Universidade Federal do Rio Grande do Norte
Centro de Ciências Exatas e da Terra
Post-Graduation Program in Systems and Computing
Masters in Systems and Computing

A Tourism Multi-user Recommendation Approach Based on Social Media Photos

Mickael Raninson Carneiro Figueredo

Natal-RN
27 November 2019

Mickael Raninson Carneiro Figueredo

A Tourism Multi-user Recommendation Approach Based on Social Media Photos

A M. Sc. Dissertation presented to the Department of Informatics and Applied Mathematics of the Center of Exact and Earth Sciences of the Federal University of Rio Grande do Norte as a partial requirement for the degree in Post-Graduation Program in Systems and Computing.

Advisor

Dr. Nélio Azevedo Cacho

Universidade Federal do Rio Grande do Norte – UFRN
Post-Graduation Program in Systems and Computing Natal-RN

27 November 2019

MICKAEL RANINSON CARNEIRO FIGUEREDO

"Uma Abordagem de Recomendação Turística Multiusuários Baseada em Fotos de Redes Sociais"

Esta Dissertação foi julgada adequada para a obtenção do título de mestre em Sistemas e Computação e aprovada em sua forma final pelo Programa de Pós-Graduação em Sistemas e Computação do Departamento de Informática e Matemática Aplicada da Universidade Federal do Rio Grande do Norte.

Presidente: **Dr. NÉLIO ALESSANDRO AZEVEDO CACHO**
(Orientador - UFRN)

Profª. Drª. Anne Magaly de Paula Canuto
(Coordenadora do Programa)

Banca Examinadora

Examinador: **Dr. ANTÔNIO CARLOS GAY THOMÉ**
(UFRN)

Examinador Interno: **Dr. BRUNO MOTTA DE CARVALHO** (UFRN)

Examinador: **Dr. DANIEL SABINO AMORIM DE ARAÚJO**
(UFRN)

Examinador: **Dr. JOSEFINO CABRAL MELO LIMA**
(UFRJ)

I dedicate this work to Robson, Marciedna, Victo and Barbara Figueredo.

Acknowledgments

I would like to thank Brazil for supporting high quality education free of charge for 5 years and now support an entire Master Education process during 2 years. I promise to honor and defend your values and fight every day to make my country a better place for next generations.

I would like to thank all the teachers who once gave me some knowledge because they are the basis for any student or great professional leaving university. A special thanks to Nélio Cacho, Andrea Cacho, Daniel Sabino, Marjory Abreu, Frederico Lopes, Carlos A. Prolo and Antônio Carlos Thomé for always giving me the best advice and knowledge they could have given. This is priceless.

I would like to thank José Lucas and Adelson Dias for all the fellowship and friendship since high school until today during entire process inside the University.

Creio que errar na forma é aceitável, desde que não se duvide jamais da intenção.

Bernardo Rocha de Rezende

Uma Abordadem de Recomendação Turística para Multiusuários baseada em Fotos

Autor: Mickael Raninson Carneiro Figueredo
Advisor: Dr. Nélio Azevedo Cacho

RESUMO

O setor de turismo é uma das atividades econômicas mais relevantes nos dias de hoje. Desta forma, é importante investir em diferentes abordagens para criar uma ótima experiência durante as viagens dos visitantes em um único destino. Em um contexto de *Smart Cities*, a idéia de *Smart Destination* aparece como uma solução para melhorar a experiência do turismo usando tecnologia para apoiar os visitantes na tomada de decisão em um *Smart City*. O estudo proposto cria uma abordagem para apoiar um *Smart Tourism Destination* para criar um melhor planejamento de viagem com base em fotos de mídias sociais. A pesquisa tem como objetivo criar recomendações para um único ou grupo de turistas utilizando técnicas de classificação de imagens e inferência fuzzy para mapear as preferências dos turistas. Através do sistema de inferência fuzzy e usando o conhecimento de especialistas em turismo dentro de um sistema de recomendação, a abordagem proposta é capaz de criar recomendações personalizadas usando atrações de uma *Smart Destination.*

Palavras-chave: *Deep Learning*, Smart Destination, Fuzzy Inference, Convolutional Neural Networks, Detecção de Preferências, Recomendação, Planejamento de Viagens.

A Tourism Multi-user Recommendation Approach Based on Social Media Photos

Author: Mickael Raninson Carneiro Figueredo
Advisor: Dr. Nélio Azevedo Cacho

ABSTRACT

The tourism sector is one of the most relevant economic activity nowadays. In this way, it is essential to invest in different approaches to create a great experience during visitors trips in one destination. In a context of *Smart Cities*, the idea of *Smart Destination* appears as one solution to improve the tourism experience using technology to support visitors in one *Smart City*. The proposed study creates an approach to support a *Smart Tourism Destination* to design a better trip planning based on photos from social media. The research aims to generate a recommendation to single or group of tourists using techniques of image classification and fuzzy inference to map tourists preferences. Through the fuzzy inference system and utilizing the tourism experts knowledge inside a recommendation system, the proposed approach can create personalized recommendations using attractions from one *Smart Destination*.

Keywords: *Deep Learning*, Smart Tourism, Fuzzy Inference, Convolutional Neural Networks, Preference Detection, Recommendation System, Decision Making.

List of Figures

List of Tables

List of Abbreviations

CNN – Convolutional Neural Networks
ILSVRC – ImageNet Large-Scale Visual Recognition Challenge
BOVW – Bag of Visual Word
SIFT – Scale Invariant Feature Transform
DoG – Difference of Gaussian
ReLU – Rectified Linear Units

Table of Contents

1 Introduction

Big data approaches have been used by many sectors to discover novel patterns and business insights that can improve public and private services. In particular, there is enormous potential in developing big data analytics in the travel and tourism sector (XIANG; FESENMAIER, 2017). Tourism is an experience-based product that requires a profound understanding of what today's travellers need and want, how they move through and interact with physical and social spaces, and what leads to their enjoyment, happiness, and the realization of personal values (XIANG; FESENMAIER, 2017). The tourism sector is also an important social and economic activity worldwide, mainly due to its capacity to generate jobs positions and create new business. According to WTTC (2017), the travel and tourism sector has been outpacing the global economy for the past six years, which is reflected in the growth figures for individual countries as well. In 2016, the sector directly contributed US$ 2.3 trillion and 109 million jobs worldwide. The tourism sector contributed US$ 7.6 trillion to the global economy and supported 292 million jobs in 2016. It was equal to 10.2% of the world's GDP, and approximately 1 in 10 of all jobs.

For that reason, it is increasingly important to understand the tourist's needs and behaviour at the destination to efficiently manage the locally available resources.

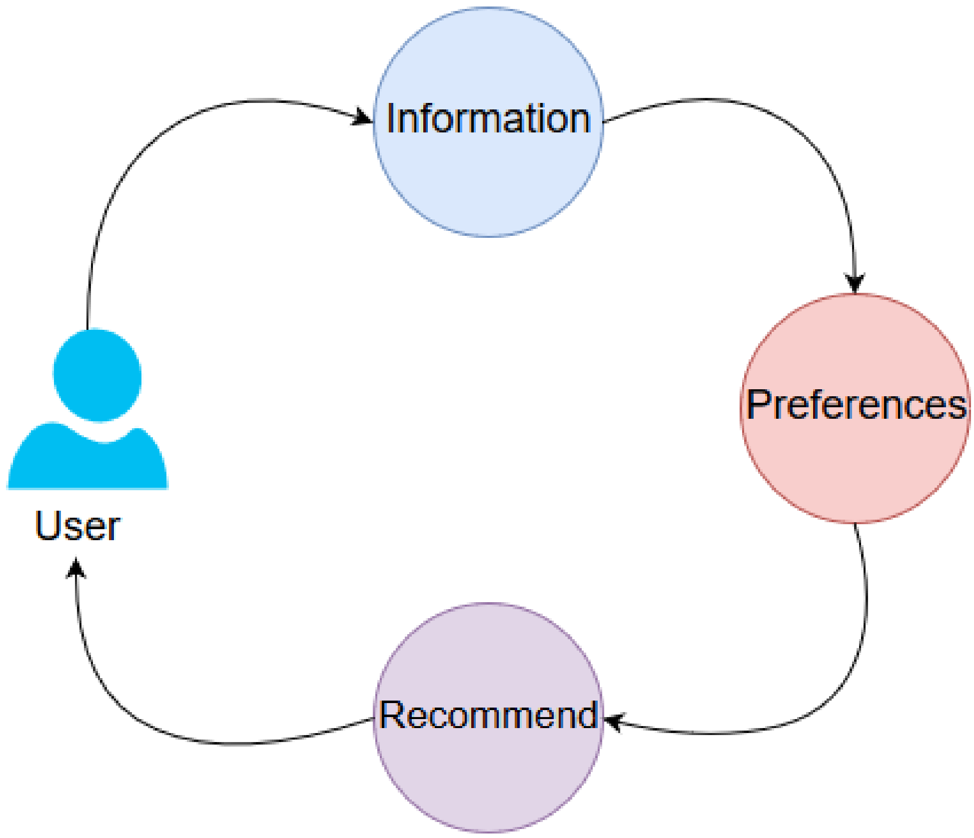


Figure 1 – The recommendation process in the pre-tourism experience

To support the tourism sector and many others, Natal has joined the IEEE Smart City initiative as an affiliated city (CACHO et al., 2016c). This initiative aims to transform Natal into a smart city through the development of systems to bolster the use of IT as a means of contributing to improving the life quality of its citizens and visitors. Regarding the tourism sector, such an initiative intends to strengthen data quality and improve information sharing aiming at creating a *Smart Tourism Destination* (BOES; BUHALIS; INVERSINI, 2015).

Smart Tourism Destination leverages the available technological tools and techniques provided by smart cities to enable demand and supply to co-create value, pleasure, and experiences for the tourist and wealth, profit, and benefits for the organizations and the destination (BOES; BUHALIS; INVERSINI, 2015). The focus of Smart Cities is on its citizens, whereas Smart Tourism Destinations emphasize the importance of enhancing the tourist experience. Moreover, Smart Tourism Destinations solutions can improve the process automation for tourism organizations to enhance destinations competitiveness. The most commonly used tool for improving the pretourism experience in a Smart Tourism Destination context are recommendation systems. These systems are based on the process shown in Figure 1 to improve the travel planning process. However, it is essential to map the user's preferences to create accurate recommendations, primarily when the recommendation aims at groups of users that have the same goal.

In this proposed study, we advocate a model able to recommend attractions to tourists or group of tourists using a fuzzy inference to map travellers' preferences through photos. First, we extract photos from social media or smartphones related to user's past travels and identify the scenarios frequented by the user. Furthermore, our algorithm relates the scenarios with five classes of tourism using the fuzzy inference approach to create a map of preferences. Finally, our approach can create recommendations based on one similarity algorithm suitable for our context. If the recommendation focus on a group of tourists, the proposed research can fit recommendations through a relationship between users. We tested our approach using real profiles and compared the results with humans evaluations to validate the preference extraction from our users. Feedback from users was collected when they tested the entire process. This feedback was used as a base to decide the best recommendation algorithm for our context. In this way, the proposed study uses feedback from tourists to improve each step of the system: image classification, fuzzy inference, recommendation and group modelling.

1.1 A Platform for Smart Tourism

Natal is located on the northeast of Brazil by the Atlantic Ocean. The capital city of the state of Rio Grande do Norte is home of approximately 862.000 thousands people. The city and the surrounding area are well known due to its sandy beaches and natural resources which attract thousands of tourists every year. According to the Brazilian Federal Tourism Organization, Natal is the fifth most wanted city by the domestic tourist who has a high income. The high number of tourists puts severe pressure on the urban infrastructure and services related to transportation, safety and water consumption. To handle such pressure, the IEEE Smart City Initiative of Natal has developed a significant synergy, between government and academia, envisioning the development of a smart tourism destination. Smart city concept covers a variety of industries, including the tourism industry (GUO; LIU; CHAI, 2014).

A smart tourism destination perceives as places utilizing the available technological tools and techniques to enable demand and supply to co-create value, pleasure, and experiences for the tourist and wealth, profit, and benefits for the organizations and the destination (BOES; BUHALIS; INVERSINI, 2015). Guo, Liu, and Chai (GUO; LIU; CHAI, 2014) argue that Smart Tourism Destination is a relevant part of the construction of the smart city's application system since it depends on the infrastructure of the smart city, utilization of information resources, and development of the intelligence industry. A smart destination should not only be limited by technological factors to provide a good experience for a tourist. Human factor and infrastructure should be considered in a smart destination context.

Smart tourism destinations are helping the development of new services in the tourism industry. A well-known case is related to the mobile industry, where smartphones have changed the tourism experience, opening up the field of advanced services applied to the travel and tourism industries (LAMSFUS et al., 2015a). For example, mobile technology enables people to travel both on the Internet and with the Internet, offering new opportunities for trip planning, and maybe providing more chances for engagement with other (LAMSFUS et al., 2015b).

Based on the importance of the mobile technology to the tourism industry (LAMSFUS et al., 2015a; LAMSFUS et al., 2015b), the Natal Smart City Initiative (CACHO et al., 2016c) designed, implemented and deployed a smart destination platform, named *Find Trip* Platform (CACHO et al., 2016b; CACHO et al., 2016a; CACHO et al., 2015). *Find Trip* platform provides technologies to collect, process, share, store, and analyses a vast amount of data coming from multipart sensing sources to turn data into powerful insights. Find Trip is the official tourism platform for the Natal Municipality and is in production since 2014. Find Trip was designed to fulfil the three phases of tourist experience (NEUHOFER; BUHALIS, 2012) (see Figure 2): pre-tourist experience, on-Site tourist experience, and post-tourist experience.

The *Find Trip* Platform (CACHO et al., 2016b; CACHO et al., 2016a; CACHO et al., 2015), a smart destination platform aimed to enrich tourists' travel experience through web and mobile applications. Find Trip intends to cover the three phases of tourist experience (NEUHOFER; BUHALIS, 2012): pretourist experience (before travelling),

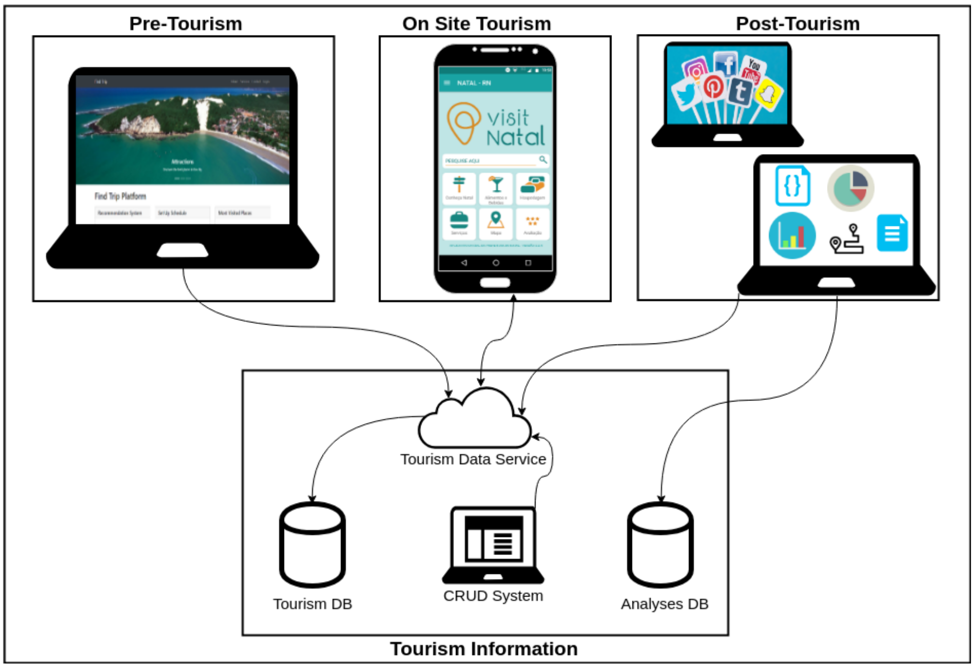


Figure 2 – Find Trip Platform

onsite tourist experience, and posttourist experience (after travelling). The previous work (CACHO et al., 2015; CACHO et al., 2016a; CACHO et al., 2016b) has presented some components of the *Find Trip* Platform that supports the onsite tourist experience and posttourist experience. *Find Natal* (CACHO et al., 2015; CACHO et al., 2016a) is a mobile application implemented in *Android/iOS* and supports the onsite tourist experience by providing tourist information and collecting the tourist's behaviour. *Social Smart Destination* (CACHO et al., 2016b) is a web dashboard that supports the posttourist experience by identifying early indicators of the tourism industry, exploring the presence of tourism issues and monitoring tourist's behaviour.

Pre-phase experience refers to all preparations and activities that tourists have before their arrival at the destination. During visit preparation, tourists use his/her social media account to log in the Web Interface. Pictures from his/her social media account are provided to the *Image Classification* component. This component classifies the tourist in five categories. Based on this classification, the *Recommendation* component provides a list of tourist attractions that are available for him/her in the mobile application during the onsite stage. This study relates to the implementation of this module from *Find Trip*.

On-Site tourist experience concerns all the touristic activities and happenings the tourists have and encounter at the destination (NEUHOFER; BUHALIS, 2012), Find Trip platform supports that stage by providing the Find Trip mobile application (CACHO et al., 2015; CACHO et al., 2016a) which is an Android and iOS application available for download at the Google Play Store and Apple Store. As depicted in Figure 2, the Mobile Tourist Guide sends user data (GPS location, language, and other values) from moving tourists to the Tourism Information System. The Tourism Information System stores tourist's data and provides a web system to manage the tourist information available at the Mobile Tourist Guide. As can be seen in Figure 2, information added, updated or removed about tourism attractions (i.e., museums, beaches) are automatically sent to the mobile tourist guide application. It allows making Find Trip users always up to date with the tourist information. This module was developed by other Students included in the entire project, and as a consequence, generate an application to be used in the on-site tourism phase.

Finally, the Post-Tourist Experience is related to the sharing of experiences through technology and helps them to their recollection and remembrance of the previously undergone travel. The Post-Tourist experience is supported by a Business Intelligence Infrastructure (CACHO et al., 2016b), which uses spatiotemporal data mining methods to extract useful information out of the moving tourist's data and from social network platforms. This component reasons on patterns and on pertinent background knowledge (tourism attractions data), evaluate patterns' interestingness, refer them to geographic information and find out appropriate presentations and visualizations. The information provided by this component allows planning traffic and public mobility systems in metropolitan areas, timely detecting problems that emerge from the movement behaviour and localizing new tourism attractions in our smart tourism destination. When a new tourist attraction is found, this information flows back to the Tourism Information System and finally to the Mobile Tourist Guide application. One example of the application of this study is identifying hotspots in the city using data collected from social media.

All the platforms that support some stage of tourism in the Find Trip Platform were developed together with several students. The focus of the proposed study is the approach that supports the first step from tourism planning. The research discusses a new approach to detecting user tourism preferences to accurate recommendations through photos. This work implements the recommendation system that supports the travel planning step on the Find Trip platform.

1.2 Problem Statement

Systems related to tourism recommendation are known to use different data types to create a preference profile and generate recommendations. Nowadays, through the advent of social media and information propagation technologies, diversity of data is generated. It can be used to model tourism preferences and create a personalized recommendation. However, it is essential that the adopted model to map user preferences be accurate to create satisfactory recommendation even in the context of groups. Also, it is crucial that the recommendation algorithm used in the approach be appropriate to the type of data used and also to the context applied.

This study is strongly motivated by the limitations to create tourism recommendations in the most common methods(HANAFI et al., 2018), such as the *Cold Start* problem. Currently, most systems seek to map tourism preferences using information that is explicitly related to tourism or general information such as demographic data. The use of information directly related to tourism generates a persistent problem: The system user must have a travel history. Also, this information must be present in the system either given by the user. In this way, it is possible to generate a preference profile. However, asking users to feed the system can be something that makes the recommendation experience somewhat frustrating for the user. On the other hand, the use of demographic data means that most of the generated profiles are stereotyped, resulting in generic applications with the use of these preference maps. In this way, this study aims to create a new way to identify, classify and create a recommendation to tourist and groups based on preferences detected from photos present in social media. The proposed approach is applied in the Natal city context. The overhead of the need for explicit information related to tourism is overcome by a public data source, in this case photos related to travels.

1.3 Research Goals

The idea of the study is to identify environments frequented by users that are related to good experiences. We can infer what types of classes of tourism the user tends to like relating it to the frequented environments. It is possible to extract tourism preferences for each system user to create a set of recommendations of Points of Interest. This study has as contributions: I - A comparative between a standard image classification technique and a modern approach in the context of scenario classification. 2 - Creation of a fuzzy model for the relationship of scenarios with types of tourism. 3 - An approach for extracting tourist preferences using photos from social media 4 - Definition of a recommendation algorithm for the tourism context 5 - The application of methods to create a recommendation for groups.

In the proposed approach, a considerable *dataset* of images containing more than 1 million images was used to create a scenario classifier using supervised learning. A study was made to identify what classifier better fit to our case. We confront two approaches. First, one standard approach based on the technique of *bag of words* large used in the natural language processing context. After, a modern approach using *deep learning* was trained using moderns techniques to improve the accuracy and reduce training cost. The two approaches were compared in Section 5 to identify what is the best method to fit in a scenarios classification context. Subsequently, a study was developed in partnership with experts from the tourism area to identify which types of tourism classes should be used to map the tourists' preferences. In this way, we can relate the types of environments and the classes of tourism for the detection of preferences. This information can be used to understand how much a tourist tends to like a type of tourism. For this, a fuzzy module was implemented. The fuzzy module is the primary key for the creation of a tourist preference map. The next step was to test the most famous recommendation algorithms in the literature and confront then with one proposed approach of recommendation. Finally, the recommendations of singles user are related using a robust approach proposed in other researches. This study validates each step of the process: image classification, fuzzy inference and recommendation. To conclude the validation, real tourists were used to creating tests situations that could be analyzed in Section 5.

2 Background

In this chapter, previous work related to Smart Tourism, Scenario Classification, Preference Detection, Recommendation and Group Modelling is discussed to introduce these five critical concepts for this work. Then we move into our approach to describe how these concepts were used to achieve our goals.

2.1 Smart Tourism

The growth of the cities brings with it complexity and management challenges to the government authorities in dealing with problems related to water supply, local waste disposal, urban traffic management system, health, education, public safety, economy, environment and tourism. In this sense, the great challenge to be faced is to ensure sustainable urbanization associated with socioeconomic progress.

Politicians around the world are seeking for answers and ways to deal with these challenges. One of the strategies proposed encompasses the creation of smart cities. The work on (CARAGLIU; BO; NIJKAMP, 2011) argue that a city can be defined as "smart", when there is an investment in human and social capital, as well as in information and communication technology (ICT) infrastructure.

Smart city incorporates a large number of systems, which represent the most basic infrastructure for integrating the real and virtual worlds. One of the significant challenges of deployment of smart cities is the extraction of relevant information from the ICT infrastructure of cities. Such mining usually relies on the use of sensors that are installed to capture the flow of vehicles, water and energy consumption. This context requires high public investment for the development of smart city (KOMNINOS; PALLOT; SCHAFFERS, 2013). In Cohen (COHEN, 2011), the author proposed what is called *The Smart City Wheel* which defines the six smartness dimensions important for the development of a Smart City, including Governance, Environment, Mobility, Economy, People and Living. With this concept, it is clear that a smart city is made not only of technological concepts but also of human, social and economic ideas.

Within this concept, the idea of Smart Tourism emerges as a new niche. A tourist destination contains within it a social and economic environment that enjoys a complex infrastructure to provide services (BOES; BUHALIS; INVERSINI, 2015). These factors make the management of resources embedded in this context a complicated task. In this way, the concept of Smart Tourism emerges as a way to improve the management of touristic cities (which become fashionable destinations) through the use of technologies that encompass the idea of smart cities. With this, a city can create a more pleasant and practical tourist experience for a tourist through the most efficient approach in terms of use of resources.

2.2 Scene Recognition

Many approaches have tried to imbibe the advanced human vision system into different models and algorithms. Scene classification and analyses are highly useful of humans, who can classify complex natural scene though the understanding of the components that belong to an image. In this way, the need to understand the context and recognize different patterns present in an image makes the task of classifying a scene one of the most difficult in the field of computer vision.

A *scene* (XIAO et al., 2010) is defined as a place in which humans can act within or a site to which a human being could navigate. However, this concept is captured in the scene classification context. The first algorithms applied to scene classification and recognition was based on low-level images features such as texture using Fourier Transform, RGB histogram and brightness. In the Renninger and Malik research (RENNINGER; MALIK, 2004), the proposed algorithm tries to mimics the human's capability to classify and identify objects in limited light exposure to classify scenarios. In this algorithm, the low-level feature used to build the classifier was the image texture extracted using the *Gabor Filter*, based on the theory of human visual perception of texture from Béla Julesz (JULESZ, 1981). Julesz defines *textons* as the strength of texture discrimination. This concept is the elements in the image that govern our understanding of texture. These *textons* could be used by the human brain to understand a set of elements that compose one image. These approaches can be very accurate when applied in simple situations, where angles, positions, colours and other patterns are standards. Low-level features are not able to describe the behaviour of inconstant changes in the context of scene recognition, where patterns are not easily detected.

Sophisticated approaches uses images descriptors (LOWE, 2004)(BAY et al., 2008a) to extract features to create robust images classifiers. In the Csurka (CSURKA et al., 2004), a generalist image classifier was developed using the SIFT (LOWE, 2004) keypoint detection algorithm to extract features from images in a training dataset. Then, a vocabulary of image descriptors is created after applying the vector quantification algorithm. A bag of keypoints which counts the occurrence of features in the vocabulary is created. Finally, a classical classification algorithm is applied to treat the bag of points as the feature vector and thus, determining which category the image belongs. In the Csurka (CSURKA et al., 2004) work, a Support Vector Machine and a Naive Bayes classifier are tested to present the best classifier. This *metafeature* approach is used instead of the original image to train the classifiers.

In the last years, image classification and object detection have improved due to advances in Deep Learning and Convolutional Neural Networks (CNN). The pioneering on this approach are Krizhevsky, Sutskever and Hinton (KRIZHEVSKY; SUTSKEVER; HINTON, 2012) that successfully conceived and proposed a Convolutional Neural Network architecture to classify 1.2 million of high-resolution images from ImageNet Dataset (DENG et al., 2009). The group discussed the architecture of the network called AlexNet, with a relatively simple layout compared to modern architectures. The network was made up of 5 convolutional layers, max-pooling layers, dropout layers, and three fully-connected layers designed to classify one image into 1000 possible categories achieving 15.6% test error rate. Many pieces of research were produced from the excellent results obtained from AlexNet. Other variations (SIMONYAN; ZISSERMAN, 2014)(HE et al., 2015) started from the main idea of increasing the depth and width of the convolutional layers, so the features are extracted in the best possible way. Recently studies obtained great success reaching to conquer the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC)(RUSSAKOVSKY et al., 2015). One of the proposed and winning ILSVRC networks were used in this study as a feature extractor. The GoogLeNet is as 22 layer CNN winner of ILSVRC15 (RUSSAKOVSKY et al., 2015) with an error of 6.7%. Convolutional Neural Networks have a standard structure with stacked convolutional, max-pooling and dropout layers and fully connected layer. The main difference in GoogLeNet module is that all filters are learned and all the layers are repeated many times, leading to the 22 deep layers cited.

However, as with any classification technique, deep learning has its shortcomings. The main one that can is the very high cost for training these networks. A considerable computational power would be necessary to prepare an accurate network with a large amount of data. Moreover, a large dataset is always required for successful training an efficient CNN.

In this chapter, two image classification techniques were detailed to confront them in Chapter 7. The first one is the *Bag of Visual Words* (CSURKA et al., 2004). In computer vision, the bag of visual words (BOVW) can be applied to image classification, by treating image features as words and train the classifier using these words. The second approach is an emerging technique. Image classification using *Deep Learning* was proposed as a way to change everything that is known about image classification. These two approaches are described in the following subsections.

2.2.1 Bag Of Visual Words

The Bag of Visual Words (BOVW) algorithm is an extension of the natural language processing algorithm applied to image classification. BOVW create a vocabulary that can best describe an image by following a few steps.

First, it is necessary to determine the features of an image and to extract them. In this way, it is possible to create a vocabulary by clustering and analyzing the features

(a) Beach with coconut trees in the left

(b) Beach with bigger coconut trees in the right

Figure 3 – Example of scale and rotation variance in beach class images

occurrences. Finally, a classifier can be trained using the extracted features occurrences histogram.

A feature to be extracted in the context of images can be considered a set or group of points that are relevant in an image. The selection of these features changes depending on the method used. The implementation of the BOVW algorithm developed in this work the Scale Invariant Feature Transform, or SIFT) is used as a feature extractor. When the dataset is composed of similar images (same scale, orientation) simple corner detectors could be enough. However, when the data contains images of different scales and rotations, is it necessary to apply a robust feature extractor, such as SIFT. One example of this variation in scale and orientation can be seen in Figure 3. The coconut element in the Figure 3a and 3b can be seen in different positions and scales. Using the SIFT descriptor, it is possible to consider both as coconut trees efficiently

The SIFT algorithm follows several steps to extract the features from images. The initial preparation of the process is the construction of the scale space. Scale-space is an essential concept in human vision. Some elements in the real world only can be understated when they are on a reasonable scale. In image processing, scale-space (LOWE,

2004) is a technique to represent an image at different scales. In SIFT algorithm, this is made by using the original image and then progressively blurred the image. Blurring, in a mathematic context, is the convolution between the Gaussian operator and the original image. This situation is described in Equation 2.1, where L is the blurred image, G is the Gaussian Blur, I is the original image, and θ is the scale parameter. Finally, to detect stable key points locations in scale-space, the SIFT algorithm uses the scale-space extrema in the Difference of Gaussian (DoG) function convolved with the image, given by the Equation 2.2

$$L(x, y, \theta) = G(x, y, \theta) * I(x, y) \tag{2.1}$$

This step works as a filter to identify locations and scales that can repeatable assigned under different views. Then, the DoG images are used to calculate Laplacian of Gaussian, that is scale-invariant, to find the *maxima* and *minima* points. The points in a DoG image are considered *key points* if it is most significant or least of all 26 neighbours.

$$D(x, y, \theta) = L(x, y, k\theta) * K(x, y, \theta) \tag{2.2}$$

After extract the features using the SIFT algorithm for all the training data set, the next step in the implemented Bag Of Words technique is to cluster the features. The clustering technique used in this implementation was the *K-Means*. Suppose there are X objects that are divided into K clusters. The input of this process can be a vector of features $X = x1, x2...xn$. The algorithm seeks to reduce the distance between each point from the dispersion cloud and its respective cluster centroid. BOVW uses a training approach that involves partitioning similar characteristics from the set of training images. In this way, a quality dataset is essential to achieve good results. This can be ensured, in this study, by using the *Places 365* (ZHOU et al., 2017) dataset.

The next important step in the BOVW algorithm is the development of vocabulary. Vocabulary can be considered a set of features that best describes a set of images. In this case, for a beach image, the set of points such as waves in the sea, sandhills, coconut trees and beach chairs can form the vocabulary of this class. In this way, there is a combination of standard weights that describes an image individually, and then each features present in an image can be used to describe the same image.

During the definition of the vocabulary, it is necessary to identify the cluster that contains the feature, which is the cluster whose centroid is closest to the current feature. Finally, we can start the training process, since each image can be represented using the frequency of each visual word. The implementation used in this study uses an SVM classifier.

2.2.2 Deep Learning

The deep learning approach takes inspiration in the biological field. The idea is to identify an object or scene based on the detection of edges, colours and forms, as made by the human visual cortex. For example, some neurons fired when exposed to vertical edges and some when shown horizontal or diagonal edges. From these ideas comes the inspiration for deep learning.

There are many variations of deep learning approaches. This study uses Convolutional Neural Networks (CNN). A CNN follows a well-defined struck composition to be used. First of all, the Convolutional Layer is always present. In this layer, one image is convolved using a kernel. In a deep learning approach, many convolutional layers filter one through during the process. This layer is used to extract the features from the images, and each layer can be used to extract different kinds of features.

Another layer usually used in CNN approaches is the *Rectified Linear Units (ReLU) Layer*. After convolving several time images, it is essential to remove the linearity from a system that only took linear operation during the process. In the past, nonlinear functions like *tanh* and *sigmoid* were used, but researchers found out that ReLU layers work far better because the network can train faster. Equation 2.3 describes the function of ReLU when applied in a image I. The layer removes the negative components from a convolution output.

$$f(x, y) = max(0, I(x, y)) \tag{2.3}$$

Usually, after a ReLU Layer commonly uses a *Pooling o Layer* in CNN approaches. This layer is responsible for the sample-based discretization process. The objective is to down-sample an input representation. In this case, the number of pixels inside a window is reduced to one value. The value depends on the kind of pooling layers used. For example, in a Maxpooling Layer, the output value for a window is the max value in that window. Another variation is the Average Pooling layers, which the mean value of the window is used as the result. Some studies (SPRINGENBERG et al., 2014) consider the pooling layer useless in favour to use more simple CNN structures and to train good generative models. On the other hand, some authors propose the pooling layers as a way to avoid overfitting in the training step.

Another critical step to avoid the overfitting in the training process is the use of *Dropout Layers*, propose by Hinton et Al.(SRIVASTAVA et al., 2014). In this kind of layer from Convolutional Neural Networks, the objective is removing a set of random values in that layer by setting it to zero. This kind of approach forces the network to be redundant, and even an image is used in different ways in the training step.

Finally, the last layer in all *Convolutional Neural Networks* is the Fully Connected Layer. This layer is the most simple to be explained. It is a simple Neural Network which receives as input the result of the processing of the other layers and has as output the class belonging to an image. The number of neurons in the output layer indicates the number of classes to be classified. How this approach was used and trained in our context is described in section 5.

2.3 Recommendation

Detect and map preferences are the most crucial step in the recommendation process. Many problems which are needed to describe the preferences of users are directly related to recommendation systems. The process of developing profile preference extraction techniques could only be advanced through the advent of recommendation algorithms. A profile must be precisely mapped.

Recommender systems are widely used for the suggestion of products, activities and services. Users in implicit and explicit ways can find recommender system every day. A large amount of data is generated by users, creating an information overload about profiles, and challenging in mapping preferences process. In this context, different kind of approaches and algorithms have been implemented and applied to suggest relevant items to users (ADOMAVICIUS; TUZHILIN, 2005) and solve the recommendation problem.

Recommender systems are personalized information agents that provide recommendations based on features or knowledge to map the preferences from users (RESNICK et al., 1994; GOOD et al., 1999). Recommender systems are divided by the recommendation inputs given by the users, which the system then aggregates and directs to appropriate recipients. The most widely used recommender approaches (HE; PARRA; VERBERT, 2016) are the *collaborative filtering*, *content-based* and *hybrid recommender*. However, there is others variation of algorithms such as *demographic based* and *knowledge based*. The first appearance of the study of these four recommendations systems happened in the 90's decade, so these terms are consolidated in the literature.

The most classical and widespread approach in the industry is the collaborative filtering(CF) based algorithms (RESNICK et al., 1994)(SHARDANAND; MAES, 1995). This technique (Figure 5) creates the user preference profile based on opinions from other users who share similar interest. The approach recognizes the similarity between users or items based on implicit or explicit information shared in their profiles (BURKE, 2007). This technique has two mains variations: the item based (SARWAR et al., 2001) collaborative filtering and the user based collaborative filtering. The great variation between these two approaches lies in the type of object where the degree of similarity is calculated, and the preference profile is generated. A classical item-based algorithm identifies the similarity between items, and recommend a similar set of items for the user who has preferences for the original item. Standard user-based recommendation recommends objects for a user based on users with similar profiles, tastes and preferences. The CF approach (MALTZ; EHRLICH, 1995)(MIDDLETON; ALANI; ROURE, 2002) is one of the most suffering

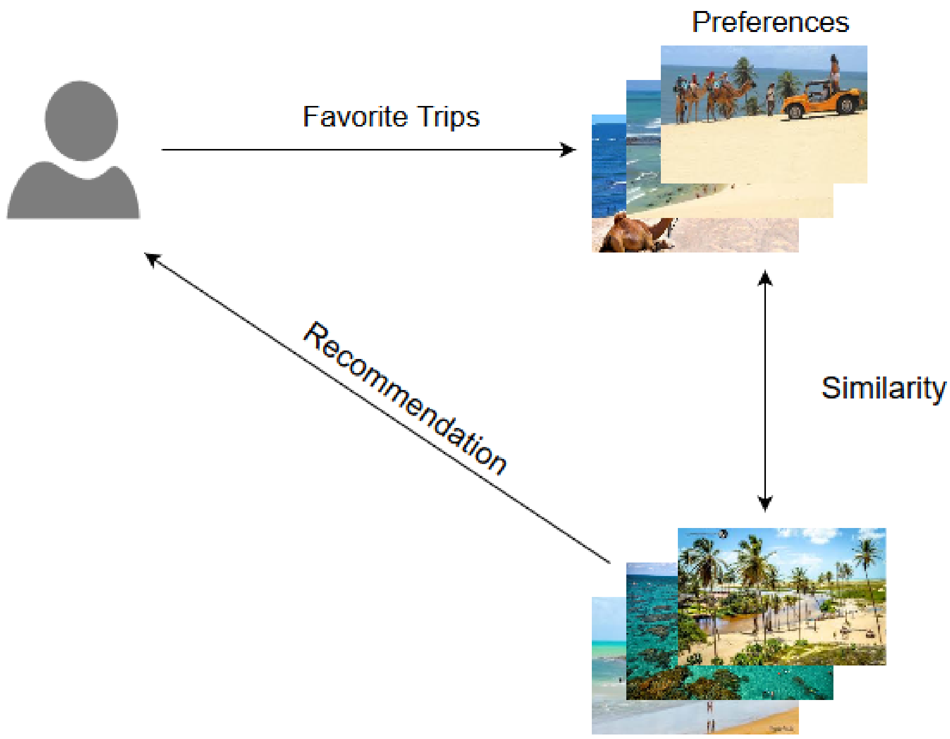


Figure 4 – Preferences detection in content based approach

from the cold start (MALTZ; EHRLICH, 1995). In this situation, users start with any preferences information in their profiles and the inferences must be made from scratch. Even with a starter profile, there is still an overhead period before the profile accurately reflects the user's preferences. In this way, a good recommendation can only be generated on this type of system after a reasonable amount of previous information about the user is collected.

The content-based technique (Figure 4) creates the user preference profile though the old preferences liked by him. The next step is to recommend objects in dataset universe that are similar to objects preferred before by the user (PAZZANI; BILLSUS, 2007). Content-based recommender systems are classifier systems derived from machine learning research. Unlike the collaborative approach, information about other users of the systems is not essential. The recommendation is based only on the content generated by the user and the objects to be recommended. In this case, it is vital how the items are mapped in the dataset. However, the cold-start problem is still a recurrent problem even in content-based systems. Because it is a machine learn-based approach, without the right amount of items in the databases, the classifiers and models are not able to generate an accurate recommendation (BURKE, 2000). Another usual problem in the machine learn-based approaches is the difficulty of change user recommendation and preferences. If one inference is generated to describe a user preference, it is kept until a refresh on rating and historical user's dataset.

The hybrid approach was proposed to overcome the cold-start problem and other

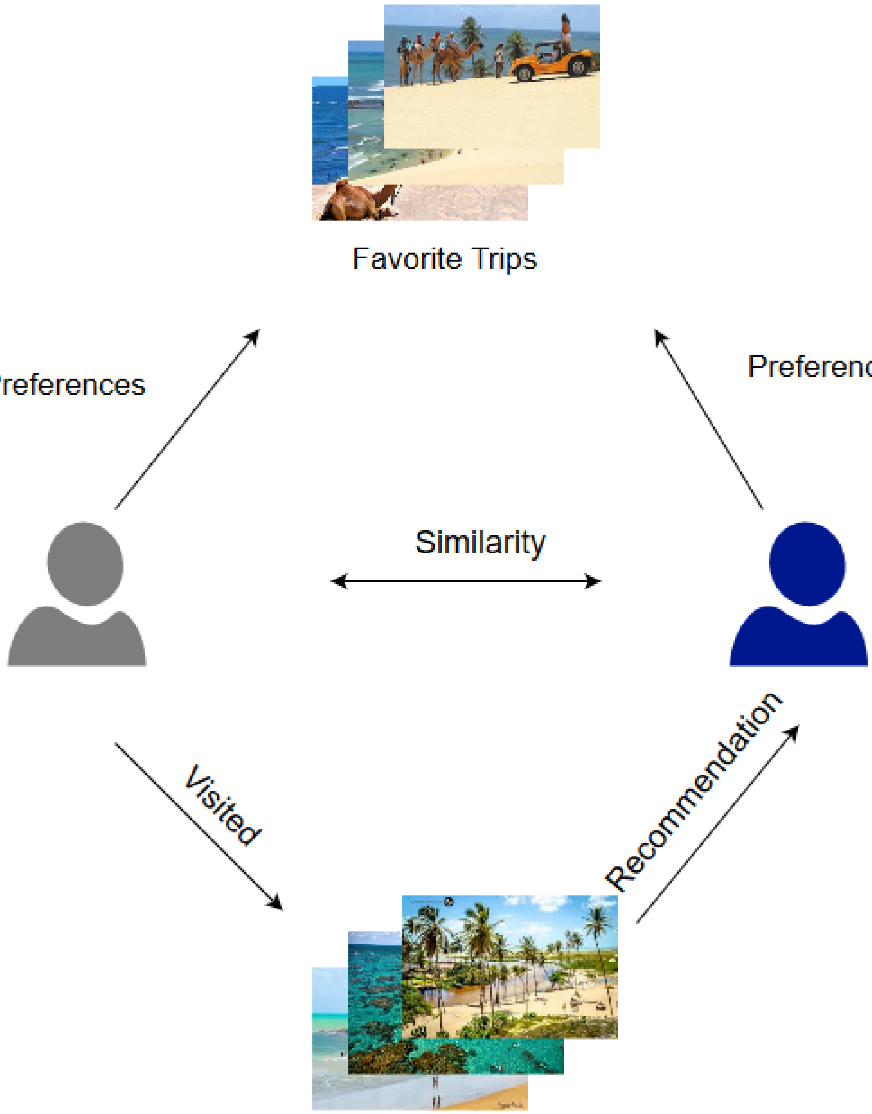


Figure 5 – Preferences detection in collaborative filtering

drawbacks such as sparseness and scalability(BURKE, 2007). A hybrid recommender system is used to describe an approach that combines multiple recommendation techniques. It is possible to hybridize all set of algorithms in the literature. Some hybrids combine systems different implementations of the same class of technique – for example, switching between two different content-based recommenders. Another approach that aims to avoid the classical model's problems is the knowledge-based recommendation (BURKE, 2000). In this approach, it is necessary to create a knowledge base about users, items and their relationship. This kind of algorithm does not suffer a ramp-up problem because it is independent of users previous ratings. The knowledge technique aims to create users and items inferences based on knowledge create by an engineering process. This knowledge can use implicit and explicit features to meet the needs of a user.

Some models to map the preference profile of a user tries to use some semantic models to create the preferences from *implicit informations*. In the usual cases for the *content base* and *collaborative filtering*, the users can inform the system directly what kind of data that he wants to receive a recommendation. In this proposed study, use *explicit information* would be requested to the user feeds the algorithms with information about previous trips. However, models such as *semantic models* and *uncertainty models* tries to use implicit information to model the user's profile. In this case, the systems do not need information directly related to tourism to recommend places on trips. Ho et al.(HSU; LIN; HO, 2012) creates one preference profile able to give the probability of one attraction be relevant to one user using feature such as age, work, trip motivation and nationality. Another recent approach used in some studies (GARCíA-CRESPO et al., 2009) is to generate the tourist profile using fuzzy logic. In this case, a set of information to understand how much a user matches with one attraction. The matches can be used to map and relate the user preferences with some other group of users.

2.4 Similarity Measure

In this study, three algorithms of similarity were applied to the recommendation. However, for the final pipeline only one algorithm is used. The choice of the algorithm is based on the validation described in the next chapters.

In this section, algorithms are described and shown how they treat attractions and tourists in the process.

2.4.1 Cosine Similarity

Also known as vector-based similarity, this formulation represents two items and their ratings as vectors and defines the similarity using the angle between these vectors. In the application context, each user and attraction are mapped into a five dimension vector. The similarity between a user u_i and an attraction a_i using the Cosine Similarity is calculated in our system using Formula 4.1.

$$similar(u_i, p_j) = cos(\vec{u_i}, \vec{p_j}) = \frac{\vec{u_i} \cdot \vec{p_j}}{|| \ u_i \ || * || \ p_i \ ||} \tag{2.4}$$

The Cosine Similarity metric is efficient when applied in a context where the vector's magnitude is not essential. However, when there is a significant variance in the vector components value, the result of a cosine recommendation approach is inferior.

In our context, to use the cosine similarity, it is essential to put the attractions a_i and users u_i in the same universe. In this case, each vector is represented at five dimension universe. Using the knowledge from tourism experts, We can transform our database of attraction into the universe of Landscape, Urban, Historical, Shopping and Sports tourism. As said, the cosine similarity will ignore the magnitude of the attraction and the user. What is important in this case is users and attraction with the same vector direction. In this way, an attraction may be relevant if it has the same tendency as the user profile, even if the relevance has different values for both cases.

2.4.2 Euclidean Distance

The Euclidean Distance Metric corresponds to everyday experience and perception. Usually, visualize how this metric is related to real-world perception is simplify when applied to 2 or 3-dimensional points. However, in this proposed platform, this metric is used to relate elements mapped into a five dimension space. The Euclidean metric calculates the straight-line distance between two points in the space. In our case, each user u_j from the universe U has this metric calculate for each attraction p_i in the universe of attractions P. This metric is calculated in our approach using the Euclidean Distance Formula, shown in the Formula.

$$dist(u_j, p_i) = \sqrt{(u_{j_1} - p_{i_1})^2 + (u_{j_2} - p_{i_2})^2 + ... + (u_{j_5} - p_{i_5})^2} \tag{2.5}$$

The Euclidean distance is a robust metric to relate elements with variance in their values for each dimension. In this case, attractions and users are not treated as vectors, but rather as points. For an attraction to be relevant to a user, the relevance values for each of the five tourism classes must be very close compared to the five values present in the user profile. In this case, the magnitude is significant, situated opposite to that found in the similarity of cosines. If users feel satisfied with this type of recommendation, it means that for an attraction to please a user it needs to have very similar pertinence compared to the profile of the users.

2.4.3 Formula

This algorithm proposed in this study seeks to be an option for both classical approaches. The idea here is to identify profile combinations that suit the attraction. However, this approach makes use of a τ value that filters only relevant values in the process.

Each Point of Interest p_n in the Point of Interest universe P has pertinences to each tourism class j(landscape, adventure, historical, urban and shopping) where n is the number of POI in our database.

Each tourist t_i in the profiles universe A has the same domain of pertinences of POI. In this way, every point of interest p_n has one score S for one specific tourist profile ti. The total score of one attraction for one specific user is the sum of the *subscores* s_j for each tourism class j. The formula that relates attractions with user is shown in Formula

2.6 where n is the POI number in our dataset.

$$S(p_n, t_i) = \sum_{j=1}^{5} s_j \tag{2.6}$$

Subscores are 0 for tourism classes j with low pertinences values in an user profile t_i. Otherwise, the *subscore* s_j is a value that measure the distance between p_{nj} and t_{ij}. Higher the similarity of these values are, greater are the *subscore* s_j. The value of τ showed in Formula 2.7 used in this research was 0.3 to avoid insignificant user pertinences t_{ij} be used in the estimation.

$$s_j = \begin{cases} 1 - |p_{nj} - t_{ij}|\,, & t_{ij} \geq \tau \\ 0, & t_{ij} < \tau \end{cases} \tag{2.7}$$

High scores can be achieved if POI has similar pertinence compared with user pertinence. In this way, for each user in our system, the entire set of recommendation is generated based on ordered scores from P universe. This method was proposed as a way of making recommendations based only on pertinence with relevant values. In this way, the value of such is significant for this filtering. If the value of such tends to zero, this type of recommendation will work similarly to the Euclidean distance. However, the higher the value, this recommendation will work using only values considered relevant in the recommendation, which can generate an accurate recommendation considering single tourism types that are important to users.

2.5 Group Modelling

Many activities, such as going to shop or planning a trip involve a group of users in the process. In such a case, recommendations systems must support a friendly approach to generate an accurate recommendation based on the entire group preference. The system should combine users preference and tastes to identify agreement points among the group. This process is divided into two primary steps: (I) Identify individual preferences without the group considered content and knowledge. (II) Find an item, in our case attractions, that the entire group reasonably accepts it. An efficient approach to identify the agreement topics in the group is to determine the success of the recommendation.

The most common algorithms for this type of study are *Aggregation* and *Intersection* (GARCíA et al., 2009) (CHRISTENSEN; SCHIAFFINO, 2011). The idea inside of these two algorithms is straightforward. Once the individual preferences for each user inside a group are modelled, they are combined in one way to create the group recommendation. These algorithms can be applied in the preference mapping for the users or in the recommendation if the recommendation list contains some score for each user. The Intersection algorithm tries to identify preferences or items in the recommendation list that are interesting for both users. However, this approach can create a group profile or a recommendation empty when users have no immediate interest. To avoid this problem, the algorithm that opposes this idea is the Aggregation recommendation. In this case, the list of preferences or recommendation is the result of the merge from both users. In this case, the entire list has the same relevance. The problem created in this classical approach is the possibility of creating generic profiles that do not please the group in a general way.

However, a more sophisticated model that was considered usable in the current context was the algorithm of *Incremental Intersection* (GARCíA et al., 2009). This proposed algorithm overcomes the problems of classic algorithms trying to balance aggregation and intersection.

2.5.1 Intersection

The intersection algorithm applied in the recommendation context for groups is the simplest of all three used in this study. Like any intersection between groups, this operation will search for elements that are present in all groups.

In the context of the tourist application, the use of the intersection algorithm is simpler when applied to the lists of recommendations of the individual users. Thus, given the sets of recommendations for individual users $L =< l^1, l^2, ..., l^n >$, where n is the size of the group, the list of final recommendations R will be given by $R = l^1 \cap l^2 \cap ...l^n$. However, each element of the recommendations list has a score generated by the similarity algorithm used. In this case, the new score for an attraction that is in the set R will be given by the average of the scores given for each attraction by each user.

On the other hand, the intersection applied to map a single profile to the group of tourists becomes more complex due to the need to define intersection in this context. Each profile is modelled as a vector of 5 numerical values that define how much a user has preference in each of the 5 tourism classes (landscape, cultural, urban, sport and shopping). In this way, we defined intersections as the situation the case where users have similar preference levels for each of the tourism classes. In this way, a value was defining to tell if an interval is small enough to be considered an intersection.

The major problem of this approach is the difficulty of supporting the growth of groups. As one group becomes larger, the smaller the likelihood of intersections occurring between them. In this way there is the risk of the recommendation being a very small or even empty list. In the case of an attraction, even though it is relevant for many users, it is not recommended for the group using the intersection algorithm because is not relevant for one user.

In our context, each tourist has your own list of N points of interest to be visited, The approach in this case is compare the lists using the top N attractions for each profile and use the attractions that appears in every single list. The worst case in the tourism situation is a group of many tourist with real different tastes, where there will be no intersection between the top N list attractions. This situation may have the chance of occurring reduced by using a small attractions database. This approach can be used in this research, as it is taken as a case study. However limiting the set of attractions to be recommended is not interesting for a production environment.

2.5.2 Aggregation

The aggregation algorithm is the most widespread among group recommendation systems. In this case, the implementation is simple in both application situations (in the profile and in the list). In the case of the application in the set of recommendation lists $L =< l^1, l^2, ..., l^n >$, the recommendation for aggregation will be given by $R = l^1 \cap l^2 \cap ...l^n$. However, in this case, the score of each attraction is placed as the largest or the lowest (dependent on the similarity algorithm used) value between the equal elements of a set.

In situations where the aggregation algorithm is applied to generate profiles, the same logic is applied to preference levels. However, the calculated value for the preference level for a tourism class is given by the group average for that class. A variation of the classical aggregation model is Aggregation without Misery, where very low or high values are disregarded in the calculation of the final average. However, the model used in this study was the classic one.

The problem faced by this type of system is the possibility of generating generalist recommendations, causing the dispersion of the individual information of each user. In the tourism context, as in the intersection approach, the score of each attraction is not used in the aggregated list. The attraction for each tourist in the group is merged to create a final list. However, apply this approach in large groups can create a large number of points of interest to be visited without considering the users preferences. In addition, the travel time factor can be a problem for this approach, since a tourist may not spend enough time in a city so that along with his group can visit the list of aggregate attractions created.

3 Related Works

Based on the amount of information available on internet reporting users experiences in different contexts, the recommendation has become an increasingly important task in a system that seeks to stand out among others. This situation can be seen in the context of tourism, where more and more tourists have more tools to report their experiences when visiting a city. Recommender systems(ADOMAVICIUS; TUZHILIN, 2005) has the purpose of filter options and provide personalized content for each particular user based on their preferences. In the tourism context, the customized content is related to Points of Interest (POIs) or attractions.

A large amount of studies(ADOMAVICIUS; TUZHILIN, 2005)(BURKE, 2000) tries to map the preferences from users for diversity of applications and using a different kind of data. In Loh et al. (LOH et al., 2003), a system to map the preferences of tourist is developed to be used by travel agents. The application aims to improve the recommendation for costumers, especially for tourists who do not know where want to travel. In this case, the focus is on identifying the best city or country which fits for the traveller taste. The study avoids using information directly related to previous trips. The centre in this application also helps tourist that do not have the knowledge and historical data about their previous trips. The approach is to use text mining from private tourist's webchats. It is a case of study where the data to map the tourist preference is indirectly related to tourism. The text mining in the webchat tries to extract exciting areas that will outline the tourist preferences. The system makes use of tourism ontology to identify the themes and exciting areas in the textual messages. However, an approach that improves the recommendation is the traveller agent, who works as a filter over the recommendation created. In this way, the recommendation does not act directly with the traveller. It is the case when the customer arrives at the travel agency without a plan or destination.

Another interesting approach is described in the Codina et al.(CODINA; CECCARONI, 2010) proposed study. In this work, the authors propose users profile modelling approach to avoid the most frequent problems contained in the most traditional kind of recommendation systems. The work presents a system whose represent the user profile using the concept of taxonomies. The approach must be able to be applying in a different context such as tourism, books, movies and music making it domain-independent. In this case, the system learns the user's interest from a variety of information sources using hybridization of user-information collecting-techniques. The first option to create the user model is to collect explicit feedback, such as user-rating and user own preference directly described. In a second case, the system tries to infer the user preference map based on implicit feedback using user's searches and events in the web browser historic. If neither of the first two approaches is viable, the latter case is to generate a stereotyped profile based on domain generalizations. Through the user interaction with the system, the algorithm must be able to learn new preferences and keep the profile modelling up-to-date.

An excellent way to create a user preference profile is to model the problem as a social choice problem, as proposed by Albanese et al.(ALBANESE et al., 2013). In this study, user profile modelling is necessary to create recommendations related to multimedia data. The study case focused is to recommend art paints. However, this study does not use any image processing technique to verify the content of the image. The proposed approach uses a set of voters to select their rank favourited images in a dataset. These set of voters will be used to create a stereotype mechanism to generate a recommendation based on voters profile. In this case, it is costly to develop the recommendations system because it is necessary a right amount of voters to create accurate recommendations. It is the way to represent the recommendation problem as a social choice it is essential completely ignoring the contents from the multimedia dataset.

The way to avoid the most usual problem in the content-based approaches is the use of hybrids methods to model and create recommendations for users. In the *SigTur*(MORENO et al., 2013) system, the profile map algorithm uses a different kind of data such as demographics, travel motivations, actions of the user in the system, rating about previous trips and options of users with similar tastes (collaborative filtering). The system uses these different sources of data to order accurate modelling about the user's feelings to avoid recommendation errors. In this case, the hybrid approach is composed of collaborative filtering and demographic and content-based recommendations. The three modelling and recommendation approaches can secure a reasonable preference inference about the taste since the user can provide such amount of information. The most relevant *SigTur* differential is to use an ontology domain to guide the recommendation in a relationship between profiles and attractions. Without the concept of ontologies, it would be impossible to establish a connection between so many different information sources and a particular attraction to be recommended.

There are some approaches to create a set of recommendations when the target are groups. The *e-Touri Tool*(GARCíA et al., 2009) tests the three classical approaches to create group models and recommendation in the tourism context. The study extracts the user's features from demographic data and the taste extract using a form. The study proposes the Generalist Recommender System Kernel (GRSK). GRSK is a domain-independent taxonomy-driven search engine that manages the group recommendation. Based on the features extracted in the first step, the tools can create recommendations using aggregation, intersection and incremental intersection methods. The GRSK model creates *interest degree* for each attraction for each user. Them, using the modelling methods, it is possible to create one single list of attraction with a group interest degree. The method to create the groups is used in this proposed study. However, we can use a different source of data to map the user's preference, avoiding the limitation of demographic and form source data.

The significant difference between our proposed work is the ability to use implicit content (photos from social media) to create the user map of preferences profiles. It is noticed that most of the approaches require significant effort on the part of the users to infer their preferences. The method proposed in this study requires much less from users, making the process much more agile and simple. Also, it is noticeable that the profile generated by fuzzy inference is not stereotyped as seen in more classical approaches in the literature. In this way, our study is able to detect the implicit preferences of users in a simple way without losing the personalization of each individual.

4 Tourism Recommendation Approach

The amount of information available on social media sites and its number of users have experienced an enormous increase in the last decade. All this information may be particularly useful for those users who plan to visit an unknown destination (BORRàS; MORENO; VALLS, 2014). According to Litvin et al. (LITVIN; GOLDSMITH; PAN, 2008), this information is also a substantial source of strategic information which can be used for the development of a number of business strategies, including enhancing visitor satisfaction through product improvement, solving visitor problems, discovering visitors' experience, analyzing competitive strategy as well as monitoring image and reputation of a tourism destination. In this context, this work uses photos gathered from social media sites to detect implicit tourist preferences as a way to support a smart tourism initiative.

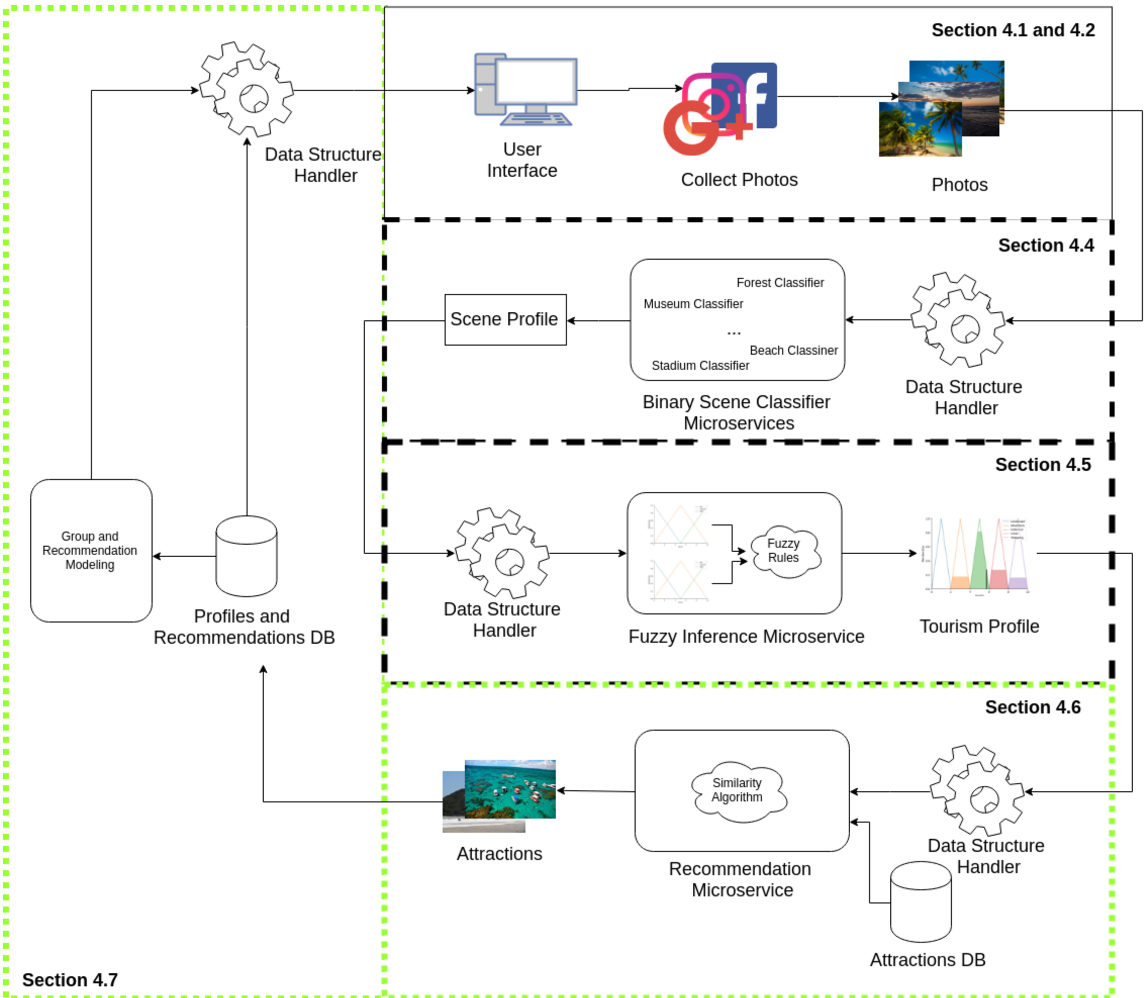


Figure 6 – Proposed multi-tourist recommendation approach

This study is the continuation of the research carried out in a previous work(FIGUEREDO, 2018). Figure 6 describes the flow of the approach reported in this text and his several changes concerning the first study (FIGUEREDO, 2018). Steps highlighted with black dotted lines in Figure 6 were improved at several points compared to the work cited. The modules highlighted with green dotted lines are brand new and have were tested in this new research. This chapter presents the steps for our tourism preference detection and recommendation approach for single and group of users.

4.1 User Interface

The user interface showed in Figure 7, is a tourism web portal that provides information regarding accommodation, Point Of Interest (POI), events, and several services (transportation, car rental, etc.). This information is gathered from the Tourism Information System. The user interface is a WEB based application developed using the Spring framework. Spring framework offers some modules that were incorporated into our solution. To cope with the significant number of users and with a large amount of data (photos) gathered from each user, we have defined a microservice architecture around short-lived processes.

A reasonable number of photos must be given for the system to be able to create a good scenario profile. However, there are problems in using pictures from social media in this application. It is common for many profiles to have a set of photos not related to tourism scenarios. The images that contain only texts and words are not relevant in our condition. In this case, this problem is avoided through the interface, where the user is asked to select favourite photos of travels, places or moments that enjoyed to be.

4.2 Data Extraction

The first set of microservices is responsible for data extraction. The data extraction comprises two steps. First, the user performs the log in the user interface using one account/password from one of the three most relevant social media or users devices. The web interface component uses the permissions given through the authentication process to gather the user's photos. The second step comprises starting the respective microservice to download the picture and store in a shared database.

The interface to select photos from social media or device is showed in Figure 7. The user has a limit of 100 most recent images from his profile to select the most interesting photos related to trips. This limitation helps to filter only recent photos, avoid pictures that express past preferences about the user.

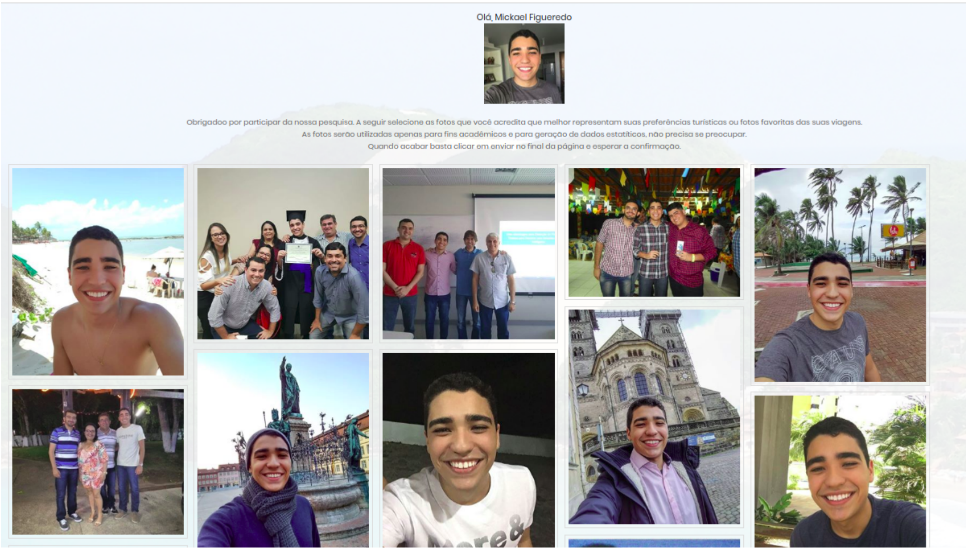


Figure 7 – User interface to select photos from social media

4.3 Tourist Classification

In general, tourists are classified according to their needs and their reasons for travelling. The tourism industry tries to meet the needs of all types of tourist by providing suitable facilities and services for each category. Moreover, marketers and planners, as well as managers of tourism businesses, consider these categories to guide their marketing, planning, development and management functions. Hence, find the right group for each tourist is essential functionality for any smart tourism recommendation. For instance, a report showed that 82,5% of business tourists stayed in hotels and resorts during their trips and 56,3% of leisure tourists chose this type of accommodation either. For tourists coming for other reasons, the kind of housing tends to be a house of friends, relatives or even rented. These figures show that the motivation and purpose of travelling impact directly on the type of services and local facilities chosen by the visitors.

Unfortunately, the tourism literature about traveller or tourist typology is scarce and sometimes confused. According to Lohmann and Panosso Netto (LOHMANN, 2017) "a great number of traveller typology has been developed, and the most notable ones include those proposed by Gray (1970), Cohen (1972), Smith (1977) and Wickens (2002)". The first three authors consider that the classifications proposed work for all types of destinations. Wickens (WICKENS, 2002) proposed a different typology, in which tourists are catalogued according to a particular tourist destination. Lohmann and Panosso Netto (LOHMANN, 2017) argue that those classifications have considerable limitations. The vast majority of these typologies were developed based on the European and American markets, and do not consider the characteristics of tourists from other regions, such as Latin America, Africa, Asia and the Middle East. Moreover, there are enormous cultural variations among different countries, which weakens their application more generally. In this context, this paper follows the approach provided by Andrade (ANDRADE, 2002) and Ignarra (IGNARRA,

1999) to classify tourists using the following typology: *Historical/Cultural*; *Adventure*; *Urban*; *Shopping*; and *Landscape*.

In order to fit the tourist into the five listed categories, the process inside the classification Component includes two steps: (i) *Scenario Classifier*, that performs the tasks of feature extraction and scene classification over each photo, (ii) *Fuzzy Inference*, that delivers an integration of all scenes classification results and maps them into the topology classification of the tourist.

4.4 Scenario Classification

Scene recognition is one of the most challenging tasks in the context of image classification. Human's visual mechanism can easily classify scenes, environments and places using the background and the knowledge about the elements inside one image. However, such a simple task for the human visual system is not so simple to provide in the computational context. Computational methods are not able to understand the semantic meaning between elements and scene. While the image classification is one of the first skills learned by humans, this task is not perfect complete for computers.

The scenario classification step is the first related to profile analyses (Figure 6) and most important to create the user tourism preferences profile. The scenarios classifiers must be accurate enough for not propagate error in the further steps. The main goal of this step is to generate a scenario preference profile. This process is described in Figure 8. Each photo is sent individually to a set of the binary classifiers. The binary classifiers approach was used to support a scalable and multilabel classification. Instead of using a single classifier, the adopted strategy can classify one image as belonging to several classes using a true or false output. Another advantage of choosing binary classifiers in this study is the scalability given by this approach. New classes can be added to the process less costly. 25 binary classifiers were created using a deep learning architecture. Table 1 shows the list of all classifiers.

The proposed study uses CNN to classify images. The convolutional architecture used was the GoogLeNet (SZEGEDY et al., 2014). The GoogLeNet is a famous deep learning architecture composed by more than 100 different kinds of layers such as Pooling, Convolutional and Dropout. A significant advantage of the GoogLeNet is the capacity to reduce the computational cost even though using parallel computing approaches. However, the essential feature of this network is the available transfer learning techniques. The CNNs have been around for a long time, but only now have they been driven by the advancement of transfer learning. *Transfer Learning* is the process of using a pre-trained model that was trained using one huge dataset of images, and retrain using your own dataset to refine the network.

In this work, we kept the convolutional structure from the GoogLeNet for all 25

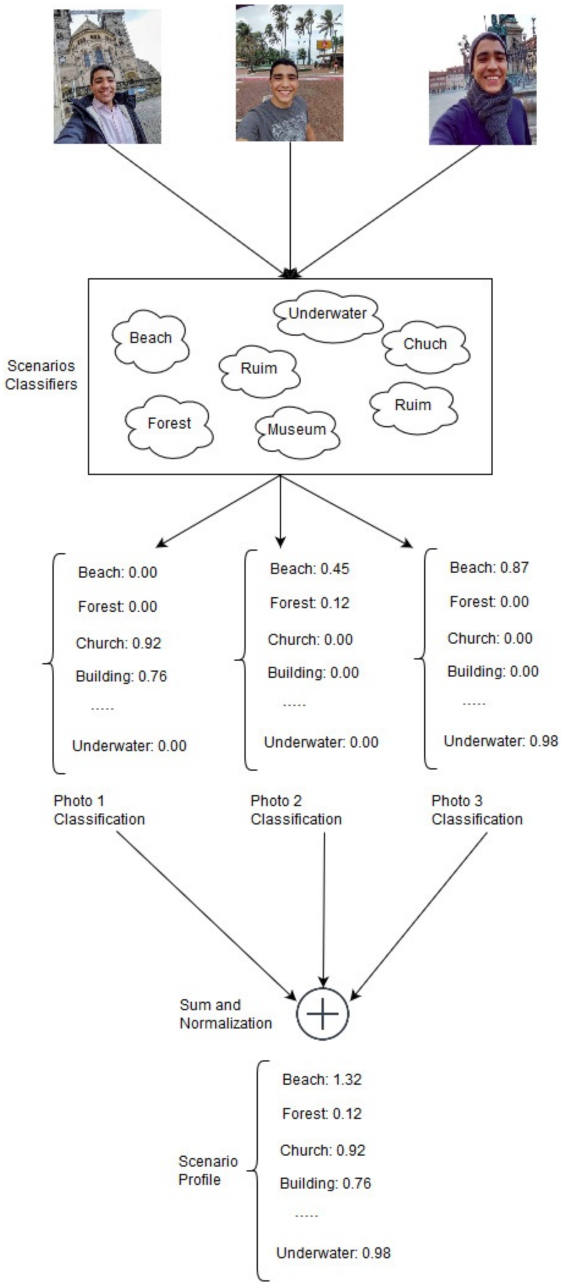


Figure 8 – Process to generate the scenario profile

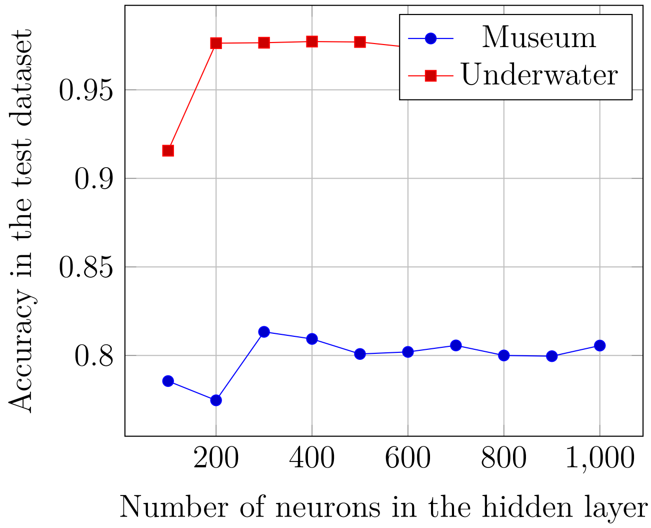


Figure 9 – Different fully connected architectures for the Museum and Underwater classes

classifiers. However, the fully connected network architecture in the final layer is changed to improve the accuracy of our application. It means that there are 25 different neural networks for classification but only one convolutional network architecture that acts as a feature extractor.

Features for each image were extracted using pre-trained GoogLeNet weights available in Keras library (CHOLLET et al., 2015). The dataset used in the training step was the Places 365 dataset. Seven thousand images were used in the training step for each classifier, resulting in a total of 175000 images to train the 25 classifiers. In this case, 3500 images represent positives images, while 3500 images are negative. The dataset of negative images is composed of 145 images from the 24 other classes.

Every fully connected part from our CNNs is composed of one hidden layer with a dropout layer using sigmoid activation function and one output neuron. However, there is a variance in the number of neurons for each neural network in the hidden layer. For each class, a range from 100 to 1000 neurons was tested to find the less expensive architecture with better accuracy. In the example described in the Graph 9, we can see that the best fully connected architecture for the Museum class is composed of 300 neurons in the hidden layer. For the Underwater, the best neural network uses 200 neurons. In both cases, there are saturation points. However, for some scenarios, the saturation point in the minimal architecture, consisting of 100 neurons. In the output from each neural network, there is one only output neuron because we are using a binary classifier approach.

As a consequence of use only one neuron in the network output, it was necessary to find a threshold that defines a classification output as true or false. For example, for the museum class, it is essential to determine what value in the output classify one image as a museum or not. In this way, a set of 200 images was classified for each class and the graph showed in Figure 10 was created. One hundred images in the collection are negative, and one hundred images are positive. The intersection between the true and false curves in

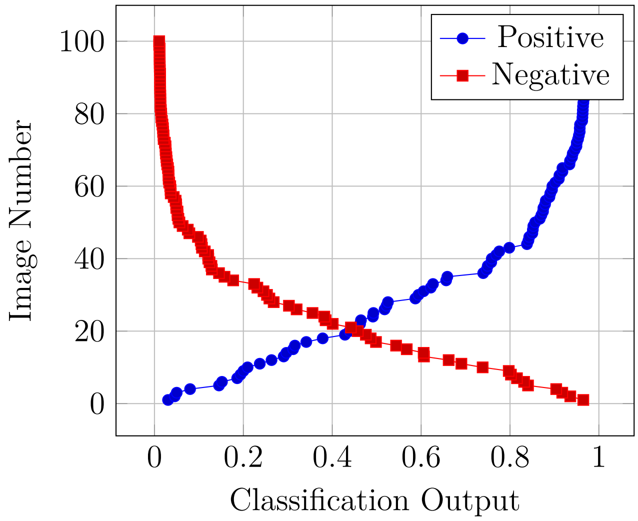


Figure 10 – Threshold analyses for museum scene class

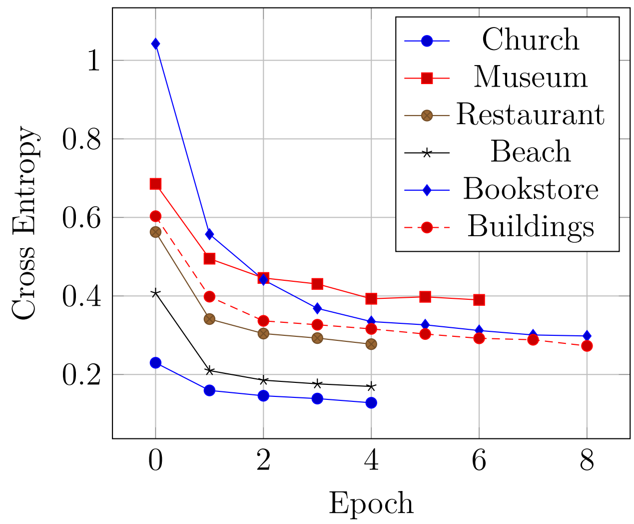


Figure 11 – Cross entropy to support early stop training

the graph from Figure 10 is the ideal value output where there is a balance between false positives and false negatives. So, the threshold used for the museum class is 0.42. This process is repeated for each scenario classifier.

After deciding the best architecture for each neural network classifier, the training cost details were analyzed. Keras API (CHOLLET et al., 2015) was used to support the *early stop* in the training process. This approach uses cross-entropy metrics at validation and training dataset to detect the decrease in learning capac2ity and handle early stop training. It reduces the probability of overfitting and expensive training cost. Max number of epoch in this training process was 15. Figure 11 express some early stop cases. Scenes easily-learned has a low training cost concluding the process in few epochs such as *Church* and *Beach* cases. For these classes, the number of epochs necessary to detect the neural network saturation was 4. However, some classes with worse results take more epochs to finish the learning step.

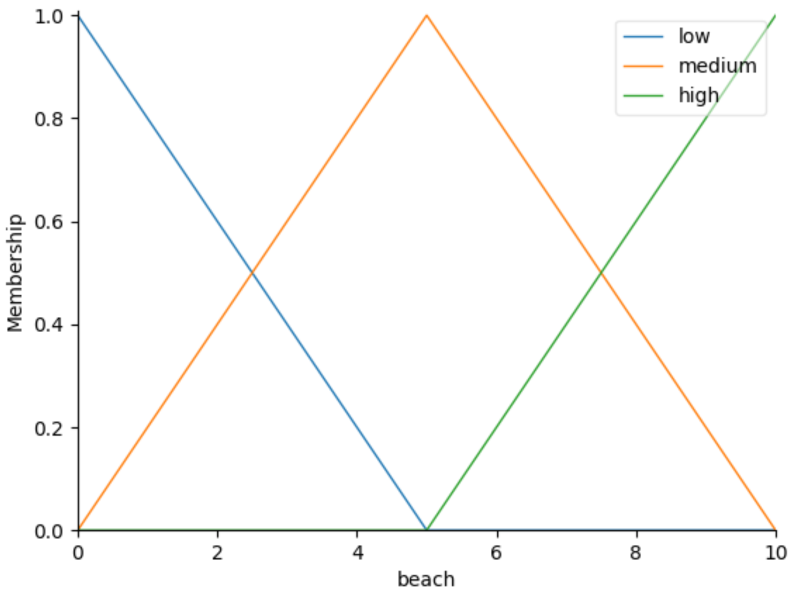


Figure 12 – Fuzzy antecedent for beach scenario.

The pipeline showed in Figure 8 is closed using the best architecture found for each class through the described methods. Each photo has its features extracted by a single convolutional architecture. Subsequently, features are used by the 25 classifiers to identify the pictures scenarios. Each classifier rates an image with a value between 0 or 1. However, values are set to 0 when lower than the threshold set for the class. If the output value of the classifier is higher than the threshold, then the output value will be kept in the output vector.

The next step is creating the complete scenario profile based on the entire set of photos from the user's social media. Each classified photo is summed to generate a single profile. The output values of the neurons are directly summed in the final scenario profile vector.

4.5 Fuzzy inference

A fuzzy approach was adopted to classify the tourist into the five classes of tourism described in Section 4.3. After the user photos analyses one vector containing 25 elements is created as showed in Figure 8. These features are the basis for the fuzzy rules. The main goal of the fuzzy inference is to infer which of the five tourism classes are relevant for the user based on the scenarios. The fuzzy approach will generate a human logic-based preferences map. The idea is relating the tourism classes with the scenes. In this case, what is essential for us to see the membership value for each tourism class. Tourism classes containing high values of memberships are considered relevant for us.

In this way, the first step in the fuzzy inference was understood the relations

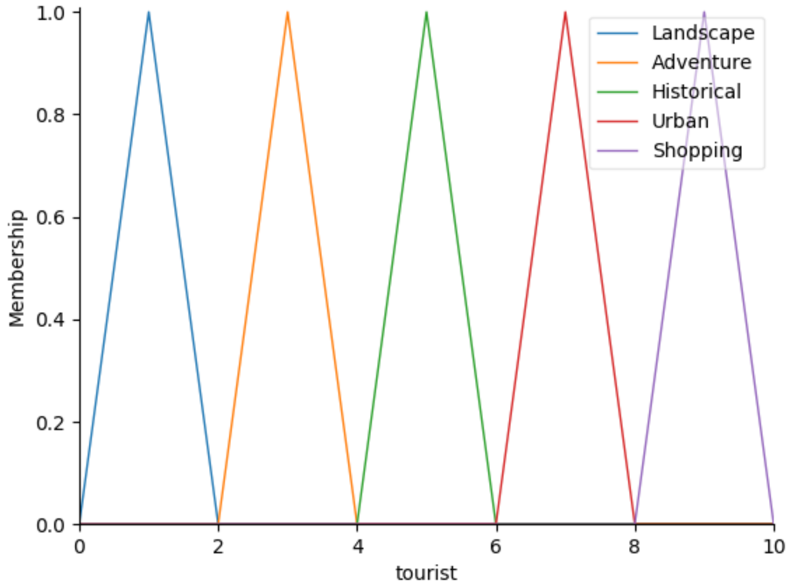


Figure 13 – Fuzzy consequence.

between scenarios and the tourism classes defined in the literature. For this, with the collaboration of tourism experts, each scenario was related to one or more tourism classes. For example, the Museum class is strongly related to *historical and cultural* tourism but is also related to the *urban* tourism, however with lower relevance. There are various situations relating one scenario with more than two classes of tourism. In the case of mountain scenario, the relations are strong with *natural* and *adventure* classes, weak with the *sport* tourism classes and is irrelevant for *shopping* and *urban* classes. The relations groups can be seen in Table 1. As it is possible to see, some scenarios established relationships with more than one class of tourism, according to experts in the area. These relationships are essential to create an accurate map about the tourist preferences using the fuzzy logic.

There are 25 antecedents (fuzzy inputs) in the fuzzy configuration. The inputs define how frequently one user has been the specific environment. The output vector from the scenario classification is our fuzzy input. The fuzzy antecedents are divided into three pieces related to the frequency: low, medium and high. The frequency has the scale 0 to 10 (Figure 12). In this case, when the scenario classification vector contains one or more scenarios with the value near 10, it means that the user usually frequents that scenario. Triangular functions were used to define *low, high and medium* frequency in one scenario. For example: If a user has in his scenario classification vector one value of 2.3 for the class beach, the fuzzy input is between low and medium. The intersections in the fuzzy antecedent functions are important to create a fuzzy idea in the process.

The fuzzy classifier consequent (Figure 13), or output, is the relation between the user scenario profile and the five tourism classes. Typically, output expected for fuzzy applications is the membership centroid. In our case, we are interested on the membership value for each category. The recommendation application context motives it. The system

Table 1 – Relations between tourism classes and scenarios.

Scenario Class	Landscape	Urban	Historical/Cultural	Shopping	Sport
Airport		Strong		Weak	
Athletic Field					Strong
Beach	Strong				Weak
Bookstore		Weak	Weak	Strong	
Building		Strong			
Candy Store			Weak	Strong	
Church		Weak	Strong		
Construction		Strong			
Department Store		Weak		Strong	
Forest	Strong				Strong
Gift Store		Weak			
Golf Field	Weak				Strong
Hockey Arena		Weak			Strong
Mountain	Strong				Strong
Museum		Strong	Strong		
Neighborhood		Strong			
Ruin	Weak	Strong	Strong		Weak
Restaurant		Weak	Strong		
Science Museum		Weak	Strong		
Shopping		Strong			Strong
Snow Field	Strong				
Football Stadium		Strong		Weak	Strong
Street		Strong			
Under Water	Strong				Strong

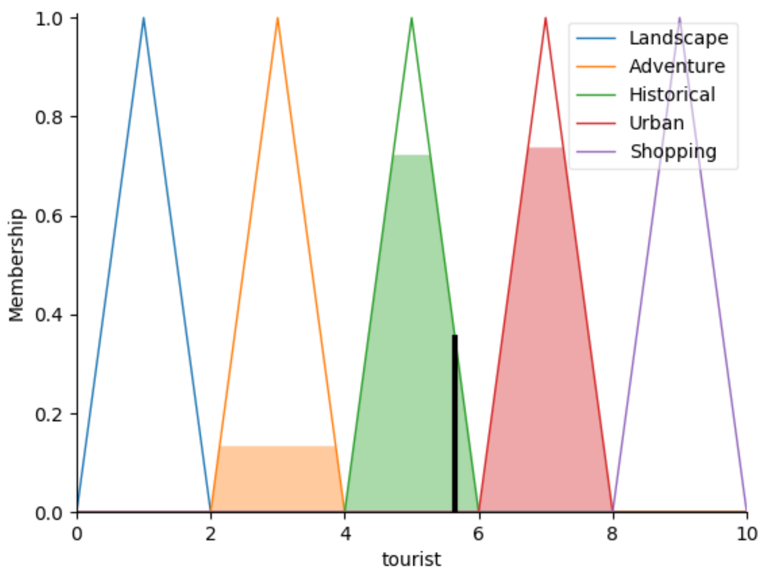


Figure 14 – Fuzzy output for one example profile

seeks to recommend attractions based on user preferences for each tourism class. The centroid classification is not interesting for us because we are not interested in finding an adequate kind of tourism for one user, but understand how much each type of tourism is relevant for him. Using the membership value (the y axis in Figure 14) the system can match different Points of Interests if they are in the same universe of the fuzzy output. The relation between the five classes in the fuzzy approach is shown in Figure 13. There is no intersection between the five types because each class should be independent in the process.

The fuzzy rules attempt to simulate general human behaviour. If a user contains many pictures of beaches, forests and mountains them the *Landscape* class hold a high pertinence in the classification process. The entire list of fuzzy rules can be seen in Appendix A. There are ten fuzzy rules to create a map of preferences that tries to cover the tourist experts knowledge.

Some approaches (JANG, 1993; JOZI et al., 2016) tries to create the fuzzy rules for these systems automatically. This kind of approach can be useful to create accurate fuzzy rules for one fuzzy classification systems. However, in this study, it was not possible to develop our fuzzy rules automatically. First of all, there was insufficient data to train our fuzzy inference module. The entire data collected about the fuzzy inference system was used to validate our approach. Also, a classic problem of automatically creating fuzzy rules could be a problem for system maintenance. The automatic rule-making approach creates a fuzzy rule for each training case. In this way, our fuzzy inference module that has ten fuzzification rules could have a much higher number of rules, which would make it difficult to maintain the system. In this way, it was decided to use only the logic of creation of rules using the knowledge of the tourism experts.

4.6 Recommendation

As described in section 2.3, in the recommendation process, all the approaches described have one point in common: the use of a similarity algorithm. The similarity algorithm seeks to relate users to other users, items with other items, or even users and items. The proposed approach aims to related user with Points of Interest(PoIs). The cosine similarity was applied for this. The cosine similarity seeks to relate the vector of each attraction to the user profile vector to find the most relevant attractions. Application of this algorithm in the approach is described below.

4.6.1 Similarity Application

The output from the fuzzy inference step is a vector containing five features relating the user with the five classes of tourism. The recommendation step connects this output vector with the attractions. First, the attractions database was described in the

Table 2 – Example of attractions modelated in the same universe of fuzzy inference output.

Attraction	Landscape	Historical	Urban	Adventure	Shopping
Midway Mall	0.0	0.0	1.0	0.0	1.0
Ponta Negra Beach	1.0	0.25	0.75	0.0	0.0
Matriz Church	0.0	1.0	0.75	0.0	0.0
Arena das Dunas Stadium	0.0	0.25	0.75	1.0	0.0

same universe of attraction. With the collaboration of the tourism experts each attraction had its membership value inferred for the classes of landscape, historical, urban, adventure and shopping. This situation is shown in Table 2. Some attractions from the database are described in the five classes of tourism. The dataset reflects the option of two tourism experts.

In this way, the vectors or points are inserted in the same profile dimension, and can thus have their similarity compared by a metric.

For a user u and an attraction p the cossine similarity can be calculated using the Formula 4.1. As described in section 2.4, the focus of this metric is the vector relation of attraction and user. In this case, the magnitude of the vector is not relevant as its direction. Importantly, the formula used in our application uses the vector module which avoids negative values. Moreover, all vector values can only be positive which means that only the first quadrant of the universe's Cartesian plane is used.

$$similar(u, p) = cos(\vec{u}, \vec{p}) = \frac{\vec{u} \cdot \vec{p}}{|| \; u \; || * || \; p \; ||} \tag{4.1}$$

All database attractions have the cosine similarity calculated for each user. The output of each iteration is a value between 0 and 1, where the higher the value the more similar the attraction is to the tourist. Thus, the output vector of this step is the descending database based on cosine similarity.

4.7 Group Modelling

In the proposed study, the system could identify the agreement among the group in two different points of the recommendation process: The fuzzy tourism profile generation and in the recommendation listing. In this research, the second approach was used. Create the group recommendation using the list is more efficient in terms of systems, as it would avoid rework to create new profiles of preference. In this case, a tourism attraction is only considered important or relevant if it is present in all individuals lists. The final pipeline uses the Incremental Intersection as the group modelling approach. The group modelling is tested and validated in previous works(GARCíA et al., 2009) (CHRISTENSEN; SCHIAFFINO, 2011) and was implemented in this pipeline as described below.

Table 3 – Example of application of Incremental Intersection.

Attraction \User	User 1	User 2	User 3	C	M	S
A	0.0	0.5	0.2	2	0.7	1.4
B	1.0	0.0	0.0	1	1.0	1.0
C	0.3	0.3	0.3	3	0.9	2.7
D	1.0	0.0	0.5	2	1.5	3.0
E	0.0	0.0	0.1	1	0.1	0.1

4.7.1 Incremental Intersection

In this implemented intersection model, the intersection has a variation that makes it more exciting and is capable of generating significant changes in the results. In this case, the situation where a recommendation list may be empty is avoided by a straightforward approach.

The elements have their repetitions counted so that recurring elements can have their final scores more relevant. In the case of the application of the algorithm in the list of recommendations $L =< l^1, l^2, ..., l^n >$, each attraction in l^n will have a multiplicative effect on the score according to the total number of appearances in list L. In cases where an attraction appears in only one list, instead of being discarded (the process at the standard intersection), it only receives a *punishment* for not being relevant. However, attractions that appear in all groups, even with low scores in all, will be more consistent in the process. However, in our context it is necessary define a value for n. The top 6 attractions for each user are used in each user's list. In this way, the intersection is made in the context of the top 6 attractions for each user with $n = 6$. The value of n may vary depending on the application of the approach, the duration of the trip or the number of people in the group. Empirically a value of 6 was used because a very long list of attractions could make the application testing process tiring.

The number of appearances C of one attraction in the top-6 list of attractions for each user is multiplied by the summed attraction score for all users. In this way, different from the intersection and aggregation, the score of each attraction for the users is not discarded and used as a way to measure the relevance of the attraction for group. The S score of an a attraction in a group recommendation is given by $C * M$ where M is the summed attraction score for users. One sample example of this approach application in our context is showed in Table 3. The idea is that attractions that are frequent in the top 8 are recommended even there are not the bests. The idea is that to please all users is average rather is better than please only a few users very much. This becomes clear when comparing attraction C and D in Table 3. The attraction that was present for all users with a reasonable value was much more interesting than an attraction that was very relevant only to one user.

The same idea of the intersection of numerical values is applied so that there can exist the idea of improvement of relevant tourism classes for all users. It is noticed that attractions that are recurring to all users are benefited by the repetition factor while attractions that have high scoring but are recommended for few users suffer a penalty spree.

5 Evaluation

In this chapter we present an evaluation of our approach. Our evaluation is comprises in three modules analyses: (I) Scenario Classifier, where we decided what kind of classifier should be used to feed our fuzzy algorithm, (II) Fuzzy inference, where we try to understand the level of accuracy of our preference detection algorithm and (III) Recommendation, where it is decided the best similarity algorithm to use in the recommendation step.

5.1 Scenario Classifier

The classification approaches were confronted focused on three faces: training cost, classification time and accuracy. The classification time is important because this classification will happen while an user interact with one interface, so it is necessary that the images classification be faster enough to not generate overhead for users. The cost of training needs to be evaluated for future system implementations. If a new class of environment needs to be classified, based on the binary system of classifiers, it is necessary that the classification method supports a training time fast enough not to harm the system environment. However, the most relevant metric in the evaluation step is the accuracy for the classifiers.

Each binary classifier was validate individually using the *K-fold cross validation* to evaluate the accuracy metric. Both approaches (Deep Learning and BOVW) were trained using the same number of images. The K value used in this study uses $K = 10$ based

Table 4 – Validation approach for each step.

Tourism Recommendation Step	**Section**	**Validation Approach**
Scenario Classifier	5.1	Performance analyses and accuracy verification using 10-Fold Cross Validation of Deep Learning and Bag of Visual Words classifier.
Fuzzy Inference	5.2	Comparative between human inference and fuzzy approach using 4 test profiles and users self analyses through mobile applications.
Recommendation	5.3	Collect 32 users recommendation preference using mobile app based on the user self analyses.

Table 5 – Training and classification time for classifiers

Classifier	Training Time (Hours)	Classification Time (μS)
Bag Of Visual Worlds	2.6	9.8
Deep Learning	0.8	350.2

Figure 15 – Gap between images from construction scenario

the study of Kohavi (KOHAVI, 1995) where the basic idea is that lower K is usually cheaper and more biased. Larger K is more expensive, less biased, but can suffer from large variability. This is often cited with the conclusion to use $K = 10$.

The 10-fold cross validation approach trains the classifier with a set of images from the 7000 total, and validate with another part. The process is repeated 10 times creating different sets of training and validation. In this way, a high accuracy in 10-fold cross validation represents a classifier without overfitting and able to generalize. All tests were produced on a computer without a graphics processing unit (GPU), processor with 4 processing cores, 4 gigabytes of RAM and using librarians who collaborated to use all cores of the processors in parallel.

5.1.1 Performance Analyses

The first step in analysing which scenario classifier the system should use in this study was the training cost and classification time. The times shown for Table 5 related to the training time are for a single class, since all classes have equal training time because they use the same amount of images. For the classification time, the result is related to one single image classification cost. In our case the classification time should be more relevant, since the profile of the tourist will be generated the moment he will send the photos to the servers. In this way it is important that the classification of the images is fast enough not to represent an overhead. The result obtained and seen in the Table 5, show that there is a great difference between classifiers in terms of classification time, however, in our context where only few images will be processing in real time, the time of classification is small since it is in microseconds for both classifiers. This means that the model adopted does not depend on this analysis, since the results are very similar. However, it can be seen that since using a series of *tricks* such as knowledge transfer and early stop training the Deep Learning classifier is more faster than Bag Of Visual Words. The time spent in training shows that adopting an SVM as the classifier makes BOVW training very expensive. However, even using a feature extractor slower (SIFT algorithm), the Bag Of Visual Words approaches seems faster in real time classification. The results could be improved if a faster point descriptor were used, such as SURF (BAY et al., 2008b) and FREAK (ALAHI; ORTIZ; VANDERGHEYNST, 2012).

From the results, we can conclude that the deep learning model extracts features more slowly than BOVW, however its training is much faster using an MLP instead of an SVM. In this way, due to the small difference in the time of classification in real time, we can say that the approach based on deep learning is more efficient due to the less time spent in the training.

5.1.2 Accuracy Analysis

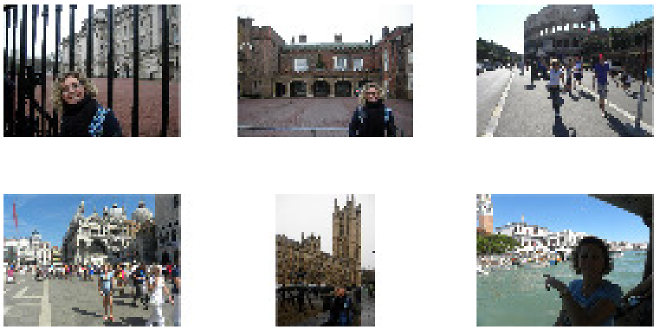
Finally, in the accuracy analyses for Bag of Visual Words and Deep Learning interesting results were achieved. The transfer learning approach used in conjunction with a standard fully connected layer resulted in accuracy scores always above 0.85. However, it can be seen that the approach to deep learning has resulted, in general, in accuracy always above 0.90. In other hand, the Bag of Visual Word classifier has most of the classifiers ranging in the range from 0.65 to 0.70 of accuracy. However, some classifiers achieved good results, such as the *Forest* (0.89), *Beach* (0.80) and *Underwater*(0.80). As can be seen, natural scenarios were better classified generally by the two classifiers. This is due to the fact that these types of environments have a very definite feature that differs from the others. It is noticed that indoors environments where the elements that identify them are small details as in the classes related to the stores. The best results achieved in these types of environment were for the Bookstore class, where 0.79 of accuracy was achieved in the classification method using BoVW and 0.91 using deep learning. One of the most surprising results between the two methods is related to the *Constructions* scenario class. In this case 30% of accuracy separates the approach based on the extraction of features using SIFT and the modern one. This is due to the great difference between the images

Table 6 – Accuracy for each binary scenario classifier using 10-Fold cross validation

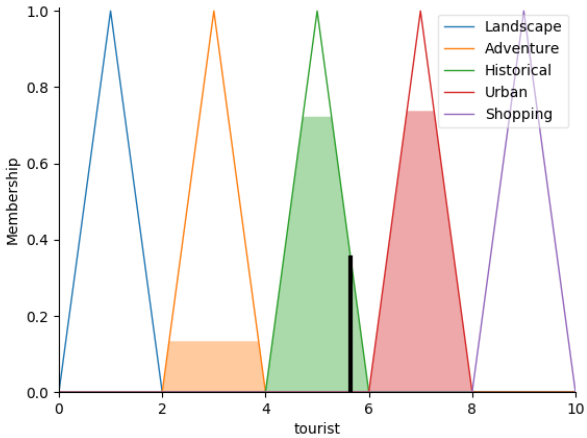
Scenario Class	Bag of Visual Words	Deep Learning
Airport	0.63	0.89
Athletic Field	0.66	0.92
Beach	0.80	0.93
Bookstore	0.79	0.91
Building	0.68	0.88
Candy Store	0.76	0.90
Church	0.67	0.95
Construction	0.57	0.87
Department Store	0.67	0.84
Forest	0.89	0.96
Gift Store	0.71	0.82
Golf Field	0.75	0.92
Hockey Arena	0.75	0.96
Mountain	0.78	0.93
Museum	0.62	0.81
Neighborhood	0.69	0.92
Ruin	0.66	0.86
Restaurant	0.66	0.86
Science Museum	0.71	0.83
Shopping	0.61	0.87
Snow Field	0.78	0.93
Soccer Field	0.64	0.93
Football Stadium	0.70	0.92
Street	0.65	0.91
Under Water	0.80	0.98

that compose the class. One difference can be seen in the image that the gaps between the class images are relevant, thus requiring a more powerful model for an acceptable results.

In this way, it is possible concluded that the *Deep Learning* is an approach less expensive in the training step using many techniques to reduce this process, such as transfer learning and early stop. Furthermore, as discussed using the Table 5, is possible affirm there is a difference in time for classification between the two approaches. The classification time can be problem for the deep learning approach if the system works with large ammount of photos in a single process. However, a big difference can be seen in the most relevant metric. In the accuracy, no classifier based on Bag of Visual Words was able to overcome Deep Learning. In this way, the scenario classifier used in this system was a Convolutional Neural Network based in the Inception Model from GoogLeNet.



(a) Photos for profile test 1



(b) Fuzzy inference for profile test 1

Figure 16 – Profile 1

5.2 Fuzzy Inference

The tourism inference system was validated in two different ways to confirm the accuracy of the approach suggested in this study. The fuzzy step has a very clear purpose: detect the classes of tourism most relevant for the users based on the scene profile. The first validation method seeks to verify how the approach follows a human logic based on the tourism experts knowledge. In the second validation method, it seeks to verify if the users feel represented by the inference made by the fuzzy module. First we describe how the data for analyzing the results were collected and then how they were analyzed.

5.2.1 Collecting Profiles

It was necessary to create a set of test profiles for the validation method that seeks to verify the machine's ability to follow the human sense through fuzzy logic. These profiles seek to be as heterogeneous as possible, aiming to test the approach in different situations. In this way, 4 profiles of tourists were created from volunteers containing different amounts of images. The variance in the number of photos is used to check the differences in results in relation to the number of photos in the profiles.

The first profile, seen in the Figure 16a, is composed of photos in historic center of Europe. Most of the photos were taken in urban centers that have great historical context, such as the photo located in the coliseum in Rome. However, one of the photos was taken in an urban region that contains in its composition a river. The inference generated by our approach, seen in the Figure 16b, assumes that the tourist enjoys to visit historic and urban scenarios. This inference is maintained even with natural environment photos in profile. In this way, due to the presence of element of ruins in some photos, the Adventure, Landscape and Shopping tourism classes received a small relevance for profile 1, based in the relation described in the Section 5.

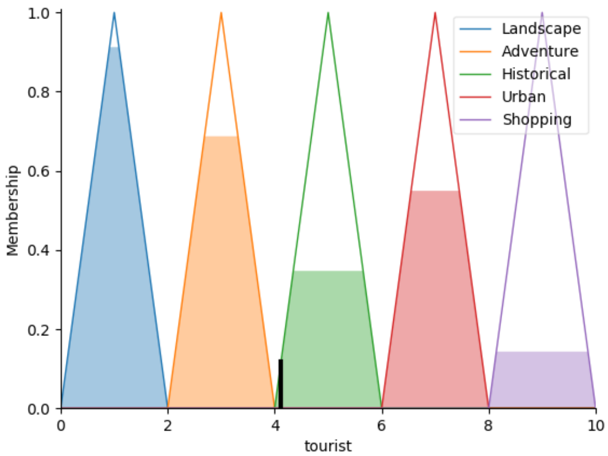
In the second profile collected, a situation more focused on natural environments was generated. The test profile is composed of strong elements of beaches, forests, mountains and aquatic environments. Only a small remnant of an unnatural environment is present, in the photo from the interior of a church. The inference generated for this user, proposes that he has very strong preferences for the landscape class, related to nature, with a reasonable preference for Sports and Adventure. This inference was generated, as expected, due to the strong presence of scenario classes related to these natural environment. However, a small preference for historical environments was generated, probably due to the photo inside a church.

The third profile used as test can be considered the most mixed of all and can be seen in Figure 17a. The profile was built to challenge the fuzzy system in cases where the tourist has a good presence in scenarios related to all classes of tourism. In this profile there are photos in historic centers, natural environments such as beaches and snow, urban environments with shopping stores. Also, this profile is the one with the most amounts of photos. As a result, seen in the Figure 17b, we can see a greater relevance for all classes, if compared with the other test profiles. However, as expected due to the greater presence of photos in natural and adventurous environments, these two classes were the main ones. In addition, the good relevance of the urban class (red triangle) can be noted. However, all this was built without disregarding the classes of tourism that are said to be less relevant (Shopping and Cultural).

Finally, the last profile (Figure 18a), is a test profile that seeks to focus on classes related to historic urban and natural centers in a balanced way. The profile has many photos in European historical centers, as well as the profile of test 1. However, a greater presence of natural environments can be seen in a good proportion. In this way, the exit of the inference seen in the Figure 18b, generated an interesting result, defining a balance between the classes of tourism more related to these environments.

Another approach used to validate the extraction of tourism preferences from users profiles was through real-time feedback about the users own profile. In this way, another Android application was developed to improve the usability of this validation. In this step, the users are asked to send their photos for an analysis of the tourism profile.

(a) Photos for profile test 3



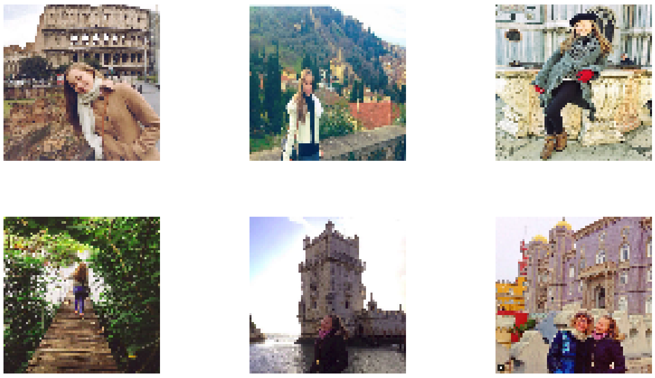
(b) Fuzzy inference for profile test 3

Figure 17 – Profile 3

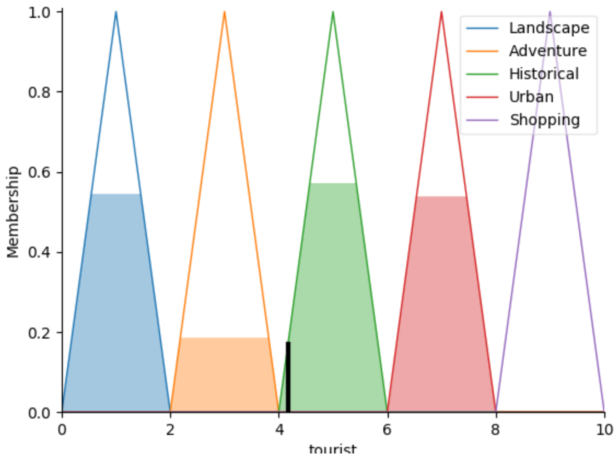
Users were asked to give priority to photos related to old trips or pictures taken in places that please them.

In this case, our application does not request photos from social media. Extract photos from the user smartphone was a solution found to avoid the limitations from socials medias APIs when it is used in non-production environment. After some scandals, such as Cabridge Analytics, there are restrictions about use the API to access users personal data. However, usually the same photos posted in social media are present in the users smart phones physical storage.

After select the photos, the users are able to upload their photos to the server to extract the tourism preferences and the recommendation. First the user is redirected to the screen showed in the Figure 19. This screen is used to collect the feedback about the tourism preference. In this way the validation approach could collect the sentiment from the user about the inference using six levels of agreement from 0 to 5. In this way the user is able to judge the profile created using the machine. This approach completes the validation pipeline of the fuzzy inference module. This way, users' opinions about third party profiles and their own profiles are evaluated as a way to demonstrate the model's ability to assimilate with the real world.



(a) Photos for profile test 4



(b) Fuzzy inference for profile test 4

Figure 18 – Profile 4

5.2.2 Evaluation

5.2.2.1 Using Test Profiles

All profiles were analysed using the same fuzzy logic considered most appropriate through empirical tests and based on relationships established by specialist in the field of tourism, which can be seen in section 5. For the Figures 16a, 17a, 17a the focus is on membership represented by the triangle, which gives us the degree of relationship of the profile with each class of tourism.

After collecting the profiles it was necessary to collect human opinions related to the four profiles. The idea was to create a ground truth to compare with the inference from the machine. In this case, human inference is our *ground truth*. An android interface, showed in the Figure 20, was developed to collect human inferences. Humans were selected randomly from the Federal University of Rio Grande do Norte. Sixty-four inferences were collected on the profiles. Each person gives certain membership values for each tourism class for each test profiles. Thus, at the end of the process, the *ground truth* for each profile is the mean of the values given by humans. Thus, if the algorithm proposed in this study reach the average generated through humans, we will have a good result.

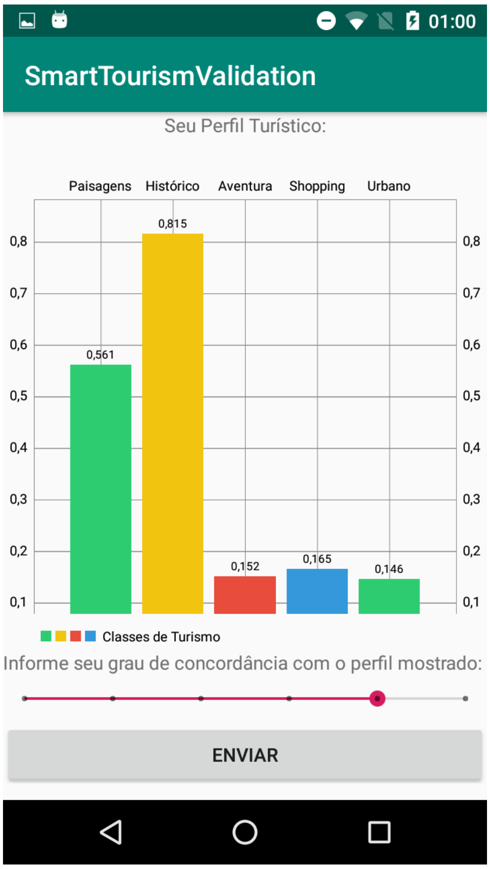


Figure 19 – Android interface to collect users agreement.

Figure 20 – Android interface to collect human inference

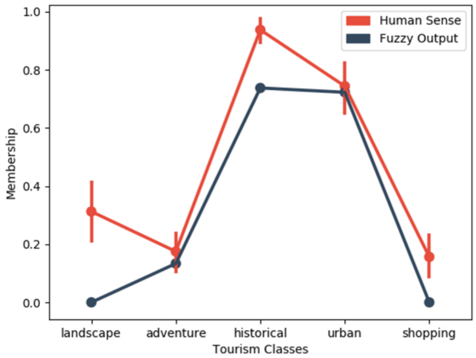
Using the ground truth and with the system classification available, it is possible verify the accuracy of the fuzzy inference. The first step was to define a metric. In this way, the Mean Absolute Error was used. In statistical, mean absolute error (MAE) is a measure of difference between two continuous variables. Assume X and Y are variables of paired observations that express the same phenomenon. In our case, what we want to calculate is the MAE of fuzzy inference compared to human inference. In this way, each class was compared individually and in the end the mean of the error is generated. The formula for the calculation used in this study given the fuzzy inference I and the average human inference collected H, is described as below:

$$MEA(I,H) = \frac{(H_l - I_l) + (H_a - I_a) + (H_h - I_h) + (H_u - I_u) + (H_s - I_s)}{5} \tag{5.1}$$

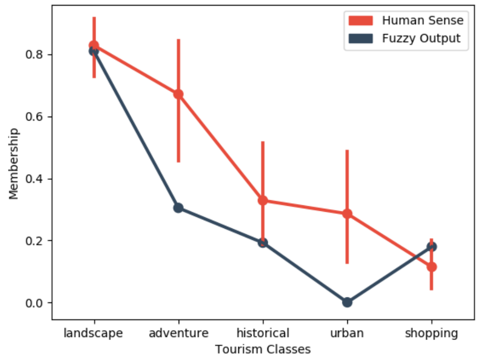
First, a visual comparative between the *ground truth* created using humans and the fuzzy output is showed in the Figure 21. The blue lines represent the same information shown in the Figures 16b, 17b and 18b but using points connected to demonstrate the overall result curve. The result collected from humans is shown on the red lines. However, in addition to this, the human inference curve show the variance of the data collected. In this way, points that have large vertical lines represent a bigger variance in the data collected. The purpose of this graph is to compare the human sense and fuzzy logic curves. Similar curves represent concordance between human and fuzzy output. In the comparison for profile 1, shown in the Figure 21a, a very similar curve between the two inferences is observed. It is possible to verify a perfect agreement between human and machine for Urban tourism class. The only notable divergence between the two inferences is noted for the Landscape class. Humans gave more importance to this class due mainly to the photo near a river seen in the Figure 16a. However, analyzing the results obtained using the MEA metric, we can see that profile 1 was the best modelled by the algorithm.

Another profile that obtained an efficient result can be seen in Figure 21b. Profile 3 has a very similar fuzzy output compared to human ground truth. The Landscape, Historical and Shopping classes were almost perfectly mapped. A small divergence is notable for the adventure class. However, we can see a relevant indecision in the human inference about this tourism class. This can be proved by the large variance in membership given by humans in this situation. Another notable divergence is for the Urban Tourism class. In this case, a difference of approximately 0.35 in this tourism class was generated. However, a low MEA value was generated for this profile, which proves the map generated efficiently on the profile 3. As seen in the Table 7, a value of 0.122 was generated from a maximum value of 1(worst case).

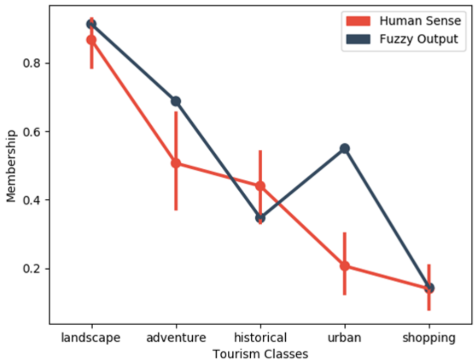
A case where the curve of the fuzzy output system follow the curve of the human sense, however not very accurately, can be seen in Figure 21d, related to profile 4. In this case, the MEA result seen in the Table 7, even been well, was one of the worst among



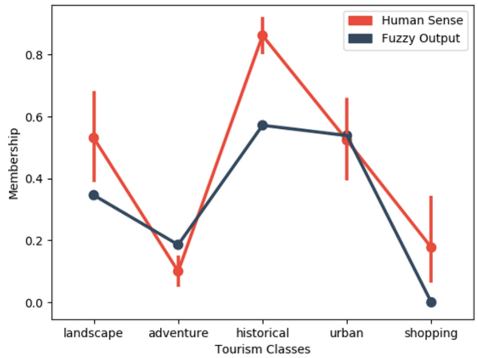
(a) Comparative for profile 1



(b) Comparative for profile 2



(c) Comparative for profile 3



(d) Comparative for profile 4

Figure 21 – Visual comparative between human sense and fuzzy output

Table 7 – Mean absolute error for fuzzy classifier

Profile	Photos Number	MAE
Profile 1	6	0.076
Profile 2	9	0.174
Profile 3	8	0.122
Profile 4	6	0.144

test profiles. However, analysing in a visual way through the Figure 21, we can confirm the quality of the output generated. The fuzzy system gave relevance to the same tourism classes that were considered relevant by humans. This means that if humans considered the Historic tourism class relevant, the machine also did the same. However, the result of the Table 7 can be justified by the difference in membership values. Taking the historical class as an example, we can see that the machine considered it relevant to the tourist. However, the value was far from that given by humans. In this way, we can conclude that the machine followed the human sense well, but not in an extremely precise way.

Finally, the profile that obtained the worst results was profile 2. In this case, the curve generated followed the human sense with some points of divergence. However, it is notorious that this profile was the one that generated the most indecision among humans,

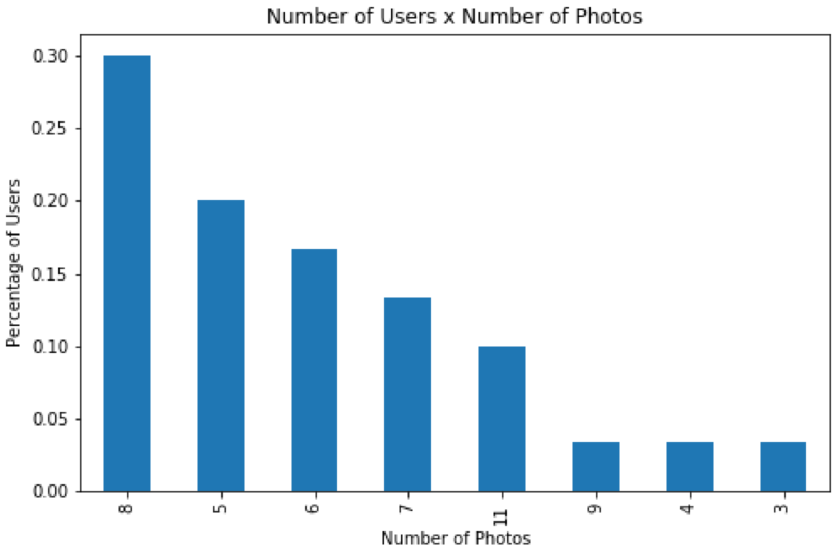


Figure 22 – Number of photos uploaded by 32 users

which can be verified by the vertical lines in the Figure 21b. In this way, we can say that the inference was good, since the mean absolute error was low (0.174 of a maximum value of 1), however the analysis was suffered with noise create by human indecision.

Information that proves the quality of the model created is the relationship between the mean absolute error and the number of photos adopted in the profiles. The increase in the number of photos is not directly or inversely related to the quality of the inference made by the machine. This shows that the model is robust enough to be used in a way that the user uses different amounts of photos to generate his profile in the system.

5.2.2.2 Collecting Users Feedback

Feedback from 32 volunteers was collected at this validation step. The volunteers were randomly selected to upload photos and judge the fuzzy inference generated. Users were asked to upload pictures related to happy times, places they like to be or old trips. No minimum or a maximum number of photos has been set.

As shown in Figure 22, users tended to send a photo number between 5 and 11. These numbers serve as the basis for the decision not to normalize the scene profile vector. With a quantity of 5 to 11 photos, it is unlikely that users would saturate the scene classification vector by 10. At least ten photos would have to occur, and all of them would be of the same type of scene. It proves that normalization in the output step of the scene classification module is unnecessary. However, the amount of photos uploaded by users directly interferes in the quality of the results seen in this approach. Users who used less than five photos for the test will not have a quality inference.

First, an analysis was made of the general tourism profile of the users who participated in the experiment (Figure 23). In a review based on the average value of

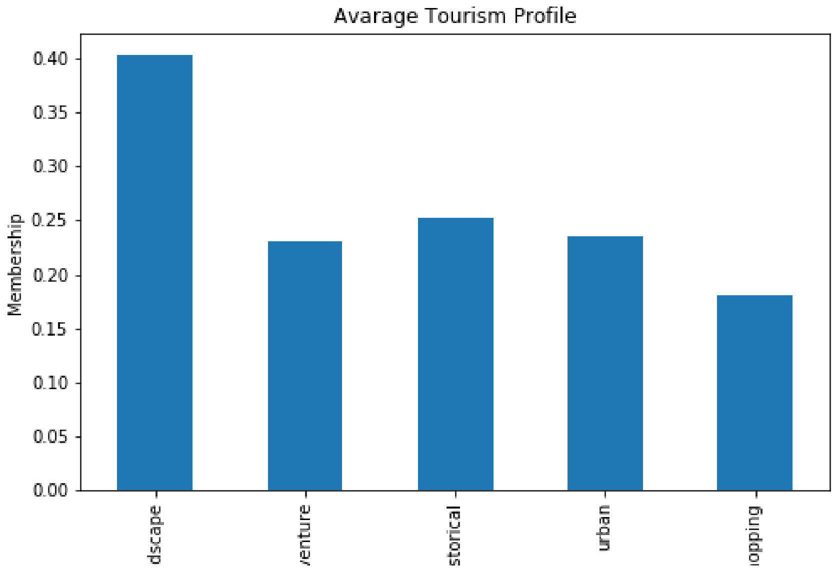


Figure 23 – Average tourism profile

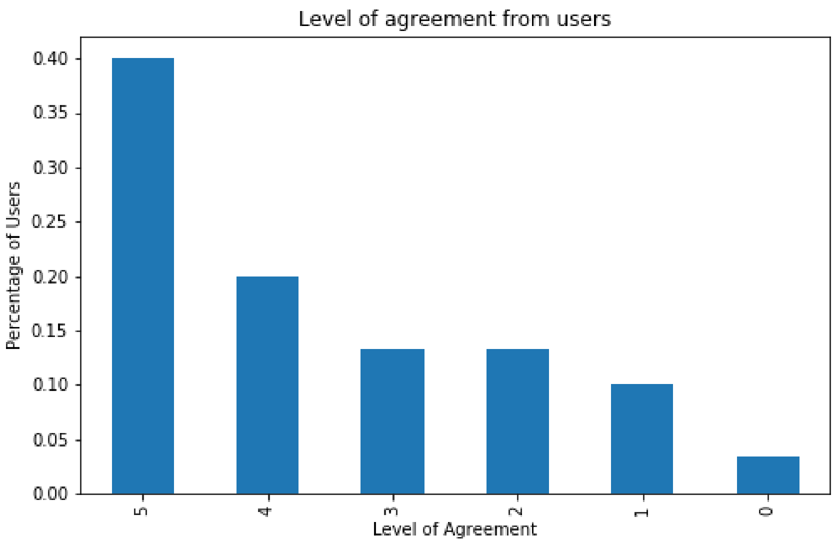


Figure 24 – User agreement level with tourist profile

Figure 25 – Photos non relevant for the fuzzy inference

each tourism class for volunteers, we can see that most users who participated in the experiment prefer Landscape tourism. It means that most photos uploaded to the system were taken in natural environments such as beaches and fields. It is justified by the type of users who participated in the experiment. The vast majority of users are from the state of Rio Grande do Norte and many of the photos were taken on the state beaches. It had a direct impact on the type of tourism profile created by fuzzy inference. It proves that the fuzzy rules were modelled correctly for the Landscape class.

In an analysis about the general level of agreement of users with the tourism profile we have as a result Figure 24. A relevant percentage of users tended to have an agreement greater than 3 of a maximum value of 5. It proves along with the other analyzes that fuzzy inference was able to describe the users' profile accurately. Besides, most users tended to have a total agreement with the fuzzy inference, entering the degree of approval 5.

Performing the analysis of users who had a low degree of agreement with the result of the inference, we can notice some specific patterns. Users who did not like the inference had a small number of photos uploaded to the system(less than four photos). As a result, the fuzzy inference system could not detect a pattern in the types of environments frequented by the user. This way, users with few uploaded photos have a low-quality inference. However, having a large number of photos uploaded did not mean good results. Some users have submitted more than five photos but disagreed with the inference. By analyzing the photos, another pattern of low-quality inferences can be seen. Pictures do not have content suitable for the approach. The photos uploaded were selfies that did not spell out the types of environments users were in, as seen in Figure 25. This experiment shows the two shortcomings of the inference system: (I) The proper number of photos must be sent (II) The photos must contain information about the environment that were taken.

Reinforcing the need for photos containing information about the environments where they were taken, we can analyze the tourism profile related to users who used photos of small relevance to the approach. In Table 8, it possible compare the output of two volunteers in opposite situations. The first user uploaded nine photos (number considered

Table 8 – Scene profile and their relation with photos relevant for the inference

User	Number of Photos	Landscape	Histor.	Advent.	Shopping	Urban	Contains Relevant Photos?
22	9	0.13	0.56	0.16	0.00	0.70	Yes
17	4	0.17	0.14	0.14	0.19	0.00	No

relevant) while the second user used four photos (low number). Also, user-uploaded photos 17 fit the situation shown in Figure 8. The result of fuzzy inference for this user is of little relevance (between 0.17 and .010) for all tourism classes. The inference module was unable to identify preferences for this user. In the case of user 22, we can see a different situation. In addition to the greater number of photos, the information about the places where they were taken made the user preference to be identified for the Historical and Cultural tourism classes with high relevance (greater than 0.5). In this context, we can conclude that the fuzzy inference module was able to make good use of the knowledge modelled from tourism experts. However, it is necessary that the user has photos that explain the environment in which they were taken, even in a small amount.

5.3 Recommendation

The same application used to collect the user feedback about the fuzzy inference was used to collect the information necessary to define what kind of recommendation the systems would use. After to obtain the agreement level about the tourism profile created by the fuzzy inference, the application redirects the user to the screen showed in Figure 26. In this step, the three options of recommendations (euclidian distance, cosine similarity and formula) are displayed showing the top 6 places to visit in the state of Rio Grande do Norte. In this case, only users from or who visited the state were select as volunteers for this step. Users who participated in the experiment could only recognize a suitable attraction for them or not if they know the state of Rio Grande do Norte. Attractions were shown to users in order of relevance to their profile according to the inference made in the previous step. All users participated in this step even though their level of agreement with the tourism profile was low.

After the users decide their preferences of places that they would like to visit, then the server receives the entire profile of the user containing their preferences and agreement level. This test was the base to select the best algorithm of recommendation for our approach.

The first analysis made is as objective as possible. An analysis of the number of users who chose each type of recommendation. As seen in Figure27, the highest percentage of users chose the Cosine Similarity algorithm as the best recommendation generator for

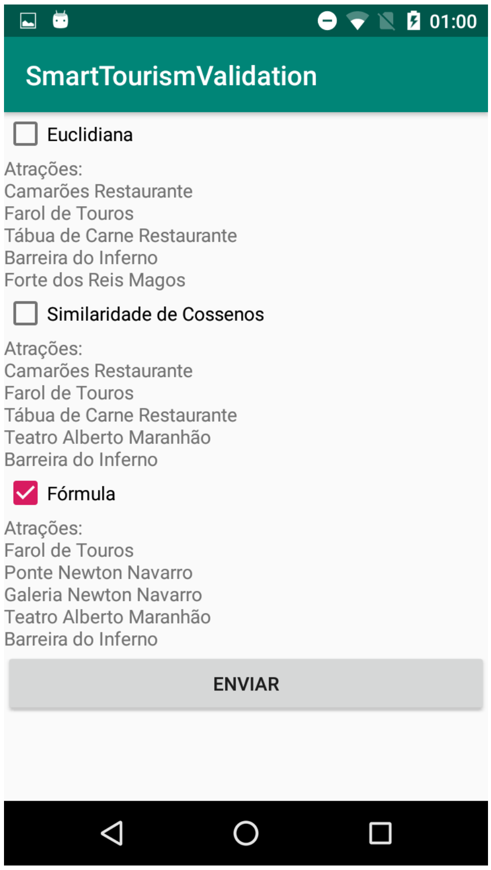


Figure 26 – Android interface to collect the recommendation preference.

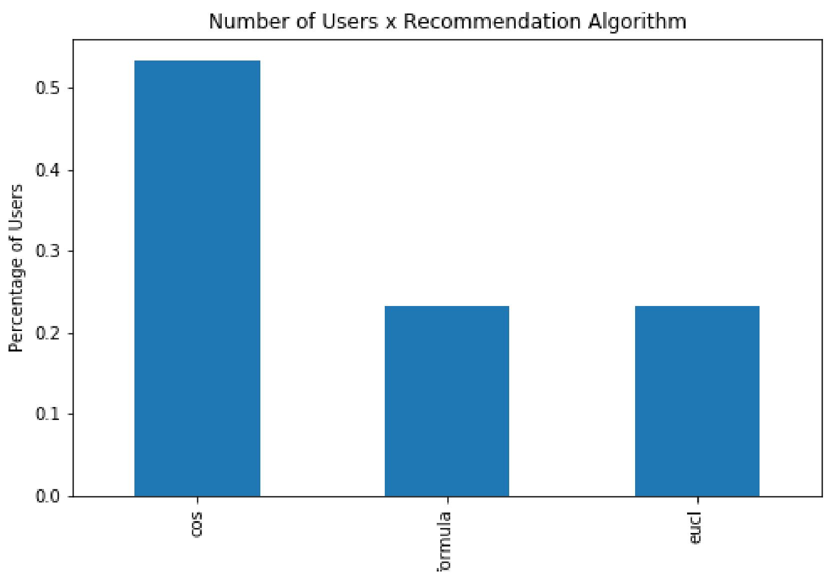


Figure 27 – Choice of recommendation algorithm

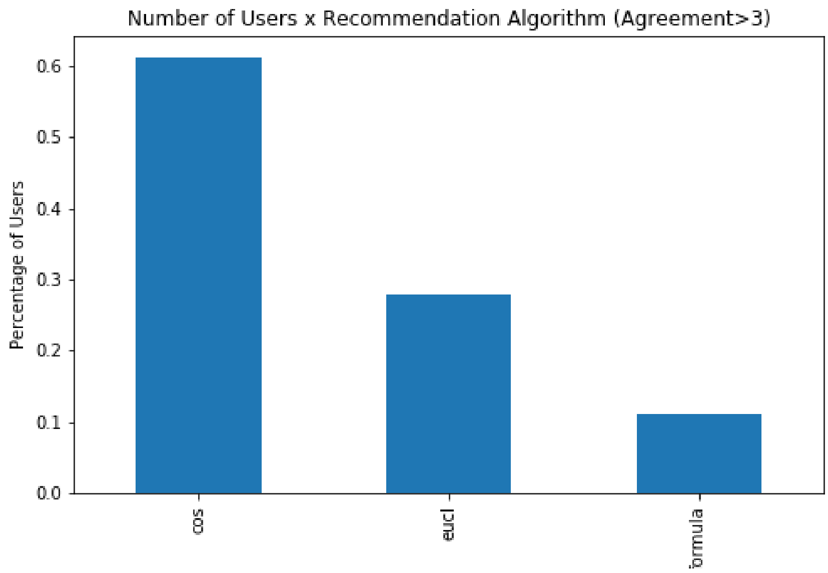


Figure 28 – Choice of recommendation algorithm by users with high level of agreement with tourism profile

their tastes.

The other two recommendation alternatives were similar between them, representing the preference of approximately 50% of users. However, there is a gap compared to cosine similarity. The same behaviour is maintained when only users with a high degree of agreement with the tourism profile. However, the formula proposed in this study has a worse result. It means that trying to use only tourist classes with a high level of relevance to the tourist is not a viable approach. The tourism classes that the tourist does not have much relation also serve as a feature for the recommendation.

In an analysis focused only on users who had a high agreement with the tourism profile generated, Figure 28 shows which behaviour of the preferred recommendation algorithm remained. The major difference is the preference given to users by the algorithm proposed in this study, showing that the algorithms that could be used in the study would be cosine similarity or Euclidean distance.

Moreover, the result performed in this step indicates that for a quality recommendation, what becomes relevant is whether the attraction vector has the same direction as the tourist vector. Values being too close is not a very important issue. An important observation is that as there are no negative values in the attraction vectors and tourism profile, it is impossible to have a similarity of negative cosines.

In this way, using a direct and objective experiment, the cosine similarity can be chosen as the recommendation used in the final pipeline of the approach. However, one might consider using a type of recommendation specific to each type of user based on the tourism profile. However, for more conclusive results on a variable type of recommendation, a more significant number of volunteers in the experiment would be interesting.

6 Concluding Remarks

Manage the tourism pre-experience is one of the most relevant challenges for the managers of smart destinations cities, especially in cities with high economic dependence on the tourism sector. Recommendation systems have been used to improve the planning step in the pre-tourism context through a huge variance of approaches.

In this context, it has become increasingly challenging to design recommendation approaches that make accurate use of different data types to create appropriate recommendations. In addition, the practicality of current systems has made such an approach require fewer user interactions. The vast majority of recommendation systems in the tourism context tend to rely on the use of many users for appropriate recommendations, characterizing the classic cold start problem. In addition, the data used for recommendations in this context have been quite limited to data that is explicitly related to the tourism context, such as asking users which places they have visited or who they like to be.

This study aims to complement and improve our prior approach to tourist preference mapping by improving the environment identification system in photos and fuzzy mapping in addition to a recommendation algorithm that can be modelled for groups. In the previous work, we have detailed the process of reproduction of the knowledge of tourism experts through a fuzzy inference system based on a vector of environments. The approach cited requested improvements that include (I) Optimization of environment classifiers for less error propagation. (II) Validation of fuzzy inference rules. (II) Feedback test with real users. (III) Definition of pipeline recommendation algorithm. (IV) Implementation of the group modelling algorithm described in well-defined research (GARCíA et al., 2009) (CHRISTENSEN; SCHIAFFINO, 2011). (V) Modelling database attractions in a universe suitable for the adopted recommendation.

In chapters 3, we introduced the Tourist Recommendation Approach and its methodology as a new way to recommend Points of Interest to users from photo-driven features. In addition, a modelling and recommendation module for groups was presented replicating an implementation presented in the research of other researchers in the area. First, the data extraction module extracts user-selected photos by granting permission or uploading photos. Subsequently, the selected photos are sent to the scene classification module that uses GoogLeNet to extract features and 25 neural networks to binary images in 25 classes of scenarios. Thus a scenario profile is created with the sum of the classifier output and used by the fuzzy module. The fuzzy inference module is the third step of the process and seeks to reduce the scene vector to a 5-position tourism profile vector representing the types of tourism types. This vector is in the same universe of attractions, which are also modelled in the 5 tourism classes with the collaboration of tourism experts. Thus, a similarity algorithm, in this case the cosine similarity, is applied between the tourism profile and the attractions to generate the tourism recommendations. The last step is the creation of group recommendation through an incremental intersection between the tourist recommendation lists through the cosine similarity scores.

The study empirical findings and validations are shown in chapter 4. We evaluate each module of the approach against its weaknesses and adversaries. The purpose of our experiments was (I) Compare two image classification techniques to define the most suitable for the context. (II) Compare fuzzy inference with human opinion. (III) Collect volunteer feedback on fuzzy inference. (IV) Detect weaknesses in the inference model. (V) Define a context-appropriate recommendation algorithm.

In our experiments we concluded that CNN classifiers used are better than Bag of Visual Word in most metrics. CNN has shown itself to be better for all tourism classes in terms of accuracy, and has faster training through knowledge transfer techniques. However, BOVW was faster for classification. However, since the purpose of the image classification module is to be as accurate as possible to be passed to the fuzzy module, CNN classifiers were adopted in the final pipeline. The fuzzy inference module, in turn, proved to be very efficient in the appropriate context. Using MAE metric, the inference error value compared to its ground truth was 0.1 overall. Moreover, most users are represented by the profile inferred by the machine, which once again demonstrates the quality of the approach. Finally, user experiments showed that the recommendation module would be more appropriate with the adoption of cosine similarity.

These results also served to show the weaknesses of the model. The approach is dependent on photos containing information from the environment where they were taken. If the scene rating module is unable to identify the types of environments users are in, it will be unable to make a good inference on the profile. In addition to quality photos the model seeks for profiles that have a good amount of photos. Profiles that add to the feature few photos with low content will have low quality inferences and therefore low quality recommendations. This also highlights another problem. The whole system bottleneck is in the fuzzy inference module. If the rules are not right the whole system can be impacted by the process.

Research has great capabilities to be further enhanced by testing in other cities with other audience profiles. In addition, testing large numbers of users on real trips to different destinations can be a huge grain for research. However, the great achievement of the survey is an alternative to a recommendation system that does not have initial user data to perform the recommendation process. Extracting photo content is a big win for systems that require a lot of effort from their users, making the process simpler and often more enjoyable.

Appendix

APPENDIX A – Fuzzy Rules

List of Fuzzy Rules Used:

1. if beach is high OR snowfield is high OR underwater is high OR mountain is high OR forest is high THEN tourist is Landscape
2. if hockey is high OR soccer field is high OR athletic field is high OR stadium is high OR golf course is high THEN tourist is Adventure
3. if museum is high OR science museum is high OR restaurant is high OR ruin is high OR church is high THEN tourist is Historical
4. if street is high OR building facade is high OR construction site is high OR airport terminal is high OR residential neighborhood is high THEN tourist is Urban
5. if gift shop is high OR bookstore is high OR shopping mall is high OR department store is high OR candy store is high THEN tourist is Shopping
6. if (beach is medium OR snowfield is medium OR underwater is medium OR mountain is medium OR forest is medium) AND (hockey is low OR soccer field is low OR athletic field is low OR stadium is low OR golf course is low) AND (museum is low OR science museum is low OR restaurant is low OR ruin is low OR church is low) AND (street is low OR building facade is low OR construction site is low OR airport terminal is low OR residential neighborhood is low) AND (gift shop is low OR bookstore is low OR shopping mall is low OR department store is low OR candy store is low) THEN tourist is Landscape
7. if (hockey is medium OR soccer field is medium OR athletic field is medium OR stadium is medium OR golf course is medium) AND (beach is low OR snowfield is low OR underwater is low OR mountain is low OR forest is low) AND (museum is low OR science museum is low OR restaurant is low OR ruin is low OR church is low) AND (street is low OR building facade is low OR construction site is low OR airport terminal is low OR residential neighborhood is low) AND (gift shop is low OR bookstore is low OR shopping mall is low OR department store is low OR candy store is low) THEN tourist is Adventure
8. if (museum is medium OR science museum is medium OR restaurant is medium OR ruin is medium OR church is medium) AND (beach is low OR snowfield is low OR underwater is low OR mountain is low OR forest is low) AND (hockey is low OR soccer field is low OR athletic field is low OR stadium is low OR golf course is low) AND (street is low OR building facade is low OR construction site is low

OR airport terminal is low OR residential neighborhood is low) AND (gift shop is low OR bookstore is low OR shopping mall is low OR department store is low OR candy store is low) THEN tourist is Historical

9. if (street is medium OR building facade is medium OR construction site is medium OR airport terminal is medium OR residential neighborhood is medium) AND (beach is low OR snowfield is low OR underwater is low OR mountain is low OR forest is low) AND (museum is low OR science museum is low OR restaurant is low OR ruin is low OR church is low) AND (hockey is low OR soccer field is low OR athletic field is low OR stadium is low OR golf course is low) AND (gift shop is low OR bookstore is low OR shopping mall is low OR department store is low OR candy store is low) THEN tourist is Urban

10. if (gift shop is medium OR bookstore is medium OR shopping mall is medium OR department store is medium OR candy store is medium) AND (beach is low OR snowfield is low OR underwater is low OR mountain is low OR forest is low) AND (museum is low OR science museum is low OR restaurant is low OR ruin is low OR church is low) AND (street is low OR building facade is low OR construction site is low OR airport terminal is low OR residential neighborhood is low) AND (hockey is low OR soccer field is low OR athletic field is low OR stadium is low OR golf course is low) THEN tourist is Shopping

Bibliography

ADOMAVICIUS, G.; TUZHILIN, A. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, v. 17, n. 6, p. 734–749, June 2005. ISSN 1041-4347.

ALAHI, A.; ORTIZ, R.; VANDERGHEYNST, P. Freak: Fast retina keypoint. In: *2012 IEEE Conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2012. p. 510–517. ISSN 1063-6919.

ALBANESE, M. et al. A multimedia recommender system. *ACM Transactions Internet Technology*, ACM, New York, NY, USA, v. 13, n. 1, p. 3:1–3:32, nov. 2013. ISSN 1533-5399. Disponível em: <http://doi.acm.org/10.1145/2532640>.

ANDRADE, J. V. Fundamentos e dimensões. Ática, São Paulo, 2002.

BAY, H. et al. Speeded-up robust features (surf). *Computer Vision and Image Understanding*, v. 110, n. 3, p. 346 – 359, 2008. ISSN 1077-3142. Similarity Matching in Computer Vision and Multimedia. Disponível em: <http://www.sciencedirect.com/science/article/pii/S1077314207001555>.

BAY, H. et al. Speeded-up robust features (surf). *Computing Vision Image Understanding*, Elsevier Science Inc., New York, NY, USA, v. 110, n. 3, p. 346–359, jun. 2008. ISSN 1077-3142. Disponível em: <http://dx.doi.org/10.1016/j.cviu.2007.09.014>.

BOES, K.; BUHALIS, D.; INVERSINI, A. Conceptualising smart tourism destination dimensions. In: ______. *Information and Communication Technologies in Tourism 2015: Proceedings of the International Conference in Lugano, Switzerland, February 3 - 6, 2015*. Cham: Springer International Publishing, 2015. p. 391–403.

BORRàS, J.; MORENO, A.; VALLS, A. Intelligent tourism recommender systems: A survey. *Expert Systems with Applications*, v. 41, n. 16, p. 7370 – 7389, 2014. ISSN 0957-4174.

BURKE, R. Knowledge-based recommender systems. v. 69, 05 2000.

BURKE, R. The adaptive web. In: BRUSILOVSKY, P.; KOBSA, A.; NEJDL, W. (Ed.). Berlin, Heidelberg: Springer-Verlag, 2007. cap. Hybrid Web Recommender Systems, p. 377–408. ISBN 978-3-540-72078-2.

CACHO, A. et al. A smart destination initiative: The case of a 2014 fifa world cup host city. In: *2015 IEEE First International Smart Cities Conference (ISC2)*. [S.l.: s.n.], 2015. p. 1–6.

CACHO, A. et al. Social smart destination: A platform to analyze user generated content in smart tourism destinations. In: ROCHA, A. et al. (Ed.). *New Advances in Information Systems and Technologies*. Cham: Springer International Publishing, 2016. p. 817–826. ISBN 978-3-319-31232-3.

CACHO, A. et al. Mobile tourist guide supporting a smart city initiative: a brazilian case study. *International Journal of Tourism Cities*, v. 2, p. 164–183, 05 2016.

CACHO, N. et al. A smart city initiative: The case of natal. In: *2016 IEEE International Smart Cities Conference (ISC2)*. [S.l.: s.n.], 2016. p. 1–7.

CARAGLIU, A.; BO, C. D.; NIJKAMP, P. Smart cities in europe. *Journal of urban technology*, Taylor & Francis, v. 18, n. 2, p. 65–82, 2011.

CHOLLET, F. et al. *Keras*. 2015. <https://keras.io>.

CHRISTENSEN, I.; SCHIAFFINO, S. Entertainment recommender systems for group of users. *Expert Systems with Applications*, v. 38, p. 14127–14135, 11 2011.

CODINA, V.; CECCARONI, L. Taking advantage of semantics in recommendation systems. In: *Proceedings of the 2010 Conference on Artificial Intelligence Research and Development: Proceedings of the 13th International Conference of the Catalan Association for Artificial Intelligence*. Amsterdam, The Netherlands, The Netherlands: IOS Press, 2010. p. 163–172. ISBN 978-1-60750-642-3. Disponível em: <http://dl.acm.org/citation.cfm?id=1893268.1893291>.

COHEN, B. *Smart cities wheel*. 2011. Disponível em: <http://www.boydcohen.com/smartcities.html>.

CSURKA, G. et al. Visual categorization with bags of keypoints. In: *In Workshop on Statistical Learning in Computer Vision, ECCV*. [S.l.: s.n.], 2004. p. 1–22.

DENG, J. et al. ImageNet: A Large-Scale Hierarchical Image Database. In: *CVPR09*. [S.l.: s.n.], 2009.

FIGUEREDO, M. R. C. Tourism preference detection approach to support a smart tourism initiative. 06 2018.

GARCíA-CRESPO, A. et al. Speta: Social pervasive e-tourism advisor. *Telematics and Informatics*, v. 26, n. 3, p. 306 – 315, 2009. ISSN 0736-5853. Mobile and wireless communications: Technologies, applications, business models and diffusion. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0736585308000683>.

GARCíA, I. et al. A group recommender system for tourist activities. p. 26–37, 09 2009.

GOOD, N. et al. Combining collaborative filtering with personal agents for better recommendations. In: *Proceedings of the Sixteenth National Conference on Artificial Intelligence and the Eleventh Innovative Applications of Artificial Intelligence Conference Innovative Applications of Artificial Intelligence*. Menlo Park, CA, USA: American Association for Artificial Intelligence, 1999. (AAAI '99/IAAI '99), p. 439–446. ISBN 0-262-51106-1. Disponível em: <http://dl.acm.org/citation.cfm?id=315149.315352>.

GUO, Y.; LIU, H.; CHAI, Y. The embedding convergence of smart cities and tourism internet of things in china: An advance perspective. v. 2, p. 54–69, 06 2014.

HANAFI, M. et al. An understanding and approach solution for cold start problem associated with recommender system: A literature review. 05 2018.

HE, C.; PARRA, D.; VERBERT, K. Interactive recommender systems: A survey of the state of the art and future research challenges and opportunities. *Expert Systems with Applications*, v. 56, p. 9 – 27, 2016. ISSN 0957-4174. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0957417416300367>.

HE, K. et al. Deep residual learning for image recognition. *CoRR*, abs/1512.03385, 2015. Disponível em: <http://arxiv.org/abs/1512.03385>.

HSU, F.-M.; LIN, Y.-T.; HO, T.-K. Design and implementation of an intelligent recommendation system for tourist attractions: The integration of ebm model, bayesian network and google maps. *Expert Systems with Applications*, v. 39, n. 3, p. 3257 – 3264, 2012. ISSN 0957-4174. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0957417411013224>.

IGNARRA, L. R. Fundamentos do turismo. Pioneira, São Paulo, 1999.

JANG, J. . R. Anfis: adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man, and Cybernetics*, v. 23, n. 3, p. 665–685, May 1993. ISSN 0018-9472.

JOZI, A. et al. Wang and mendel's fuzzy rule learning method for energy consumption forecasting considering the influence of environmental temperature. In: *2016 Global Information Infrastructure and Networking Symposium (GIIS)*. [S.l.: s.n.], 2016. p. 1–6. ISSN 2150-329X.

JULESZ, B. Textons, the elements of texture perception, and their interactions. v. 290, p. 91–7, 04 1981.

KOHAVI, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: *Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1995. (IJCAI'95), p. 1137–1143. ISBN 1-55860-363-8. Disponível em: <http://dl.acm.org/citation.cfm?id=1643031.1643047>.

KOMNINOS, N.; PALLOT, M.; SCHAFFERS, H. Special issue on smart cities and the future internet in europe. *Journal of the Knowledge Economy*, Springer, v. 4, n. 2, p. 119–134, 2013.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. In: *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*. USA: Curran Associates Inc., 2012. (NIPS'12), p. 1097–1105. Disponível em: <http://dl.acm.org/citation.cfm?id=2999134.2999257>.

LAMSFUS, C. et al. Smart tourism destinations: An extended conception of smart cities focusing on human mobility. In: TUSSYADIAH, I.; INVERSINI, A. (Ed.). *Information and Communication Technologies in Tourism 2015*. Cham: Springer International Publishing, 2015. p. 363–375. ISBN 978-3-319-14343-9.

LAMSFUS, C. et al. Going mobile. *Journal of Travel Research*, v. 54, n. 6, p. 691–701, 2015.

LITVIN, S. W.; GOLDSMITH, R. E.; PAN, B. Electronic word-of-mouth in hospitality and tourism management. *Tourism Management*, v. 29, n. 3, p. 458 – 468, 2008. ISSN 0261-5177.

LOH, S. et al. A tourism recommender system based on collaboration and text analysis. *Journal of IT and Tourism*, v. 6, p. 157–165, 2003.

LOHMANN, G. *Tourism theory : concepts, models and systems*. [S.l.: s.n.], 2017. ISBN 9781780647166.

LOWE, D. G. Distinctive image features from scale-invariant keypoints. *Intelligence Journal Computing Vision*, Kluwer Academic Publishers, Hingham, MA, USA, v. 60, n. 2, p. 91–110, nov. 2004. ISSN 0920-5691. Disponível em: <https://doi.org/10.1023/B:VISI.0000029664.99615.94>.

MALTZ, D.; EHRLICH, K. Pointing the way: Active collaborative filtering. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA: ACM Press/Addison-Wesley Publishing Co., 1995. (CHI '95), p. 202–209. ISBN 0-201-84705-1. Disponível em: <http://dx.doi.org/10.1145/223904.223930>.

MIDDLETON, S. E.; ALANI, H.; ROURE, D. D. Exploiting synergy between ontologies and recommender systems. *CoRR*, cs.LG/0204012, 2002. Disponível em: <http://arxiv.org/abs/cs.LG/0204012>.

MORENO, A. et al. Sigtur/e-destination: Ontology-based personalized recommendation of tourism and leisure activities. *Eng Apply Artificial Intelligence*, Pergamon Press, Inc., Tarrytown, NY, USA, v. 26, n. 1, p. 633–651, jan. 2013. ISSN 0952-1976. Disponível em: <http://dx.doi.org/10.1016/j.engappai.2012.02.014>.

NEUHOFER, B.; BUHALIS, D. Understanding and managing technology-enabled enhanced tourist experiences. In: ______. *The 2nd Advances in Hospitality and Tourism Marketing and Management*. Cham: [s.n.], 2012.

PAZZANI, M. J.; BILLSUS, D. The adaptive web. In: BRUSILOVSKY, P.; KOBSA, A.; NEJDL, W. (Ed.). Berlin, Heidelberg: Springer-Verlag, 2007. cap. Content-based Recommendation Systems, p. 325–341. ISBN 978-3-540-72078-2. Disponível em: <http://dl.acm.org/citation.cfm?id=1768197.1768209>.

RENNINGER, L. W.; MALIK, J. When is scene identification just texture recognition? *Vision Research*, v. 44, n. 19, p. 2301 – 2311, 2004. ISSN 0042-6989. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0042698904001919>.

RESNICK, P. et al. Grouplens: An open architecture for collaborative filtering of netnews. In: *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work*. New York, NY, USA: ACM, 1994. (CSCW '94), p. 175–186. ISBN 0-89791-689-1. Disponível em: <http://doi.acm.org/10.1145/192844.192905>.

RUSSAKOVSKY, O. et al. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision (IJCV)*, v. 115, n. 3, p. 211–252, 2015.

SARWAR, B. et al. Item-based collaborative filtering recommendation algorithms. In: *Proceedings of the 10th International Conference on World Wide Web*. New York, NY, USA: ACM, 2001. (WWW '01), p. 285–295. ISBN 1-58113-348-0. Disponível em: <http://doi.acm.org/10.1145/371920.372071>.

SHARDANAND, U.; MAES, P. Social information filtering: Algorithms for automating “word of mouth”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA: ACM Press/Addison-Wesley Publishing Co., 1995. (CHI '95), p. 210–217. ISBN 0-201-84705-1. Disponível em: <http://dx.doi.org/10.1145/223904.223931>.

SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. *CoRR*, abs/1409.1556, 2014. Disponível em: <http://arxiv.org/abs/1409.1556>.

SPRINGENBERG, J. T. et al. Striving for simplicity: The all convolutional net. *CoRR*, abs/1412.6806, 2014. Disponível em: <http://arxiv.org/abs/1412.6806>.

SRIVASTAVA, N. et al. Dropout: A simple way to prevent neural networks from overfitting. *Journal Machine Learning Res*, JMLR.org, v. 15, n. 1, p. 1929–1958, jan. 2014. ISSN 1532-4435. Disponível em: <http://dl.acm.org/citation.cfm?id=2627435.2670313>.

SZEGEDY, C. et al. Going deeper with convolutions. *CoRR*, abs/1409.4842, 2014. Disponível em: <http://arxiv.org/abs/1409.4842>.

WICKENS, E. The sacred and the profane: a tourist typology. *Annals of Tourism Research*, v. 29, p. 834–851, 2002.

XIANG, Z.; FESENMAIER, D. R. Analytics in tourism design. In: ______. *Analytics in Smart Tourism Design: Concepts and Methods*. Cham: Springer International Publishing, 2017. p. 1–10. ISBN 978-3-319-44263-1.

XIAO, J. et al. Sun database: Large-scale scene recognition from abbey to zoo. In: *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2010. p. 3485–3492. ISSN 1063-6919.

ZHOU, B. et al. Places: A 10 million image database for scene recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, IEEE, 2017.

Universidade Federal do Rio Grande do Norte - UFRN
Sistema de Bibliotecas - SISBI
Catalogação de Publicação na Fonte. UFRN - Biblioteca Central Zila Mamede

Figueredo, Mickael Raninson Carneiro.
A Tourism Multi-user Recommendation Approach Based on Social Media Photos / Mickael Raninson Carneiro Figueredo. - 2020.
77 f.: il.

Dissertação (mestrado) - Universidade Federal do Rio Grande do Norte, Centro de Ciências Exatas e da Terra, Programa de Pós-Graduação em Sistemas e Computação, Natal, RN, 2020.
Orientador: Prof. Dr. Nélio Alessandro Azevedo Cacho.

1. Smart destination - Dissertação. 2. Deep learning - Dissertação. 3. Fuzzy inference - Dissertação. 4. Turism - Dissertação. 5. Social media photos - Dissertação. I. Cacho, Nélio Alessandro Azevedo. II. Título.

RN/UF/BCZM CDU 621.391:338.48

Elaborado por Ana Cristina Cavalcanti Tinôco - CRB-15/262