed for the client to enjoy the service. In addition, a user
 6 area is made available to the customer to facilitate the consultation of the travel informa-
 7 tion and its associated legal documentation. For the second case, similar functionalities
 8 are provided. However, travel agents have the initiative here, generating multiple travel
 9 services for the customers. Note that more details about the sale processes are addressed
 10 in the next sections.

11 3.2. Algorithm for the detection of potential customers

12 The *Tourist service manager* module contains a propensity purchase estimation algorithm.
 13 It classifies the potential customers according to their propensity to commit the book of a
 14 tourist service. This algorithm provides valuable information to the travel agency and en-
 15 ables the travel agent to differentiate between good and bad clients (customers who book a
 16 tourist service or not). The algorithm consists of three key elements: variable transforma-
 17 tion, clustering, and classification. The first element prepares the inputs for the probability
 18 estimation, conducted by clustering and classification methods. The second element pro-
 19 vides an initial purchase probability, and the last element refines the probability in some of
 20 the clusters. The algorithm's output categorizes the probability into five groups according
 21 to the conversion rate (i.e., the proportion of clients who finally book a tourist service).

22 The variable transformation element is carried out by Weight of Evidence (WoE) en-
 23 coding [45]. This type of transformation and its benefits have been studied within the
 24 scope of credit scoring deeply [5, 47]. It helps deal with categorical and numerical data
 25 simultaneously; besides, it can lead to missing data with any additional technique. Tra-
 26 ditionally, WoE encoding is used to separate bad clients from good clients, which fits
 27 perfectly with an algorithm of propensity purchase estimation. The equation of the WoE
 28 transformation is:

$$WoE = \ln \frac{\%non_events}{\%events} \quad (7)$$

29 where *events* value refers to good clients, and *non-events* value represents bad clients.

30 One of the most powerful advantages of WoE for the current scope is that variables
 31 with too many discrete values lead to numeric variables, and they will be ordered ac-
 32 cording to a monotonic relationship to the dependent variable. WoE provides a numeric
 33 variable that regroups huge discrete values into densely populated categories that express
 34 information for the new category. The other great advantage that benefits the proposed
 35 solution is that WoE is a standardized value. The former allows using Decision Tree (DT)
 36 algorithms. Note that a tree-based algorithm can reduce its predictive potential with cat-
 37 egorical variables with too many categories [15]. The latter allows using distance-based
 38 algorithms like *k-means* clustering algorithm, which needs variables with comparable val-
 39 ues because it employs Euclidean distance [33].

40 The algorithm is a bi-stage model as depicted in Figure 5, that first applies *k-means*
 41 clustering to achieve three clusters with different conversion rates (low, regular, high).
 42 Secondly, for more accurate results, a DT classifier acts over the low and high clusters

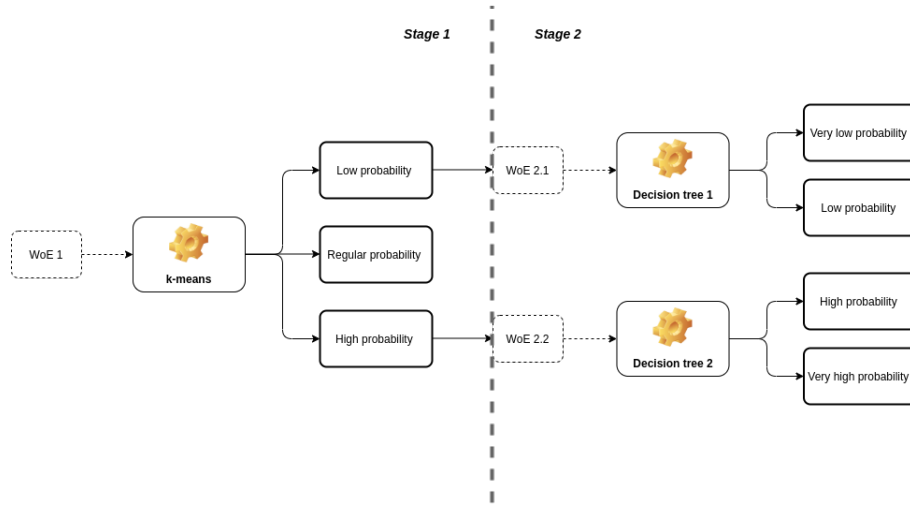


Fig. 5. Bi-stage model for the probability estimation of propensity to commit the book of a tourist service.

1 by subdividing them into low and very low, and high and very high respectively. Hence,
 2 the proposed algorithm eventually provides five output categories: very low, low, regular,
 3 high, and very high. This type of output is calculated on the base of the prior probability
 4 of propensity to purchase [37], and it offers an interpretable result for a travel agent.

5 The key to the model is how to group the variables in the two stages. In the first stage,
 6 WoE transformation is applied to the set of variables that contains information about how
 7 the customer is and where they come from. Examples of these kinds of variables are:
 8 means of contact (e.g., web, email, physical presence in the agency or phone), getting
 9 information for short term, or large term, the travel is related to a particular event (e.g.,
 10 wedding or anniversary). With these variables, clustering is developed to first obtain three
 11 groups (low, regular, and high), which later will be split into the final desired five groups.
 12 Then, in the second stage, WoE transformation is applied to variables inherent in travel in-
 13 formation like the number of places to be visited, trip duration, flexibility, final price, etc.
 14 Note that WoE encoding must be applied separately to low and high-probability clusters,
 15 even concerning the same variable, because the goal is to capture the patterns related to
 16 each independent cluster. Finally, two DT applied into low and high probability clusters
 17 subdivide those clusters into two groups each, resulting in five categories with different
 18 propensity purchases.

19 **3.3. Bio-inspired MAS organization**

20 Intelligent agents are independent entities that are used by *Pharaoh* to tackle the ML
 21 methods. They establish cooperative activities organized through the hierarchical archi-
 22 tecture of a MAS.

23 The proposed MAS consists of an anthill that comprehends agents playing three types
 24 of ant-roles: queen, soldiers, and workers (see Fig. 6). Therefore, it is organized following

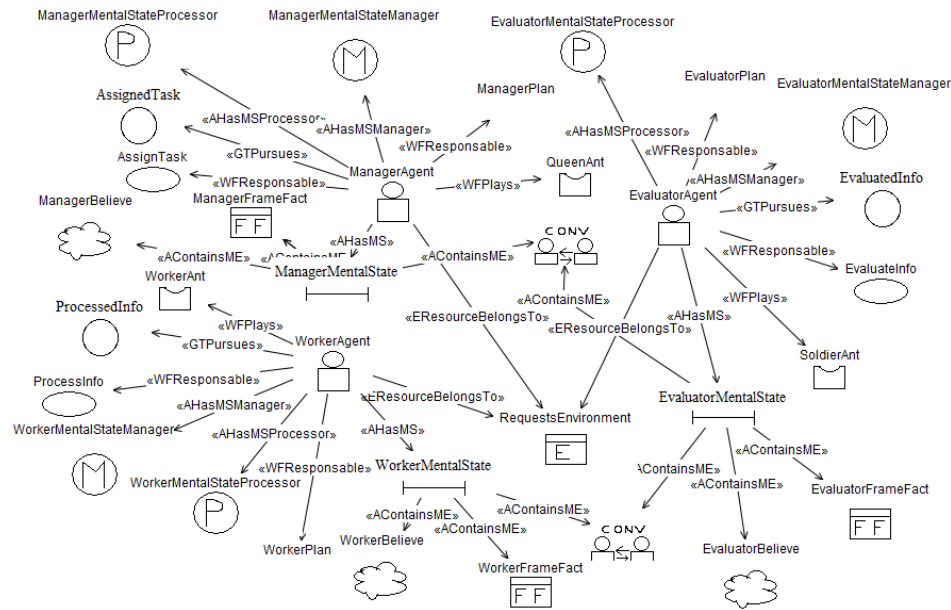


Fig. 6. Excerpt of the main entities involved in the anthill MAS.

1 a bio-inspired ant social structure. The queen (i.e., the manager agent) is responsible for
 2 the anthill, receiving the requests to analyze. Notice that several requests can be sent to this
 3 agent. These requests come from the customers' preferences for a specific tourist service.
 4 Then, the queen agent assigns the activity of evaluating the elements of the request to one
 5 soldier agent (i.e., the evaluator agent). This agent is in charge of assigning the relevant
 6 information to the worker agents. Thus, this information is decomposed into more simple
 7 requests organized by the different features indicated by the customer. Worker agents
 8 process the information and apply the ML methods provided by the system. Notice that
 9 a copy of the ML model is available for each worker and they have access to the relevant
 10 information provided by the suppliers to configure the possible tourist packages that are
 11 offered to the customer. Once the worker agents have concluded their task, the soldiers
 12 join the result if necessary (protecting and supervising the result obtained by a worker or
 13 a set of them) and they return the result to the queen agent.

14 This process is based on the Belief-Desire-Intention (BDI) model [11] where each
 15 agent presents a goal or a set of them that must be satisfied to complete its life cycle. In
 16 this sense, the queen has the *assignedtask* goal, the soldiers have the *evaluatedinfo* goal
 17 and the workers have the *processedinfo* goal. All the goals usually include associated tasks
 18 that are actions to achieve. Agents have at least one task that solves their corresponding
 19 goals. These tasks are applied in the shared environment. In this case, the environment
 20 is formed by the requests of the possible customers. Each agent incorporates a mental
 21 state, its processor and manager, and a set of beliefs (motivations). The queen has rules
 22 and a plan to manage the requests, while soldiers present similar rules to distribute the
 23 information about the requests between the workers. However, workers include the ML
 24 model in their mental states to achieve the evaluation of the information, a plan, and

1 some simple rules to organize the process in the beliefs. Finally, interactions between the
 2 individuals follow the hierarchical structure which mainly consists of direct conversations.
 3 Notice that in this case, workers do not need to establish conversations with other workers
 4 since they tackle their commitment individually according to the orders of the soldier.

5 Regarding the design of the anthill model, it has been addressed through the INGE-
 6 NIAS agent methodology. Then, the resulting composition has been transformed to be
 7 compliant with the MESA framework. The conversations and interactions of the agents
 8 have been implemented following the Foundation for Intelligent Physical Agents (FIPA)
 9 standards [31].

10 **3.4. Internal system processes**

11 The *Pharaoh* framework presents three main processes according to its architecture. The
 12 first process encompasses the different actions achieved by the travel agent and the cus-
 13 tomer until the sale of a tourist service is completed. The second process refers to the
 14 feedback task. Finally, the third process comprehends the steps to update the information
 15 about rates, services, and other relevant information.

16 The sale of a tourist service process consists of a set of steps most of them common
 17 to both types of end-users (i.e., travel agent and customer). It occurs due to both roles
 18 interacting several times during the process (see Fig. 7). Initially, the end-users are logged
 19 into the system. If the users are customers, they can request a tourist service and ask the
 20 travel agents about it. However, if the users are travel agents, they can take the initiative
 21 (the customers are physical with them) or they answer the possible requests made by
 22 customers previously. Then, the system analyzes the customers to decide if they are prone
 23 to sale or not. In case they are prone, the travel agents prioritize their demands. If the
 24 customers respond to the answer provided by the travel agents and one of the tourist
 25 services is interesting to them, the system finishes the process of sending the documents
 26 with the transaction and the tourist service-related information. However, the customers
 27 could not respond to the tourist services stipulated by the travel agents. In this case, the
 28 process concluded. Moreover, the customers could not find interesting the offers made by
 29 the travel agents. In this case, the travel agents should produce new tourist services more
 30 adjusted to the desires of the customers.

31 The feedback process follows a set of steps where the customers introduce feedback
 32 about a tourist service previously bought and concluded (see Fig. 8). It is a simple and
 33 optional process carried out only by the customers where their opinion is stored in the
 34 system. The information provided is used to update the formulas for the recommendations
 35 that are considered by *Pharaoh* to select the best options.

36 The updating information process is the typical maintenance work in the system. It
 37 mainly modifies the stored knowledge in the two repositories: the *Suppliers knowledge*
 38 and the *Tourist services knowledge*. This task must be daily performed to maintain the
 39 *Pharaoh* framework up to date. This process is achieved by a complete Extract, trans-
 40 form and load (ETL) architecture that automatizes the steps to manage the actions and
 41 simplifies the task for the administrator. Note that the administrator is not considered an
 42 end-user by the system (see Fig. 7).

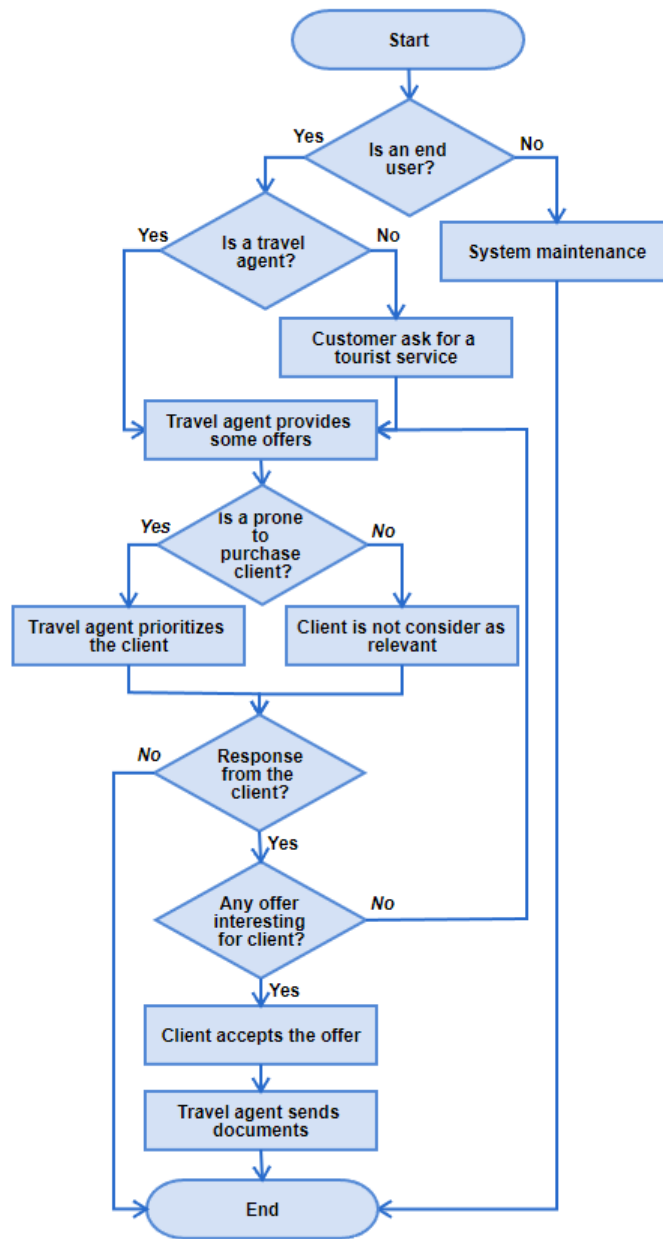


Fig. 7. Offer and sale processes of a tourist service provisioning.

1 4. Experiments

2 This section details the experiments carried out to validate the performance of the *Pharaoh*
 3 framework. Three experiments have been considered and developed in a real environment.

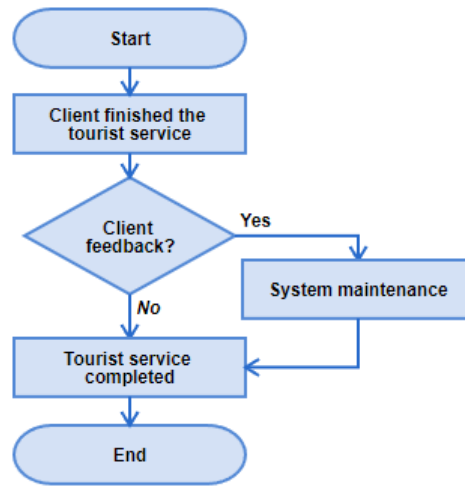


Fig. 8. Information updating process based on the feedback of the customer.

1 *JOF ASSOCIATES SLU* also called *Madox Viajes S.L.*¹ was the selected travel agency
 2 used to develop the experiments. Notice that this is the first company where the system
 3 has been implanted in production once the development stage has been completed.

4 Regarding the MAS included in the system, it presents a basic configuration adapted
 5 to the needs of the company. These needs have been previously tested by adjusting the
 6 parameters before the presented experiments. Thus, the MAS comprises two anthills, with
 7 five soldiers and ten workers.

8 The first experiment covers the parameters configuration and the setup of the sys-
 9 tem. Moreover, a basic test is presented to evaluate the satisfaction level produced by the
 10 system between travel agents and customers according to the recommendations made.
 11 The second experiment details the performance of the detection of potential customers
 12 algorithm. A comparison with other methods of the state-of-the-art is included. The third
 13 experiment evaluates the performance of the system once it is implanted in the travel
 14 agency. Comparisons between the profits of the travel agency using and not using the
 15 *Pharaoh* framework are addressed. This fact illustrates the relevance of the system during
 16 the management of a travel agency.

17 **4.1. Parameters configuration and initial validation of the system**

18 The first experiment consists of a set of steps to show how the proposed system works. Ini-
 19 tially, the parameters of the *Pharaoh* framework were configured to produce recommen-
 20 dations about tourist services to customers and travel agents. These values were provided
 21 by 3 experts in the domain from the selected travel agency (i.e., the people responsible
 22 for the company). Then, 10 customers (preserving their anonymity) and 3 travel agents
 23 of the company (i.e., some of the workers of the company) evaluated the obtained recom-
 24 mendations for their specific tourism desires. Finally, these actors provided feedback

¹ <https://www.madoxviajes.com/>

Table 1. Satisfaction values provided by the travel agents.

Role	Satis VAS
<i>Agent A</i>	0.85
<i>Agent B</i>	0.93
<i>Agent C</i>	0.73
Mean \pm S.D.	0.84 \pm 0.10

Table 2. Satisfaction values provided by the customers.

Role	Satis VAS
<i>Customer 1</i>	0.71
<i>Customer 2</i>	0.62
<i>Customer 3</i>	0.89
<i>Customer 4</i>	0.76
<i>Customer 5</i>	0.95
<i>Customer 6</i>	0.58
<i>Customer 7</i>	0.78
<i>Customer 8</i>	0.69
<i>Customer 9</i>	0.92
<i>Customer 10</i>	0.76
Mean \pm S.D.	0.77 \pm 0.12

1 about the performance of the system and their satisfaction with the obtained results. This
 2 satisfaction has been measured according to a Satisfaction Visual Analogue Scale (Satis
 3 VAS) [29]. This VAS was included in an anonymous document. The result of the VAS
 4 was converted to a range of values from 0 to 1 to obtain the resulting opinion.

5 Regarding the parameters configuration (see Section 3.1), the travel agency experts
 6 decided the next values for them. For θ_T , θ_A , and θ_E , the values were fixed to 0.3, 0.45
 7 and 0.25 respectively. This decision was motivated by the assumption that customers give
 8 more importance to the hosting and the location than the rest of the features of the tourist
 9 service. Then, $w_{u,j}^C$ and $w_{e,j}^C$ were fixed to 0.4 and 0.6, giving a moderate relevance to the
 10 feedback provided by previous customers. Lastly, the weights of the elements of each cat-
 11 egory (transport, accommodation, and extras) are established. For the transport category,
 12 $w_{quality}^T$, $w_{stopovers}^T$, w_{time}^T , and w_{price}^T are configured to 0.3, 0.1, 0.2, and 0.4 respec-
 13 tively. This decision of the experts was motivated by the assumption that users usually
 14 consider very relevant the price of the transport over the other features. For the accom-
 15 modation category, $w_{quality}^A$, $w_{location}^A$, $w_{services}^A$, and w_{price}^A are configured to 0.35, 0.25,
 16 0.15, and 0.25 respectively. These values correspond to the assumption that users usually
 17 consider as relevant the price and the quality of the hostage unless the location and the
 18 provided services are also very important. For the extras category, $w_{quality}^E$, $w_{flexibility}^E$,
 19 $w_{services}^E$, and w_{price}^E are configured to 0.45, 0.15, 0.15, and 0.25 respectively. This de-
 20 cision follows the assumption that the users give more importance to the price and the
 21 quality of the extras.

22 Once the parameters of the system are completely fixed, the three travel agents make
 23 petitions to *Pharaoh* to obtain recommendations according to their preferences. These

1 preferences were selected as heterogeneous as possible to validate the system through a
2 wide spectrum of possibilities.

3 Next, the customers provided indications of the system being assisted by the travel
4 agents to solve possible problems related to the operation of the system. Travel agents
5 also provided users with the best configurations obtained by the system. If a customer
6 agrees with one of the offered tourist services, the system completes the corresponding
7 booking process.

8 Finally, the anonymous document with the VAS was given to the participants. The
9 results are shown in Table 1 and Table 2. The average satisfaction of the travel agents was
10 0.84, while the average of customers was 0.77. Similar Standard Deviations were obtained
11 in both cases. Thus, the global average satisfaction of the system was 0.80 ± 0.11 .

12 In conclusion, it can be said that the result is very acceptable from both perspectives
13 (customers and travel agents). Furthermore, customers are more pressing and demanding
14 than the travel agents in the evaluation of the satisfaction produced by the recommenda-
15 tions made by the *Pharaoh* framework.

16 **4.2. Propensity purchase estimation algorithm validation.**

17 This experiment addresses the performance of the proposed propensity purchase estima-
18 tion model (*bi-stage* model). The proposal is compared with the most typical state-of-
19 the-art ML classifiers [39, 19]: Random Forest (RF), Gradient Boosting (GB), k-Nearest
20 Neighborhood (kNN) and Logistic Regression (LR).

21 *JOF ASSOCIATES SLU* has provided the dataset for the development of this exper-
22 iment. It contains about 10^4 instances presenting an initial conversion rate of 14%, that
23 is, the percentage of customers who finally purchased the tourist service (i.e., the data
24 are unbalanced). Each instance includes a variety of attributes: contact details, contact
25 channel, service, travel destination, type of trip, etc. In this particular case, more pre-
26 cise information of the data is not available because of confidentiality. All the selected
27 algorithms are trained with 80% of the whole data and tested using the remaining 20%.
28 This division is carried out with stratified sampling. The hyper-parameters of the ML
29 alternatives are selected by 10-fold-cross-validation. The output of each algorithm is pre-
30 sented into five categories attending to the output probability distribution on the train set.
31 The performance is measured in terms of conversion rate and relative frequency on each
32 classification group, expressed in percentage of the overall total. The predictive power is
33 evaluated by checking the conversion rate in the different groups, and the frequency pro-
34 vides the significance of the group. Finally, to assess the possible overfitting train and test
35 agreement are evaluated.

36 The results are shown in Table 3. The RF has a poor agreement between train and
37 test, in both frequencies and conversion rates. Indeed, one of the groups has no represen-
38 tation in the test dataset. The kNN has similar behavior in train and test agreement, with
39 differences up to 21%. This situation makes no sense to evaluate any other performance
40 measure. The GB has a generally good agreement between train and test instances, except
41 in conversion rates for *high* and *very high* categories with differences higher than 20%.
42 These differences make, like in RF and kNN, no sense to evaluate the rest of the evalua-
43 tion items. The LR has the best agreement between train and test, in both frequency and
44 conversion rates. The data distribution is highly unbalanced throughout the categories,

Table 3. Performance of the ML classification algorithms in terms of relative frequency (freq.), expressed in percentage about the overall total, and conversion rate (c.rate), for each category.

		RF		GB		kNN		LR		Bi-stage model	
		freq.	c.rate	freq.	c.rate	freq.	c.rate	freq.	c.rate	freq.	c.rate
Train	very low	75.0	0.0	74.2	5.8	0.0	100.0	72.5	11.5	16.6	1.7
	low	9.9	0.0	10.3	25.2	54.2	0.0	9.0	17.8	17.1	5.1
	regular	9.9	89.1	10.3	34.5	29.8	20.6	13.2	19.6	44.7	10.9
	high	3.9	100.0	4.1	63.4	11.9	42.2	4.2	26.7	15.8	32.8
	very high	1.2	100.0	1.0	87.3	4.1	100.0	1.1	28.2	5.9	48.0
Test	very low	57.5	4.9	73.6	7.0	0.0	-	72.3	6.7	17.0	2.5
	low	17.1	16.5	10.0	25.0	75.2	7.5	2.0	25.0	17.0	6.7
	regular	25.2	30.9	9.8	28.8	17.4	26.7	5.8	22.9	46.1	10.2
	high	0.0	-	4.8	41.4	5.3	40.6	7.3	34.1	14.6	34.2
	very high	0.2	100.0	1.8	63.6	2.0	58.3	12.6	35.5	5.4	51.5

1 where *low probability* group comprises more than 70% of the data. Besides, the five cate-
2 gories do not have enough distance between them in the conversion rate term, 2% in some
3 cases, making it difficult to differentiate between close groups. Eventually, the proposed
4 *bi-stage* model has an excellent agreement between train and test instances, with a max-
5 imum deviation of 3%. The data distribution is the most balanced of the models tested,
6 where the *regular* category is the densest with 45% of the data and the *very high* the lower
7 dense with 6% of the data for train, and 46% and 5% for test respectively. The probabil-
8 ity categories obtained are different with five representative probability values, and it is
9 validated in the test dataset with minimum deviations.

10 In conclusion, the experiment carried out shows that this particular problem needs an
11 ad-hoc approach, such as the *bi-stage* model presented. This solution is adequate to solve
12 the propensity to purchase problem provided by *JOF ASSOCIATES SLU*. The *bi-stage*
13 model is the one that best classifies in the different probability categories with balanced
14 frequencies. In addition, it is the one that provides the best concordance between training
15 and testing, thus guaranteeing the non-existence of overfitting. Hence, according to the
16 performance requirements, the proposed method serves better than the state-of-the-art
17 classifier for the current aim.

18 4.3. Efficiency evaluation of the complete system

19 The last experiment checks the efficiency of *Pharaoh* in a real situation. The setup of the
20 system with the parameters configuration has been maintained, being the same ones as
21 the first experiment.

22 The experiment starts by organizing randomly and proportionally 184 potential cus-
23 tomers into two groups: *Traditional* and *Innovators*, for making comparisons between
24 them. All of them are looking for the same type of tourist services: *National Coasts*, *Na-*
25 *tional Islands*, *Dubai-Maldives*, and *Safaris*. In addition, web canal and other canals such
26 as phone calls, family recommendations, in-person visits to the office, etc., are considered.
27 Note that these services comprehend a set of specific tourist services.

Table 4. Summary of the obtained results for *Traditional* and *Innovators* groups for each tourist service. n is the number of potential customers in each service and group. Time is the average time (in hours) each travel agent spends processing the tourist service until it is accepted or dismissed. The conversion rate (c.rate) and the number of recommendations (Rec.) are presented.

		Traditional				Innovators			
	Services	n	Time	c.rate	Rec.	n	Time	c.rate	Rec.
	N. Coasts	20	3.2	15.0	2	20	2.9	25.0	3
Web	N. Islands	20	5.2	10.0	2	20	4.8	15.0	3
	Dubai-Maldives	16	10.1	6.3	1	16	9.6	12.5	2
	Safaries	12	11.2	16.7	1	12	10.7	16.7	2
	N. Coasts	8	4.2	37.5	1	8	2.7	50.0	2
Other	N. Islands	8	5.7	25.0	0	8	3.0	37.5	1
	Dubai-Maldives	4	12.0	25.0	1	4	8.0	50.0	1
	Safaries	4	12.5	25.0	0	4	9.0	25.0	1

1 *Traditional* group consists of 92 customers who were attended maintaining the tra-
2 ditional way that the company has been used (i.e., without the support of the *Pharaoh*
3 framework). On the other hand, *Innovators* group consists of 92 customers who attended
4 using the *Pharaoh* framework to carry out the requests made by the users and the config-
5 urations provided by the travel agents. Both groups worked with the same travel agents to
6 minimize possible variations in the results.

7 The experiment considers three main variables: the average time (in hours) each agent
8 spends processing the tourist service offered to customers, the conversion rate, and the
9 number of recommendations. These results are presented in Table 4 for both groups.

10 For *Traditional* group, the results show that the average time to manage the services is
11 less in national destinations. That is, the customers are more demanding when the travel
12 destination is in the same country. The conversion rate fluctuates in the web canal, hav-
13 ing a low value for the service *Dubai-Maldives* (i.e., only a few customers complete the
14 purchase), while *Safaris* and *National Coasts* have a high percentage. A total of 8 recom-
15 mendations have been obtained for this group.

16 For *Innovators* group, the service *Dubai-Maldives* has the lowest conversion rate and
17 *National Coasts* has the highest one in the web canal. Relevance differences in the num-
18 ber of inverted hours, the number of recommendations, and the conversion rate between
19 *Traditional* and *Innovators* have been detected. The system incremented acceptably the
20 conversion rate, the number of hours was reduced significantly, and the number of recom-
21 mendations was also higher. A total of 15 recommendations have been obtained for the
22 *Innovators* group. The changes in the conversion rate and the number of hours spent are
23 motivated by the propensity purchase estimation algorithm. It allowed the selection of the
24 best customers for the travel agents, discarding the customers with less interest in the sale.
25 Moreover, the visual information provided by the system, and the precise recommenda-
26 tions could have also done that customers more easily accepted the tourist service.

27 In conclusion, *Pharaoh* has demonstrated that it works perfectly with the daily issues
28 of a travel company. Moreover, it could increase the conversion rate that is translated into

1 profits for the company, and reduce the number of hours until the customer accepts or dis-
 2 miss the different possibilities offered for a specific tourist service. This fact is translated
 3 into less stressful situations for employees and more agility in the booking process.

4 5. Conclusions

5 The *Pharaoh* framework, a system devoted to completely covering the management of the
 6 sales of tourist services from the side of the travel agent and the customer, has been pre-
 7 sented. It includes as a novelty a bio-inspired MAS based on the structure of anthills and
 8 a proneness customer model based on ML techniques. This latter can classify customers
 9 according to their proneness of completing a booking (i.e., it detects the most interest-
 10 ing customers). The model consists of three elements: variable transformation, clustering,
 11 and classification. The output categorizes the probability into five groups based on the
 12 conversion rate.

13 The system has been tested empirically through several experiments. They have illus-
 14 trated that *Pharaoh* enhances the performance of a travel agency, increments the quality
 15 of the service, and generates interesting profits. Thus, it can be said that the framework
 16 provides very desirable functionalities and new insights in the tourism domain thanks to
 17 the proposed bio-inspired MAS and the ML techniques.

18 Future research will focus on the communication between the customer and the travel
 19 agent, trying to improve it by simplifying the steps during the conversation. The introduc-
 20 tion of improvements at the graphical level both during the sale and feedback processes
 21 will also be considered. Moreover, in the domain of intelligent agents, it could be very
 22 interesting to include a complete MAS. It could cover all the functionalities of the system
 23 to distribute and emulate the competencies of travel agents. This will lead to develop-
 24 ing a completely independent intelligent framework that could interact with customers
 25 becoming a virtual travel agent.

26 Acknowledgments

27 This work has been partially supported by the Spanish MICINN under the XMIDAS
 28 project (PID2021-122640OB-I00) and the donation of the Titan V GPU by NVIDIA.
 29 Special thanks to JOF ASSOCIATES INT S.L.U and Madoxviajes S.L.

30 References

- 31 1. Adomavicius, G., Tuzhilin, A.: Toward the next generation of recommender systems: A sur-
 32 vey of the state-of-the-art and possible extensions. *IEEE transactions on knowledge and data*
 33 *engineering* 17(6), 734–749 (2005)
- 34 2. Adriana, S.M.: The customer orientation principle - a prerequisite for increased efficiency in
 35 tourism. *Ovidius University Annals, Economic Sciences Series* 18(1), 405–408 (2018)
- 36 3. Al Fararni, K., Nafis, F., Aghoutane, B., Yahyaouy, A., Riffi, J., Sabri, A.: Hybrid recommender
 37 system for tourism based on big data and ai: A conceptual framework. *Big Data Mining and*
 38 *Analytics* 4(1), 47–55 (2021)
- 39 4. Alrasheed, H., Alzeer, A., Alhowimel, A., Althyabi, A., et al.: A multi-level tourism destination
 40 recommender system. *Procedia Computer Science* 170, 333–340 (2020)

- 1 5. Anderson, R.: The credit scoring toolkit: theory and practice for retail credit risk management
2 and decision automation. Oxford University Press (2007)
- 3 6. Bakar, N.A., Rosbi, S.: Effect of coronavirus disease (covid-19) to tourism industry. *International Journal of Advanced Engineering Research and Science* 7(4), 189–193 (2020)
- 4 7. Bedi, P., Agarwal, S.K., Jindal, V., Richa: Marst: Multi-agent recommender system for e-
5 tourism using reputation based collaborative filtering. In: *Databases in Networked Information*
6 *Systems: 9th International Workshop, DNIS 2014, Aizu-Wakamatsu, Japan, March 24–26,*
7 *2014. Proceedings 9.* pp. 189–201. Springer (2014)
- 8 8. Bergenti, F., Caire, G., Monica, S., Poggi, A.: The first twenty years of agent-based software
9 development with jade. *Autonomous Agents and Multi-Agent Systems* 34(2), 1–19 (2020)
- 10 9. Cabezas, J., Fernandez-Isabel, A., Fernández, R.R., González-Fernández, C., Alonso, A.,
11 de Diego, I.M.: Bio-inspired agent-based architecture for fraud detection. In: *Proceedings of*
12 *the 2020 3rd International Conference on Information Management and Management Science.*
13 *pp. 67–71 (2020)*
- 14 10. Camilleri, M.A.: The tourism industry: An overview. *Travel marketing, tourism economics and*
15 *the airline product* pp. 3–27 (2018)
- 16 11. De Silva, L., Meneguzzi, F.R., Logan, B.: Bdi agent architectures: a survey. In: *Proceedings*
17 *of the 29th International Joint Conference on Artificial Intelligence (IJCAI), 2020, Japão.* pp.
18 4914–4921 (2020)
- 19 12. Dinu, A.M., et al.: The importance of transportation to tourism development. *Academic Journal*
20 *of Economic Studies* 4(4), 183–187 (2018)
- 21 13. Fernández-Isabel, A., Fuentes-Fernández, R., de Diego, I.M.: Modeling multi-agent systems
22 to simulate sensor-based smart roads. *Simulation Modelling Practice and Theory* 99, 101994
23 (2020)
- 24 14. Forouzandeh, S., Berahmand, K., Nasiri, E., Rostami, M.: A hotel recommender system for
25 tourists using the artificial bee colony algorithm and fuzzy topsis model: A case study of tri-
26 padvisor. *International Journal of Information Technology & Decision Making* 20(01), 399–
27 429 (2021)
- 28 15. Friedman, J.H.: The elements of statistical learning: Data mining, inference, and prediction.
29 *springer open* (2017)
- 30 16. García, I., Sebastia, L., Onaindia, E., Guzman, C.: A group recommender system for tourist
31 activities. In: *International conference on electronic commerce and web technologies.* pp. 26–
32 37. Springer (2009)
- 33 17. Garro, A., Mühlhäuser, M., Tundis, A., Baldoni, M., Baroglio, C., Bergenti, F., Torroni, P.,
34 et al.: Intelligent agents: Multi-agent systems. *Encyclopedia of Bioinformatics and Computa-*
35 *tional Biology: ABC of Bioinformatics* 315 (2018)
- 36 18. Garro, A., Mühlhäuser, M., Tundis, A., Mariani, S., Omicini, A., Vizzari, G.: Intelligent agents
37 and environment. *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioin-*
38 *formatics* p. 309 (2018)
- 39 19. González, S., García, S., Del Ser, J., Rokach, L., Herrera, F.: A practical tutorial on bagging
40 and boosting based ensembles for machine learning: Algorithms, software tools, performance
41 study, practical perspectives and opportunities. *Information Fusion* 64, 205–237 (2020)
- 42 20. González-Fernández, C., Cabezas, J., Fernández-Isabel, A., Martín de Diego, I.: Combining
43 multi-agent systems and subjective logic to develop decision support systems. In: *International*
44 *Conference on Information Processing and Management of Uncertainty in Knowledge-Based*
45 *Systems.* pp. 143–157. Springer (2020)
- 46 21. Hamid, R.A., Albahri, A.S., Alwan, J.K., Al-qaysi, Z., Albahri, O.S., Zaidan, A., Alnoor, A.,
47 Alamoodi, A., Zaidan, B.: How smart is e-tourism? a systematic review of smart tourism
48 recommendation system applying data management. *Computer Science Review* 39, 100337
49 (2021)
- 50

- 1 22. Herzog, D., Laß, C., Wörndl, W.: Tourrec: a tourist trip recommender system for individuals
2 and groups. In: Proceedings of the 12th ACM Conference on Recommender Systems. pp. 496–
3 497 (2018)
- 4 23. Hsu, P.F., Lan, K.Y., Tsai, C.W.: Selecting the optimal vendor of customer relationship man-
5 agement system for medical tourism industry using delphi and ahp. *International Journal of*
6 *Enterprise Information Systems (IJEIS)* 9(1), 62–75 (2013)
- 7 24. Intayoad, W., Kamyod, C., Temdee, P.: Reinforcement learning for online learning recommen-
8 dation system. In: 2018 Global Wireless Summit (GWS). pp. 167–170. IEEE (2018)
- 9 25. Kang, E.y., Kim, H., Cho, J.: Personalization method for tourist point of interest (poi) rec-
10 ommendation. In: Knowledge-Based Intelligent Information and Engineering Systems: 10th
11 International Conference, KES 2006, Bournemouth, UK, October 9-11, 2006. Proceedings,
12 Part I 10. pp. 392–400. Springer (2006)
- 13 26. Kazil, J., Masad, D., Crooks, A.: Utilizing python for agent-based modeling: the mesa frame-
14 work. In: International Conference on Social Computing, Behavioral-Cultural Modeling and
15 Prediction and Behavior Representation in Modeling and Simulation. pp. 308–317. Springer
16 (2020)
- 17 27. Kler, B.K.: Tourism and restoration. In: Philosophical issues in tourism, pp. 117–134. Channel
18 View Publications (2009)
- 19 28. Loh, S., Lorenzi, F., Saldaña, R., Liethnow, D.: A tourism recommender system based on col-
20 laboration and text analysis. *Information Technology & Tourism* 6(3), 157–165 (2003)
- 21 29. Lynskey, S.J., Ling, F., Greenberg, A.M., Penny-Dimri, J.C., Sutherland, A.G.: The influence
22 of patient resilience and health status on satisfaction after total hip and knee arthroplasty. *The*
23 *Surgeon* 19(1), 8–14 (2021)
- 24 30. Meehan, K., Lunney, T., Curran, K., McCaughey, A.: Context-aware intelligent recommenda-
25 tion system for tourism. In: 2013 IEEE international conference on pervasive computing and
26 communications workshops (PERCOM workshops). pp. 328–331. IEEE (2013)
- 27 31. Melo, L.S., Sampaio, R.F., Leão, R.P.S., Barroso, G.C., Bezerra, J.R.: Python-based multi-
28 agent platform for application on power grids. *International transactions on electrical energy*
29 *systems* 29(6), e12012 (2019)
- 30 32. Moguerza, J.M., Martín-Duque, C., Fernández-Muñoz, J.J.: The importance of service quality
31 as an instrument for client customization: a methodological and practical approach within the
32 hotel sector. *Quality & Quantity* pp. 1–12 (2021)
- 33 33. Mohamad, I.B., Usman, D.: Standardization and its effects on k-means clustering algorithm.
34 *Research Journal of Applied Sciences, Engineering and Technology* 6(17), 3299–3303 (2013)
- 35 34. Moliner, M.A., Sánchez, J., Rodríguez, R.M., Callarisa, L.: Relationship quality with a travel
36 agency: The influence of the postpurchase perceived value of a tourism package. *Tourism and*
37 *Hospitality Research* 7(3-4), 194–211 (2007)
- 38 35. Moreno, R., Viajes, M., Fernández-Isabel, A., de Diego, I.M., Moguerza, J.M., Lancho, C.,
39 Cuesta, M.: Automatic detection of potential customers by opinion mining and intelligent
40 agents. In: 2022 17th Conference on Computer Science and Intelligence Systems (FedCSIS).
41 pp. 93–101. IEEE (2022)
- 42 36. Novikova, N.G., Kortunov, V.V., Krasnova, O.N., Zorina, N.M., Kireenkova, Z.A., et al.:
43 A client with special needs in tourist services. *Journal of Environmental Management and*
44 *Tourism* 6(2), 338 (2015)
- 45 37. Pons, J., Granados, A., Espinas, J., Borrás, J., Martín, I., Moreno, V.: Assessing open heart
46 surgery mortality in catalonia (spain) through a predictive risk model. *European journal*
47 *of cardio-thoracic surgery: official journal of the European Association for Cardio-thoracic*
48 *Surgery* 11(3), 415–423 (1997)
- 49 38. Ray, B., Garain, A., Sarkar, R.: An ensemble-based hotel recommender system using sentiment
50 analysis and aspect categorization of hotel reviews. *Applied Soft Computing* 98, 106935 (2021)

- 1 39. Sen, P.C., Hajra, M., Ghosh, M.: Supervised classification algorithms in machine learning: A
2 survey and review. In: *Emerging technology in modelling and graphics*, pp. 99–111. Springer
3 (2020)
- 4 40. Siswahyudi, P., Kurniawan, T.A., Sugiarto, V.: Agent-oriented methodologies comparison: A
5 literature review. *Advanced Science Letters* 24(11), 8710–8716 (2018)
- 6 41. Subramaniaswamy, V., Logesh, R., Vijayakumar, V., Keerthana, K., Rakini, K., Swarnamalya,
7 B.: Dynamo: dynamic multimodal route and travel recommendation system. In: *2018 Inter-
8 national Conference on Recent Trends in Advance Computing (ICRTAC)*. pp. 47–52. IEEE
9 (2018)
- 10 42. Tureac, C.E., Turtureanu, A.: Types and forms of tourism. *Acta Universitatis Danubius. Eco-
11 nomica* 4(1) (2010)
- 12 43. Uddin, M., Wang, Q., Wei, H.H., Chi, H.L., Ni, M.: Building information modeling (bim),
13 system dynamics (sd), and agent-based modeling (abm): Towards an integrated approach. *Ain
14 Shams Engineering Journal* 12(4), 4261–4274 (2021)
- 15 44. Wang, J., Hong, Y., Wang, J., Xu, J., Tang, Y., Han, Q.L., Kurths, J.: Cooperative and com-
16 petitive multi-agent systems: From optimization to games. *IEEE/CAA Journal of Automatica
17 Sinica* 9(5), 763–783 (2022)
- 18 45. Weed, D.L.: Weight of evidence: a review of concept and methods. *Risk Analysis: An Interna-
19 tional Journal* 25(6), 1545–1557 (2005)
- 20 46. Zapata, G., Murga, J., Raymundo, C., Dominguez, F., Moguerza, J.M., Alvarez, J.M.: Business
21 information architecture for successful project implementation based on sentiment analysis in
22 the tourist sector. *Journal of Intelligent Information Systems* 53(3), 563–585 (2019)
- 23 47. Zdravevski, E., Lameski, P., Kulakov, A.: Weight of evidence as a tool for attribute transforma-
24 tion in the preprocessing stage of supervised learning algorithms. In: *The 2011 international
25 joint conference on neural networks*. pp. 181–188. IEEE (2011)
- 26 48. Zheng, X., Luo, Y., Sun, L., Zhang, J., Chen, F.: A tourism destination recommender system us-
27 ing users' sentiment and temporal dynamics. *Journal of Intelligent Information Systems* 51(3),
28 557–578 (2018)