

Discovering a tourism destination with social media data: BERT-based sentiment analysis

Discovering a
tourism
destination

Marlon Santiago Viñán-Ludeña and Luis M. de Campos
*Departamento de Ciencias de la Computación e Inteligencia Artificial,
ETSI Informática y de Telecomunicación, CITIC-UGR,
University of Granada, Granada, Spain*

907

Received 15 October 2021
Revised 27 January 2022
12 May 2022
Accepted 30 July 2022

Abstract

Purpose – The main purpose of this paper is to analyze a tourist destination using sentiment analysis techniques with data from Twitter and Instagram to find the most representative entities (or places) and perceptions (or aspects) of the users.

Design/methodology/approach – The authors used 90,725 Instagram posts and 235,755 Twitter tweets to analyze tourism in Granada (Spain) to identify the important places and perceptions mentioned by travelers on both social media sites. The authors used several approaches for sentiment classification for English and Spanish texts, including deep learning models.

Findings – The best results in a test set were obtained using a bidirectional encoder representations from transformers (BERT) model for Spanish texts and TweetEval for English texts, and these were subsequently used to analyze the data sets. It was then possible to identify the most important entities and aspects, and this, in turn, provided interesting insights for researchers, practitioners, travelers and tourism managers so that services could be improved and better marketing strategies formulated.

Research limitations/implications – The authors propose a Spanish-Tourism-BERT model for performing sentiment classification together with a process to find places through hashtags and to reveal the important negative aspects of each place.

Practical implications – The study enables managers and practitioners to implement the Spanish-BERT model with our Spanish Tourism data set that the authors released for adoption in applications to find both positive and negative perceptions.

Originality/value – This study presents a novel approach on how to apply sentiment analysis in the tourism domain. First, the way to evaluate the different existing models and tools is presented; second, a model is trained using BERT (deep learning model); third, an approach of how to identify the acceptance of the places of a destination through hashtags is presented and, finally, the evaluation of why the users express positivity (negativity) through the identification of entities and aspects.

Keywords Social media, Deep learning, Tourism, Sentiment analysis, Data analysis

Paper type Research paper

© Marlon Santiago Viñán-Ludeña and Luis M. de Campos. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licenses/by/4.0/legalcode>

This work has been funded by the Spanish Ministerio de Ciencia e Innovación, Agencia Estatal de Investigación under project PID2019-106758GB-C31, and the European Regional Development Fund (ERDF-FEDER).



使用社交媒体数据发现旅游目的地：基于 bert 的情感分析。

摘要

研究目的 – 这项工作的主要目的是使用情感分析技术和来自 Twitter 和 Instagram 的数据来分析旅游目的地, 以便找到最具代表性的实体 (或地点) 和用户的感知 (或方面)。

研究设计/方法/途径 – 我们使用 90,725 个 Instagram 帖子和 235,755 个 Twitter 推文来分析格拉纳达 (西班牙) 的旅游业, 以确定旅行者在两个社交媒体网站上提到的重要地点和看法。我们使用了几种方法对英语和西班牙语文本进行情感分类, 包括深度学习模型。

研究发现 – 测试集中的最佳结果是使用来自 Transformers (BERT) 模型的双向编码器表示 (BERT) 用于西班牙语文本和 Tweeteval 用于英语文本, 这些结果随后用于分析我们的数据集。然后可以确定最重要的实体和方面, 这反过来又为研究人员、从业人员、旅行者和旅游管理者提供了有趣的见解, 从而可以改进服务并制定更好的营销策略。

研究局限性 – 我们提出了一个用于执行情感分类的西班牙旅游 BERT 模型, 以及通过主题标签找到地点并揭示每个地点的重要负面方面的过程。

实践意义 – 该研究使管理人员和从业人员能够使用我们发布的西班牙旅游数据集实施西班牙-BERT 模型, 以便在应用程序中采用该数据集, 以找到正面和负面的看法。

研究原创性 – 本研究提出了一种如何在旅游领域应用情感分析的新方法。首先, 介绍了评估不同现有模型和工具的方法; 其次, 使用 BERT (深度学习模型) 训练模型; 第三, 提出了如何通过标签识别目的地地点的接受度的方法, 最后通过实体和方面的识别来评估用户表达积极性 (消极性) 的原因。

关键词 情感分析, 社交媒体, 数据分析, 旅游, 深度学习
文章类型 研究型论文

1. Introduction

Researchers use different methods on various social media sites such as Twitter, Facebook, Instagram and TripAdvisor (Viñan-Ludeña, 2019) to analyze and identify patterns. This information has been conceptualized as user-generated content (UGC) (Daugherty *et al.*, 2008; Krumm *et al.*, 2008), and it acts as an additional source of data that travelers consider as part of their search information process (Cox *et al.*, 2009).

The ability to analyze social media information is extremely useful for visitors, administrators and tour operators. In Widmar *et al.* (2020), Netbase platforms are used to study the number of online posts from Twitter and other sites, including blogs, news releases and online publications associated with Walt Disney World and SeaWorld to quantify social media data, and their findings were compared with publicly available performance measures. In van Dijck (2009), the effect of photo themes on facilitating social media user engagement on Facebook brand pages was assessed and the importance of designing images for developing destination marketing strategies was emphasized.

Public opinion about a product or service is a very important resource used by managers to evaluate it and discover what is missing, what is wrong or why users are not satisfied when visiting a place or using a tourist service. Because traditional methods such as interviews and surveys are very expensive and inefficient, social networking data in the form of tweets (Twitter), posts (Instagram, Facebook), forums, travel reviews (TripAdvisor), etc. are, therefore, widely used to mine public opinion about services or products. There are many supervised and unsupervised machine learning-based techniques that have been developed to identify text polarity and, in recent years, more accurate deep learning methods have been developed and used in text classification (Hao *et al.*, 2020a, 2020b; Kirilenko *et al.*, 2018; Liu *et al.*, 2019).

In view of the above, the main aim of this work is to analyze the main techniques in sentiment analysis (SA) using tourism social media data, select the best options for English and Spanish, extract the popular and undervalued places and services of tourist

destination and give relevant insights to managers and practitioners to improve destination image and enhance the reputation management. Thus, our work has the following research contributions: first, we analyze the performance of SA tools for English and Spanish texts; second, we propose a deep learning model for Spanish texts, and the information for training the model was collected from reviews of the 30 most recommended Spanish destinations on TripAdvisor and tweets from a Workshop on Semantic Analysis (TASS)[1] (2019 edition), and the TripAdvisor data used in this research has been shared in a free access database so that it may be used by the community to further research; third, the highest performance tool is selected to classify our tweets and Instagram posts about Granada connected with tourism in the region to analyze any negative information, to find the shortcomings of the tourist destination, to identify the hashtags that refer to a tourist attraction so that corresponding tweets or posts may be collected and the features or important characteristics of the negative data be identified through detailed text analysis; and fourth, to compare the results between Twitter and Instagram information.

This paper is organized in the following way: Section 2 provides an overview of SA in the tourism sector; Section 3 examines the data collection, data classification and tourism data analysis; Section 4 examines our results and discussion; and, finally, Section 5 outlines our conclusions and their implications, and presents our future lines of research.

2. Literature review

2.1 *Social media and sentiment analysis*

SA “is the field of study that analyzes people’s opinions, sentiments, evaluations, appraisals, attitudes and emotions towards entities such as products, services organizations, individuals issues, events, topics and their attributes” (Liu, 2012). In other words, SA, or more precisely opinion mining, consists in identifying attitudes, moods and emotions. These concepts are very important when it comes to understanding the social psychology of how a group or an individual might modify their beliefs, choices and perceptions of the world (Liu, 2015). One example of this is “influencers,” who persuasively promote company brands on social media (Viñán-Ludeña *et al.*, 2020).

The immense amount of data that is generated in social media has allowed researchers to analyze it in different domains such as politics and tourism marketing. The value of this information is immense. In marketing, for example, Kräussl and Mirgorodskaya (2017) investigate the impact of media pessimism on financial market returns; their results claim that negative (positive) social media data are associated with negative (positive) market returns. In the next subsection, we analyze SA in tourism.

2.2 *Sentiment analysis and tourism*

Because of the large amount of data generated from social media (blogs, micro-blogs, reviews, etc.), it is necessary to summarize this data to obtain useful information for tourism managers and travelers (Hao *et al.*, 2020a, 2020b) and SA provides people’s opinion or feedback about a product or destination. Alaei *et al.* (2019) examine different SA approaches to identify the performance of each one using different data sets. In Gu *et al.* (2018), sentiment emotion was used to build a system that consists in analyzing tourism blog data, and to summarize and visualize their content, thereby reducing the time spent reviewing travel blogs. An exploratory analysis is presented in Saura *et al.* (2018), where the authors use Twitter data to identify positive, neutral and negative factors that affect user experience when visiting hotels in Spain. The use of Twitter information also enabled to perform correspondence analysis and a statistical technique to understand the sentiments and opinions about English tweets about tourism in Peru (Cajachahua and Burga, 2017). In times

of recession, it is important to analyze the destination image and (Gkritzali *et al.*, 2018) used SA to analyze TripAdvisor forum messages posted about Athens to understand the destination image of the city.

Feature extraction enables us to reduce the amount of redundant data and recognize entities, thereby improving opinion classification (Abirami and Askarunisa, 2016; Afrizal *et al.*, 2019; Gunathilaka *et al.*, 2019).

This paper evaluates different SA methods. The best one is then chosen and applied to the tourism data relating to the province of Granada gathered on Twitter and Instagram to analyze the shortcomings of the tourist destination and the differences between user perceptions on each social networking site.

3. Methodology

In this section, we introduce the framework used for data extraction, data classification and tourism management. We use a crawler application to collect data from Twitter and Instagram. In Figure 1, we summarize the approach used.

First, the data extraction phase is defined from the social media platforms. A number of SA methods or tools are then evaluated to choose the tools with the best accuracy for each language. These tools are described in the Section 3.2. Finally, the polarity information obtained from the Twitter and Instagram data with the chosen tools is analyzed to identify places through hashtags and aspects (nouns), and to propose recommendations to practitioners and travelers about the tourist destination. In the following subsections, each stage of the proposed framework is described in detail.

The process to apply SA to our tourism data consists of three stages. First, the initialization step that includes data collection, preprocessing data (tokenization, stop-words removal and special characters removal) and the sentiment-level selection, in our case at sentence level. Second is the learning step that consists in the selection of the model, fine-tuned process (if it is deep-learning-based) and the training process; it also includes that we need a training data set and validation process using the test data set. Last, the summarization of results – in our case, the identification of the acceptance of places and services.

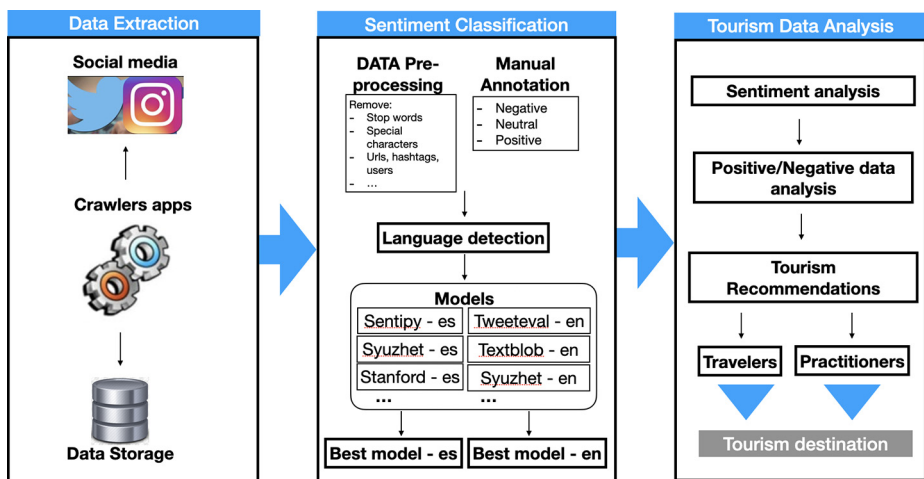


Figure 1. Tourism sentiment analysis framework

3.1 Data collection

In the case of Twitter, the use of a python scraping tool called “Twint” is recommended and this enables us to obtain tweets without the Twitter API. Because Instagram has certain restrictions regarding the use of its API, the development of a Java program for this analysis enables us to obtain data using the API and a job-scheduling process to deal with the daily data access restrictions.

The keywords chosen for Spanish data in both Twitter and Instagram are *granadaturismo*, *teenseñomigranada*, *alhambracultura*, *#alhambra*, *granada turismo*, *turismogranada*, *gastronomia granada*, *gastronomiagranada*, *hoteles granada*, *hotelesgranada*, *granadahoteles*, *restaurantes granada*, *restaurantesgranada*, *granadarestaurantes*, *#planesgranada*, *#albaicín* and *#sierranevada*. The keywords chosen for English in both platforms are *#welovegranada*, *#granadatrip*, *#granadatourism*, *granadatour*, *granadatours*, *granadatourisme*, *granadatourtravel*, *granadatourcenter*, *granadatourtravels*, *travelgranada*, *granadatourtraveler*, *travelinggranada*, *sixsensestravelsgranada*, *granadatourtraveltips*, *triptogranada*, *thingstodoingranada*, *granadathingstodo*, *granadahotels*, *granadaluxuryhotels*, *cheaphotelsgranada* and *granadarestaurant*.

We apply all the recommendations mentioned in [Viñán-Ludeña and de Campos \(2022\)](#).

3.2 Sentiment classification

The Twitter data has been chosen to manually label the polarity of 1,000 English tweets and 1,000 Spanish tweets, and then to evaluate the behavior of the different SA tools with a random sample of our tourism data about Granada. The most used approaches in SA are the supervised learning approach and the lexicon-based approach. In this paper, both approaches have been chosen for both languages, and the following SA tools have been chosen to classify polarity in Spanish and English:

- *Senti-py* [2]: This tool was chosen because of its popularity among the Github community (Spanish).
- *Stanford CoreNLP* [3]: The lexicon-based approach with Stanford Sentiment Treebank corpus is used in this tool (Spanish).
- *Syuzhet* [4]: This R package extracts sentiment from the text, and the tool is a lexicon-based model (both languages).
- *Textblob* [5]: A supervised learning approach is used in this Python library (English).
- *TweetEval* [6]: This supervised learning framework is trained with Twitter data (English).

The three categories of positive, negative and neutral are chosen in this work for sentiment classification. The metrics to be analyzed are accuracy, precision, recall and the F1 score ([Lecun et al., 2015](#)). Once these metrics have been calculated, we proceed to analyze and evaluate the best tool that will be used in the next phase.

3.3 Tourism data analysis

The final part of this approach uses the best method obtained in the previous section to obtain the polarity of the tweets and posts, and then to relate this polarity with the hashtags appearing in them referring to different entities (e.g. specific places within the tourism site, or concepts such as gastronomy). An important part of this research is to identify those tweets and posts which are considered as negative, and so the characteristic hashtags of a place can be grouped and the problems that users express on social media about this place can be identified. This information is important so that services can be improved and travelers can gather accurate information about the tourist destination.

4. Results

Data was collected from Twitter and Instagram. Because both social media platforms are similar, it is possible to analyze the published texts in the same way. Of the 90,725 Instagram posts collected, 7,717 of these are written in English and 56,247 in Spanish, and the 26,761 posts written in other languages were discarded. We also collected 235,755 tweets, and these include 19,340 English tweets and 144,947 Spanish tweets, and we discarded the 71,468 tweets in other languages. The language of the tweets and posts was identified using gclid3 [7] (Google Compact Language Detector v3). We used Twitter data to choose the best sentiment classification method: 1,000 tweets in English and 1,000 tweets in Spanish were manually classified (as being positive, negative or neutral), and each classifier was then evaluated on this data set. The SA tools used in Spanish were Sentipy, Syuzhet and CoreNLP. The results are shown in Table 1.

Of the 1,000 Spanish tweets, 163 are positive, 796 neutral and 41 negative, and so this data set seems quite unbalanced. The results obtained are rather modest (with an accuracy lower than 60% and an F1 score of around 0.34). Although the Syuzhet classifier has the best accuracy and F1 score, it almost always tends to predict each post as neutral with very few hits in the negative and positive classes. The Sentipy tool, on the other hand, performs best for precision and recall because it performs better with positive and negative classes which are in the minority. As a result of its bad performance with the majority neutral class, however, this classifier has a worse accuracy and F1 score.

The classifiers selected for English data were Tweeteval, Syuzhet and Textblob. The obtained results are shown in Table 2.

Of the 1,000 English tweets, 109 are negative, 658 neutral and 233 positive. The classifier that stands out with English tweets is Tweeteval, followed by Textblob and Syuzhet. These results are significantly better than those in Spanish, which seems to suggest that there are better sentiment classifiers in English than in Spanish. The results obtained by Tweeteval are good, with an accuracy of over 75% and an F1 score of more than 0.7.

From this preliminary analysis, we can conclude that Tweeteval is a useful tool for analyzing our English tweet/post data about tourism in Granada, but we believe that to obtain more reliable results for our Spanish data, it is necessary to improve the Spanish classifiers. In the following subsection, we study three types of classifiers that use deep learning models.

4.1 Deep learning sentiment classifiers for Spanish data

4.1.1 Training data set. We propose that various classifiers based on deep learning technology be used with our Spanish tourism data. The data set was labeled at sentence

Table 1.
Model results for the
test set of Spanish
tweets

Model	Accuracy	Precision	Recall	F1
CoreNLP (Stanford)	0.5720	0.3564	0.3958	0.3402
Syuzhet R	0.6050	0.3691	0.3906	0.3404
Sentipy	0.3590	0.4064	0.5521	0.3176

Table 2.
Models results for the
test set of English
tweets

Model	Accuracy	Precision	Recall	F1
Tweeteval	0.7560	0.6862	0.7660	0.7157
Textblob	0.5480	0.5296	0.6171	0.5128
Syuzhet R	0.4810	0.5325	0.6157	0.4807

level and we used three labels (positive, negative and neutral). We therefore identified the 30 most visited places in Spain recommended by TripAdvisor, and collected reviews about each place using a Python script (30,805 reviews in total). Each review was rated by the TripAdvisor users in range between 1 and 5 (being 5: Excellent, 4: Very good, 3: Average, 2: Poor and 1: Terrible); therefore, we changed this classification, i.e. reviews that have a rating of 5 and 4 were labeled as Positive; reviews that have been rated with 3 were labeled as Neutral and reviews with ratings of 2 and 1 were labeled as Negative. Additionally, we manually labeled 3,316 tweets about Granada (this set of tweets is not included in our data set proposed in our analysis); the labeling process was done by the first author and reviewed by the second author of this work, and 4,800 tweets were tagged in the 2019 edition of the Spanish workshop on SA. We obtained a total of 38,921 labeled texts (Tripadvisor data + Granada's Tweets + Tass data), and 5,219 of these are negative, 9,096 neutral and 24,606 are positive, and these were then used to train our models.

This data set has been made publicly available for research purposes, and to the best of the authors' knowledge, it is the only publicly available Spanish tourism data set (TSD)[8].

4.2 Deep learning architectures

To compare the performance of deep learning technology, we have used three architectures based on Krohn *et al.* (2020) using a Python library called Keras [9]: stacked bidirectional long short-term memory [10] (BiLSTM), multi-convolutional network and the bidirectional encoder representation from transformers (BERT) model.

We use stacked multiple BiLSTM-family layers in our work, and this architecture has produced quite acceptable results when performing text classification. We have based our research on the architecture presented in Khron *et al.*'s book. We then implemented various modifications to the model and its hyperparameters with the addition and removal of layers to improve the results.

One of the most common models is the one generated by convolutions, and the convolutional neural network (Convnet or CNN) is an artificial neural network that presents one or more convolutional layers that enable spatial patterns to be efficiently processed (Lecun *et al.*, 2015).

One of the latest language representation models is BERT, which uses a masked language model that randomly masks various tokens from the input and fuses the left and right context, thereby enabling a deep, bidirectional transformer (Devlin *et al.*, 2019).

To apply this approach to our work, we use two data sets: one for pre-training and the other for training. We use BETO, a BERT model trained on a Spanish corpus (Cañete *et al.*, 2020) as the pre-trained data set. The TSD proposed in this work was used as the training data set. We use this pre-trained model with the same classification fine-tuning method used by Liu *et al.* (2019).

Table 3 shows the metrics of the three architectures analyzed on the 1,000 Spanish tweets used previously.

With little training data (approximately 38,000), it is clear that the BERT model for Spanish performs well and stands out not only among the other deep learning architectures analyzed but also among the models used in Table 1.

Model	Accuracy	Precision	Recall	F1
Stacked Bi-LSTM	0.3400	0.3479	0.3500	0.2877
MultiConvnets	0.5450	0.3195	0.3274	0.3043
Spanish-BERT	0.7570	0.5635	0.6278	0.5875

Table 3.
Deep learning models
results for the test set
of Spanish tweets

4.3 Entities and aspect extraction

Using the BERT model for Spanish data and Tweeteval for English tweets/posts, we have obtained the sentiment for each one. As the data does not have geo-location information, to observe the polarity of each entity (place, event, heritage, gastronomy or tourist activity) of the tourist destination, we have associated the polarity obtained for each tweet/post with the hashtags appearing in it. In this way, if a tweet that mentions, for example, both #alhambra and #generalife is classified as positive, then this polarity is also associated to the two hashtags. Moreover, different hashtags that refer essentially to the same entities were grouped together. By way of example, a number of different hashtags are used to refer to the Alhambra, the most important tourist sight in Granada: “#alhambra,” “#laalhambra,” “#alhambradegradanada” and “#thealhambra.” In this way, 38 entities have been identified [11], and their polarity frequency counts are shown in Table 4 for Twitter data.

The most popular or most mentioned entities in Granada (although not necessarily the best valued ones) can be identified as the “Alhambra,” “Albaicin” (two World Heritage sites), “sierranevada” and “gastronomy.” Table 4 shows the places or entities ordered according to the positive percentage. Generally speaking, the number and percentage of negative tweets are rather small (especially in comparison with the positive tweets), and so the degree of satisfaction with the entities tends to be high.

It is important to note that although places such as “salobreña,” “cisterns,” “plazabibrambla” or “sannicolasviewpoint” are not often mentioned on social media, the tourists who do visit such places have quite a positive perception of them. Destinations such as “salobreña,” “plazabibrambla,” “sannicolasviewpoint,” “cathedral,” “almuñecar,” “lecrinvalley,” “lanjaron,” “costatropical,” “alpujarra” and “cisterns” can therefore be regarded as relatively undervalued places with an important tourist potential.

We also identified the entities for Instagram data, and their polarity frequency counts are shown in Table 5.

According to these data, the most popular entities are again “Alhambra,” “Albaicin” and “sierranevada,” with the inclusion of “holyweek.” We can also see that the percentage of neutral Instagram posts is considerably lower than that of Twitter, and this may perhaps reflect the greater use of Twitter to spread objective information. The percentage of negative posts is also lower, and that of positive posts is greater than the number of Twitter posts. In the same way, the results of the Instagram SA shown in Table 5 rate “realejo,” “lecrinvalley,” “almuñecar,” “sannicolasviewpoint,” “sacromonte” and “flamenco” as highly valued entities but ones that are not excessively mentioned, and so again these have a good tourist potential.

To identify the most important aspects of negative data, we use the following process. First, we clean each publication to remove any links and special characters. We use the parts of speech (POS) in this process whereby the entities were identified as nouns and the adjectives provide us with the positive and negative perceptions of these entities. The data are already classified using both Tweeteval for the English texts and the BERT model for the Spanish data, but we complement it by trying to specify the negative aspects by identifying the nouns and adjectives. We determined the relation or frequency between entity and adjectives by detecting whether both the entity (noun) and aspect (adjective) are present in a certain tweet/post. The size of the result that the previous query returns constitutes the frequency of the noun–adjective pair. For example, if we were to look for tweets containing both “Granada” (entity) and “wrong” (aspect), the result would show that these words appear in seven tweets. The NLTK [12] Python library was used to recognize the POS (nouns and adjectives) of each tweet/post written in English and the Python framework called Spacy [13] to identify nouns and adjectives written in Spanish. The earlier process was performed only for negative adjectives. To do so, the sentiment score was

Hashtag	Negative	Neutral	Positive	Total	%Neg	%Neu	%Pos
<i>Places</i>							
salobreña	1	97	99	197	0.51	49.24	50.25
plazabibrambla	4	80	57	141	2.84	56.74	40.43
sannicolásviewpoint	4	131	88	223	1.79	58.74	39.46
cathedral	0	109	68	177	0.00	61.58	38.42
almuñecar	9	463	274	746	1.21	62.06	36.73
lecrinvalley	3	229	126	358	0.84	63.97	35.19
lanjaron	1	98	52	151	0.66	64.90	34.44
costatropical	2	236	101	339	0.59	69.62	29.79
alpujarra	11	682	285	978	1.12	69.73	29.14
alhambra	381	15,292	6,207	21,880	1.74	69.89	28.37
generalife	11	565	225	801	1.37	70.54	28.09
granada	1,209	43,810	17,261	62,280	1.94	70.34	27.72
guadix	1	228	83	312	0.32	73.08	26.60
sierranevada	306	10,807	3,827	14,940	2.05	72.34	25.62
patiodelosleones	3	199	66	268	1.12	74.25	24.63
sacromonte	9	490	162	661	1.36	74.13	24.51
monachil	20	224	76	746	6.25	70.00	23.75
albaicin	59	2,661	838	3,558	1.66	74.79	23.55
guejardelasierra	1	108	29	138	0.72	78.26	21.01
motril	3	200	53	256	1.17	78.13	20.70
realejo	3	187	43	233	1.29	80.26	18.45
palaciocarlosv	3	202	40	245	1.22	82.45	16.33
<i>Heritage</i>							
cisterns	4	94	105	203	1.97	46.31	51.72
culture	18	1,204	413	801	1.10	73.64	25.26
lorca	11	257	87	355	3.09	72.39	24.51
flamenco	4	419	271	533	0.75	78.61	20.64
<i>Activities – Events</i>							
ski	26	1,147	459	1,632	1.59	70.28	28.13
crosses-in-Granada	1	97	37	135	0.74	71.85	27.41
holy-week	36	1,103	256	1,395	2.58	79.07	18.35
corpus	8	244	53	305	2.62	80	17.38
<i>Gastronomy</i>							
restaurant	68	1,601	836	2,505	2.71	63.91	33.37
tapas	18	709	271	998	1.80	71.04	27.15
gastronomy	74	3,708	1,380	5,162	1.43	71.83	26.73

Table 4.
Sentiment analysis
results for Twitter
data (English–
Spanish)

calculated for each adjective found using Texblob for tweets/posts in English, and CoreNLP Stanford for Spanish, thus selecting only negative adjectives.

Table 6 summarizes the most important [entity adjective] pairs extracted from the 224 English negative tweets. The number shown after each adjective is the frequency that each pair appears in this sub-data set. For example, some users relate Granada with adjectives such as wrong, impossible, disappointed, bad, terrible, stupid and horrible.

In the other section of Table 6, we can see that a number of problems relate to the purchase of tickets. For example, the tweet “alhambra cultura incredibly unprofessional ticket sale for #alhambra in #granada web doesn’t work and they hang up, phoning from australia!!!” refers to the purchase of Alhambra tickets, something which some travelers find difficult. Other users mention the booking process, “I can’t believe how much time I had to spend on booking two tickets to visit alhambra cultura. Very poor online booking system

Table 5.
Sentiment analysis
results for Instagram
data (English–
Spanish)

Hashtag	Negative	Neutral	Positive	Total	%neg	%neu	%pos
<i>Places</i>							
realejo	0	6	18	24	0	25	75
lecrinvalley	0	28	60	88	0	31.82	68.18
almuñecar	0	10	20	30	0	33.33	66.67
sannicolásviewpoint	0	28	44	72	0	38.89	61.11
sacromonte	0	51	69	120	0	42.5	57.49
granada	28	8,241	8,334	16,603	0.17	49.64	50.19
alpujarra	0	43	42	85	0	50.59	49.41
sierranevada	20	1,263	1,078	2,361	0.85	53.49	45.66
monachil	0	11	9	30	0	55	45
albaicin	7	705	551	1,263	0.55	55.82	43.63
cathedral	0	97	74	171	0	56.73	43.27
guadix	0	17	12	29	0	58.62	41.38
costatropical	0	12	8	20	0	60	40
alhambra	12	2,577	1,725	4,314	0.28	59.74	39.98
generalife	0	146	85	231	0	63.20	36.79
palaciocarlosv	0	12	5	17	0	70.58	29.41
patiodelosleones	1	26	11	38	2.63	68.42	28.95
paseodelostristes	0	16	6	22	0	72.73	27.28
<i>Heritage</i>							
flamenco	1	35	89	98	1.02	35.71	63.26
culture	1	199	182	382	0.27	52.09	47.64
<i>Activities – Events</i>							
holyweek	4	1,049	1,051	2,104	0.19	49.86	49.95
ski	2	75	54	131	1.53	57.25	41.22
<i>Gastronomy</i>							
restaurant	1	38	126	165	0.61	23.03	76.37
Tapas	0	104	89	193	0	53.89	46.11
Wine	0	11	8	19	0	57.89	42.12
Gastronomy	0	24	13	37	0	64.86	35.13

Ticketmaster.” There are also some negative comments about the narrow streets in the city. We also found illustrative (negative) noun phrases such as:

- people can’t order tickets;
- unprofessional ticket sale;
- terrible customer service;
- white elephant construction; and
- 2hr queue.

Table 7 summarizes the most important [entity adjective] pairs obtained from the 4,000+ Spanish negative tweets.

Many of the Spanish tweets refer to the management of the tourist destination. One such example is the tweet “En Granada: El dinero de la Alhambra y la sierra se lo lleva Sevilla, intolerable. Veis como los catalanes no pueden ser más españoles?,” where the user is annoyed by the policies applied in the management of tourism in Granada. Or the tweet *Otra cosilla @Granada Limpia! A ver si Uds pueden eliminar de la acera los restos de los líquidos que se derraman del contenedor verde, que a veces huele fatal y da una sensación terrible!*, in

Entity	Adjectives
<i>Places</i>	
Granada	wrong 7; impossible 6; disappointed 9; due 8; bad 8; terrible 6; other 8; stupid 3; horrible 3; long 2 disappointing 2; miserable 2; few 1; little 3; bloody 1; confusing 1; sad 7; pale 1; narrow 1; unnecessary 1 anxious 1; lazy 1; sorry 4; angry 1; annoyed 1; fake 2 weird 2; awful 2; sharp 1; violent 1; due 8; green 1; sick 1; past 2; dead 2; useless 2; confused 1; dreadful 1; sloppy 1; dangerous 1; vicious 1; annoying 1; guilty 1
Alhambra	wrong 13; impossible 7; disappointed 15; due 10; bad 10; terrible 8; other 7; stupid 4; horrible 3; long 4; disappointing 2; miserable 3; few 3; little 3; bloody 3; confusing 2; sad 8; pale 1; narrow 1; unnecessary 1; anxious 1; lazy 1; sorry 7; angry 1; annoyed 1; fake 3; weird 2; awful 4; sharp 1; violent 1; due 8; green 1; sick 1; past 4; dead 2; poor 4; useless 3; expensive 2; confused 1; dreadful 1; sloppy 1; dangerous 1; vicious 1; annoying 1; guilty 1
Sacromonte	due 1; pale 1; narrow 1; anxious 1
albaicin	confusing 1
sierranevada	impossible 1; bad 1
<i>Activities – Events</i>	
holy-week	disappointed 1; stupid 1
<i>Others</i>	
Ticket	wrong 2; impossible 2; disappointed 6; bad 1; terrible 1; other 1; stupid 1; long 1; disappointing 1; few 1; confusing 2; angry 1; fake 1; awful 2; poor 1; dreadful 1
booked	wrong 1; disappointed 1; due 3; bad 1; stupid 1; angry 1
museum	disappointed 1; due 1; pale 1; narrow 1; anxious 1
narrow streets	due 1; pale 1; narrow 1; anxious 1; due 1

Table 6.
Most important
negative features
(English-Twitter)

which the user talks about the liquid leaking from the recycling bin which gives a bad impression and can smell terrible.

The same analysis was carried out with Instagram. The most important entities found in posts written in English are *Granada*, with some relevant adjectives such as “useless,” “ugly,” “wrong,” “bad,” “terrible,” “weird” or “dark”; *Alhambra*, with adjectives such as “wrong,” “bad” and “terrible”; *Sierra Nevada* with adjectives such as “horrific,” “bad,” “terrible,” “dark” and “small” as adjectives; and *Albaicin* with adjectives such as “ugly,” “wrong,” “little” or “terrible.”

It is much more difficult to identify the polarity in Instagram posts because the text is more extensive than tweets. An additional factor is that the same text might contain the same content in different languages, and that often the entity is described but not specifically named as it can be seen in the images uploaded to Instagram.

In a similar way, we carry out the process of identifying the most important entities and adjectives with the posts written in Spanish. The most important entities are *Granada* with adjectives such as “grave,” “terrible,” “cruel” or “rival” and *Alhambra* with adjectives such as “cruel,” “rival” or “victima.”

5. Discussion and conclusions

5.1 Conclusions

The study revealed the most popular touristic entities at the destination and these tended to be the same on Twitter and Instagram, and also the best (and worst) valued. Our study also

Table 7.
The most important
negative features
(Spanish-Twitter)

Entity	Adjectives
<i>Places</i>	
granada	terrible 15; artificial 2; cruel 4; grave 33; impersonal 3; mediocre 5; manual 1; error 29; intolerable 8
alhambra	terrible 13; cruel 4; grave 9; impersonal 2; mediocre 4; manual 1; error 30; intolerable 2
sierranevada	terrible 2; artificial 1; grave 3; mediocre 1; intolerable 1
albaicin	terrible 3; grave 1
cathedral	terrible 1
patidelosleones	terrible 1
monachil	grave 1
palaciocarlosv	grave 1
<i>Heritage</i>	
culture	terrible 8; artificial 1; grave 7; impersonal 1; error 12; terrible 8
lorca	mediocre 1
<i>Activities – Events</i>	
holy-week	mediocre 1; error 1
<i>Gastronomy</i>	
restaurant	mediocre 1

revealed a number of less mentioned entities that had been fairly positively assessed, and these undervalued places are considered to have an important tourist potential.

We analyzed Instagram and Twitter because these platforms are used by travelers to comment on a tourist destination. We have also observed that the percentage of neutral posts in Instagram is considerably lower than that of Twitter, which perhaps may reflect a greater use of Twitter to spread objective information. The percentage of negative Instagram posts is lower and the percentage of positive Instagram posts is higher than on Twitter.

To try to explain the reasons why some users have negative feelings about their trip to Granada, we have also conducted a more detailed study of the negative tweets/posts, using POS methods to identify associations between entities (names) and features (aspects).

The reliability of our work is in concordance with the evaluation of the learning models; thus, if the accuracy of the classifiers is high, then the reliability is high. We obtained a reasonably good accuracy for the selected models: 0.757 for our Spanish-BERT model and 0.756 for Tweeteval. Finally, public access to both the data and the model code is provided, so the results are highly reproducible.

5.2 Research implications

Our work presents a novel approach to analyze the tourist destinations, which involves computational methods such as SA. We examine machine-learning-based tools and deep-learning approaches to evaluate the sentiment in posts uploaded to social media platforms and we propose a TourismSpanish-BERT-model that was trained with tourism data to improve Spanish sentiment classification. By processing each and every post and tweet, we are able not only to identify important entities and aspects and highlight the most visited places and the undervalued important places, but also to determine why travelers might have negative perceptions about the tourist destination.

Our work builds a solid foundation for future research, i.e. increasing sample sizes, larger destinations, different social media platforms, development of a specific taxonomy regarding classification process and integration of different types of data; all of these features can help to

enhance the results. For example, information on transport, weather conditions, special events, crises and other features may have importance on visitor satisfaction; in addition, the integration of these data might be useful to build a Smart-Destination application that integrates the data of the city/destination and social media tourism data improving the quality of visits.

5.3 Practical implications

The Spanish-BERT model we propose has important practical implications for the tourism domain. Thanks to the code provided with our TSD, it is easy to implement and use. Because our Spanish model can also be adopted in practical applications with destination management tools by managers and tourism organizations and improved with future research, our training data is fully available for such purposes. SA for entity and aspect detection can provide managers, organizations or travel agencies with information about what things should be changed or improved in terms of a specific place or service provider in a certain tourist destination. This, in turn, enables new marketing strategies to be devised and better policies implemented so that these places may become smart tourism destinations.

Moreover, managers could monitor periodically their services and destinations through applications based on our approach to update constantly the policies and strategies of destination management organizations. For example, through a software application, our model can be incorporated to identify the shortcomings in a service/place; then, fix or improve them and finally show the improvements through social media or Web platforms allowing to improve the promotion of the service/place of a tourist destination.

Our approach contributes to the knowledge on the role of tourism information in social media on destination image formation, providing meaningful insights for tourism marketers to build new strategies to enhance the tourist places, services, etc. and to attract travelers through social media. Furthermore, this work contributes to understanding the tourism experience using Twitter and Instagram with the aim to enhance mechanisms for managing customer experiences. Finally, the posts classified as negative can be useful to managers and practitioners developing recovery strategies on customers' satisfaction and trust.

5.4 Limitations and future research

This study has a number of limitations. First, the task of data acquisition is extremely complicated because of the limits imposed by Instagram on the number of posts that can be obtained, and so it is necessary to run a crawler each day for this task. It is also important to find automatic methods to detect bias in social media, as fake posts, reviews, news, etc. It is therefore interesting to have a method based on deep learning technology (using a training data set) to perform this task and improve the quality of results.

The SA process can be improved with a much larger training data set and can be combined with a standardized taxonomy to differentiate positive, negative and neutral allowing to be used in any language, adding greater consistency and validity to the results obtained. Although our work focuses in English and Spanish language, it can be extrapolated to any other language.

In the future, we will try to use a tailored BERT model to identify entities and aspects to improve the performance of this process so that it is possible to identify the reasons why users have positive or negative feelings about a tourist destination.

Notes

1. <http://tass.sepln.org/>

2. <https://github.com/ayllote/senti-py>
3. <https://stanfordnlp.github.io/CoreNLP/sentiment.html>
4. <https://cran.r-project.org/web/packages/syuzhet/index.html>
5. <https://textblob.readthedocs.io/en/dev/>
6. <https://github.com/cardiffnlp/tweeteval>
7. <https://pypi.org/project/gclid3/>
8. <https://bitbucket.org/msvl/tourismspanishdataset/src/master/>
9. <https://keras.io>
10. Which means that this model contains two Bi-LSTM layers.
11. We only considered those significant entities with a number of mentions above a threshold.
12. www.nltk.org
13. <https://spacy.io>

References

- Abirami, A. and Askarunisa, A. (2016), "Feature based sentiment analysis for service reviews", *Journal of Universal Computer Science*, Vol. 22 No. 5, pp. 650-670.
- Afrizal, A., Rakhmawati, N. and Tjahyanto, A. (2019), "New filtering scheme based on term weighting to improve object based opinion mining on tourism product reviews", *Procedia Computer Science*, Vol. 161, pp. 805-812.
- Alaei, A.R., Becken, S. and Stantic, B. (2019), "Sentiment analysis in tourism: capitalizing on big data", *Journal of Travel Research*, Vol. 58 No. 2, pp. 175-191.
- Cajachahua, L. and Burga, I. (2017), "Sentiments and opinions from twitter about peruvian touristic places using correspondence analysis", *CEUR Workshop Proceedings*, Vol. 2029, pp. 178-189.
- Cañete, J., Chaperon, G., Fuentes, R., Ho, J.-H., Kang, H. and Pérez, J. (2020), "Spanish pre-trained bert model and evaluation data", *Pml4dc at iclr*.
- Cox, C., Burgess, S., Sellitto, C. and Buultjens, J. (2009), "The role of user-generated content in tourists' travel planning behavior", *Journal of Hospitality Marketing and Management*, Vol. 18 No. 8, pp. 743-764.
- Daugherty, T., Eastin, M.S. and Bright, L. (2008), "Exploring consumer motivations for creating user-generated content", *Journal of Interactive Advertising*, Vol. 8 No. 2, pp. 16-25.
- Devlin, J., Chang, M.-W., Lee, K. and Toutanova, K. (2019), "BERT: pre-training of deep bidirectional transformers for language understanding", *Proceedings of the 2019 conference of the north American chapter of the Association for Computational Linguistics: Human language technologies*, Vol. 1, (long and short papers), pp. 4171-4186.
- Gkritzali, A., Gritzalis, D. and Stavrou, V. (2018), "Is xenios zeus still alive? Destination image of Athens in the years of recession", *Journal of Travel Research*, Vol. 57 No. 4, pp. 540-554.
- Gu, Y.H., Yoo, S.J., Jiang, Z., Lee, Y.J., Piao, Z., Yin, H. and Jeon, S. (2018), "Sentiment analysis and visualization of Chinese tourism blogs and reviews", *2018 International Conference on Electronics, Information, and Communication (ICEIC)*, pp. 1-4.
- Gunathilaka, D., Pathirana, S., Senarathne, S., Weerasekara, J. and Silva, T. (2019), "Feature based opinion mining for hotel profiling", *Communications in Computer and Information Science*, Vol. 890, pp. 219-231.

- Hao, J.-X., Fu, Y., Hsu, C., Li, X.R. and Chen, N. (2020a), "Introducing news media sentiment analytics to residents' attitudes research", *Journal of Travel Research*, Vol. 59 No. 8, pp. 1353-1369.
- Hao, J.-X., Wang, R., Law, R. and Yu, Y. (2020b), "How do mainland Chinese tourists perceive Hong Kong in turbulence? A deep learning approach to sentiment analytics", *International Journal of Tourism Research*, Vol. 23 No. 4, pp. 478-490.
- Kirilenko, A.P., Stepchenkova, S.O., Kim, H. and Li, X. (2018), "Automated sentiment analysis in tourism: comparison of approaches", *Journal of Travel Research*, Vol. 57 No. 8, pp. 1012-1025.
- Kräussl, R. and Mirgorodskaya, E. (2017), "Media, sentiment and market performance in the long run", *The European Journal of Finance*, Vol. 23 No. 11, pp. 1059-1082.
- Krohn, J., Beyleveld, G. and Bassens, A. (2020), "Deep learning illustrated", *A Visual, Interactive Guide to Artificial Intelligence*, Pearson's Addison-Wesley, Boston.
- Krumm, J., Davies, N. and Narayanaswami, C. (2008), "User-generated content", *IEEE Pervasive Computing*, Vol. 7 No. 4, pp. 10-11.
- Lecun, Y., Bengio, Y. and Hinton, G. (2015), "Deep learning", *Nature Publishing Group*, Vol. 521 No. 7553, pp. 436-444.
- Liu, B. (2012), "Sentiment analysis and opinion mining", *Synthesis Lectures on Human Language Technologies*, Vol. 5 No. 1, pp. 1-184.
- Liu, B. (2015), "Preface", *Sentiment Analysis*, Cambridge University Press, Cambridge, pp. 11-14.
- Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L. and Stoyanov, V. (2019), "Roberta: a robustly optimized bert pretraining approach", *arXiv*.
- Saura, J., Palos-Sanchez, P. and Martin, M. (2018), "Attitudes expressed in online comments about environmental factors in the tourism sector: an exploratory study", *International Journal of Environmental Research and Public Health*, Vol. 15 No. 3, p. 553.
- van Dijck, J. (2009), "Users like you? Theorizing agency in user-generated content", *Media, Culture and Society*, Vol. 31 No. 1, pp. 41-58.
- Viñan-Ludeña, M.S. (2019), "A systematic literature review on social media analytics and smart tourism", *Smart Tourism as a Driver for Culture and Sustainability*, Springer, Cham, pp. 357-374.
- Viñan-Ludeña, M.S. and de Campos, L.M. (2022), "Analyzing tourist data on twitter: a case study in the province of Granada at Spain", *Journal of Hospitality and Tourism Insights*, Vol. 5 No. 2, pp. 435-464, doi: [10.1108/JHTI-11-2020-0209](https://doi.org/10.1108/JHTI-11-2020-0209).
- Viñan-Ludeña, M.S., de Campos, L.M., Jacome-Galarza, L.R. and Sinche-Freire, J. (2020), "Social media influence: a comprehensive review in general and in tourism domain", in Rocha, A., Abreu, A., de Carvalho, J., Liberato, D., González, E. and Liberato, P. (Eds), *Advances in Tourism, Technology and Smart Systems. Smart Innovation, Systems and Technologies*, Springer, Singapore, Vol. 171, doi: [10.1007/978-981-15-2024-2_3](https://doi.org/10.1007/978-981-15-2024-2_3).
- Widmar, N.O., Bir, C., Clifford, M. and Slipchenko, N. (2020), "Social media sentiments an additional performance measure? Examples from iconic theme park destinations", *Journal of Retailing and Consumer Services*, Vol. 56, p. 102157.

Corresponding author

Marlon Santiago Viñan-Ludeña can be contacted at: marsantovi@gmail.com

For instructions on how to order reprints of this article, please visit our website:

www.emeraldgrouppublishing.com/licensing/reprints.htm

Or contact us for further details: permissions@emeraldinsight.com