Kejo Starosta / Cristian Bogdan Onete / Sonia Budz / Michael Krutwig

Differences in travelers' perceptions of popular tourist destinations estimated by a LSTM neural network: A comparison between the UK and Germany

Abstract

This study examines the differences and similarities in the perceptions of popular tourist destinations between the English and German media in an empirical and quantitative way. The perception of a tourist destination country is measured by the sentiment in the media in the tourists' countries of origin — either the UK or Germany. The study measures such perceptions with the help of a Long Short Term Memory (LSTM) artificial neural network that analyzes the perceptions in publicly available news streams in the UK and Germany. The results show that the online media indeed generates different perceptions for some of the popular travel destinations. An equally interesting finding is that there is, in general, a broad consensus between the sentiments in the British and German media. Businesses in the tourism and hospitality sectors, news providers, political actors, and policymakers can use the methodology to analyze the source of this information gap and the impact on their operations and policies.

Key words: tourism; online media; sentiment analysis; machine learning; UK; Germany

Introduction

While the European countries share many common values, standards, and views, there are also many differences between them. Two of these differences are the behavior of tourists across different European countries and their preferred travel destinations. The World Tourism Organization (UNWTO), EUROSTAT and national statistics offices provide timely data for tourism flows in popular tourist destinations for Europeans that tracks these differences. This study tries to analyze whether conventional tourism demand models can explain all these differences or whether there are additional differences in the perceptions of tourist destinations that influence the tourism flows.

There is a large number of studies that nowcast, forecast or explain the tourism flows in almost all countries with a considerable number of inbound travelers. These studies utilize many different methodologies ranging from econometric and time-series models, to the analysis of historic reasons and the analysis of the behaviour of potential travellers in the web (see e.g. Dogru, Sirakaya-Turk & Crouch, 2017; Dinis, Costa, Pacheco & Osvaldo, 2017; Li, Pan, Law & Huang, 2017; Song & Li, 2008; and many more). Conventional models analyze the differences by means of exchange rates, disposable incomes, the distance to travel, and some more factors. Further studies analyze the historical and individual reasons that go beyond the conventional models for tourism flows. These studies include, for instance, refugee flows, immigration and emigration flows in the past, or relations from the colonial

Kejo Starosta, M.Sc., The Bucharest University of Economics Studies (ASE), Romania; E-mail: k@ksfx.de

Prof. **Cristian Bogdan Onete**, PhD, The Bucharest University of Economics Studies (ASE), Bucharest, Romania; E-mail: bogdan.onete@com.ase.ro

Sonia Budz, The Bucharest University of Economics Studies (ASE), Bucharest, Romania; E-mail: sonia.budz@yahoo.com

Michael Krutwig, The Bucharest University of Economics Studies (ASE), Romania; E-mail: michael@krutwig.com

period (see e.g. Moufakkir, 2014; Seetaram, 2012). Lately there are more and more studies that analyse the behaviour of potential travellers in the web based on big data to estimate the tourism demand. For instance, Dinis, Costa, Pacheco and Osvaldo (2017) could show that the search engine query volumes are a good indicator for the tourism demand in Portuguese regions.

Most differences in tourist flows of different countries are explained well by the existing studies. In this study, we try to complement this existing knowledge about tourism flows with knowledge about different perceptions of European travelers in respect of popular tourist destinations. We analyze whether there are different perceptions in relation to popular tourist destinations and whether these differences influence tourism flows. A preceding study analyzed the impact of the sentiment in the German-speaking media on tourist arrivals in popular tourist destinations for Europeans (Starosta, Budz & Krutwig, 2019). This study shows that the German-speaking media exerts a strong influence and there is a strong correlation with tourist arrivals in popular tourist destinations for Europeans. A further study shows how different types of news reports in the German-speaking media and different types of events in the destination countries lead to different reactions on the part of the travelers, and how these news reports and events can be classified (Starosta, Budz & Krutwig, 2018). This study classifies different types of news reports which lead to stronger or weaker reactions and arousal levels of travelers. Starosta et al.'s (2019) findings on the relationship of the German-speaking online news sentiment with tourist arrivals in popular tourist destinations for Europeans lead to the question whether the news reporting in other countries of origin is similar and whether differences in the perception are reflected in the tourist arrivals in the destination countries as well. This analysis will show if the perceptions are reflected in the tourist arrivals or whether conventional models comprehensively explain the differences in the behavior of British and German travelers.

The UK and Germany are both western European countries, have some common tourist destinations (See Table 1), a similar median disposable income (Germany $25,140, UK $21,576), and the distances to the destination countries are not too different, as the average distance between the UK and Germany is only ~800km. Furthermore, both countries are not located at the Mediterranean Sea, and do not host any domestic summer holiday hotspots for Europeans. Because of their similar tourism relevant characteristics, a comparison between the UK and Germany is a reasonable choice for this study.

Instead of directly surveying the perceptions of the citizens in the UK and Germany, we measured the sentiments in the online media of the UK and Germany toward popular tourist destinations for Europeans. This is a cost-effective way compared to a direct survey of the perceptions. The sentiment in the media is a very good proxy for the perceptions of the citizens and travelers to measure the sentiment toward foreign countries. Several studies show that the online media in a country is the opinion-former par excellence, and even that the opinions in the media and the perceptions of the consumers — in this case, travelers — are mutually reinforcing.

We measure the perceptions of British and German travelers in a computer-based manner, with the help of a LSTM artificial neural network that analyzes the perceptions in the British and German media toward popular tourist destinations for Europeans. The algorithm classifies the news of two very comprehensive news streams in UK and Germany. At first, the software determines — based on full text searches — if a news article belongs to one of the destination countries under observation and then classifies the sentiment of the news article as either positive or negative. The information about the number of positive and negative news articles per day is used to generate the sentiment indices and to display the development of the sentiment over time. To create the sentiment indices, we aggregate the number of positive and negative news into a single time-series index. Each tourist destination gets two sentiment indices, one with the British and one with the German sentiment. These time-series can then be compared between the UK and Germany to find out the similarities and differences in the sentiment and also with the corresponding tourist arrival series. The result provides a clear view of the differences (and similarities) in the perceptions between British and German travelers and also shows how this is reflected in the tourist arrivals. Hence, we can see whether there is an impact of the different perceptions on the tourist arrivals.

With this new information, we want to show that Germany and the UK have different perceptions of some travel destinations, which might also affect tourist arrivals that go beyond the classical models — which overlook tourist perceptions. This information can be useful to analyze the causes for the different perceptions and acknowledge the issues that might arise for businesses and politics, based on these different views.

Theoretical foundations

Starosta et al.'s (2019) finding that the tourist arrivals in many popular tourist destinations for Europeans are strongly connected to the sentiment in online media provided the impulse to this study. Their study shows that the sentiment in the German-speaking online media is a good indicator of the number of tourist arrivals in the most popular destinations for Europeans. If this is the case, however, the perception of destinations in other European countries and other language areas should be similar, and differences should arise in the tourist arrivals as well. Many studies show how the perceptions for international relations are formed by the mass media and how important their role is in forming the opinions of citizens. McComb and Reynolds (2002) and Wanta, Golan and Lee (2004) show how good the media is in setting the agenda for the citizens and at determining what people should talk about. Wanta et al. (2004) points out that the citizens are much more likely to think in a negative manner about a country if it receives negative news coverage, and that this, of course, influences travelers' booking decisions. Brewer, Graf, and Willnat (2003) found out that a "news story that presents a frame linking an issue to a foreign nation in a way that suggests a particular evaluative implication may shape how audience members judge that nation." Zeitzoff (2016) even found evidence that armed conflicts are influenced by social media. While our focus is on the influence of the media on international relations, many other different studies analyze various topics of media influence in different areas.

Apart from Starosta et al. (2019), who try to explain the tourist arrivals based on the perceptions of the travelers on the destination country, the domain of tourism research constitute the foundations of the conventional tourism demand models and the conventional explanatory variables. Song and Li (2008) summarize most conventional tourism demand models based on time-series and econometric models that have been used between 2000 and 2008. Dogru et al. (2017) show ways on how to enhance the existing tourism demand models. Apart from these conventional models, some other recent studies in the domain of tourism research apply models that use machine-learning, artificial intelligence, and big data. The studies range from sentiment analysis to the analysis of unstructured data from the web and the use of search-engine query volumes. Li et al. (2017) create a method to forecast the tourism demand based on search-engine query volumes. Dinis et al. (2017) use a similar approach to forecast the tourism in different regions of Portugal. They found that the number of searches for a specific region and the overnight stays correlate strongly. While search engine query volumes and the analysis of social media posts are a reasonable approach to tourism forecasting, we still expect that the sentiment analysis of online media has a stronger lead to the tourist arrivals than the social media posts and search engine query volumes. The longer lead can be explained by the decision-making models of consumers and travelers of Mathieson and Wall (1982), Howard (1963), Lattin and Roberts (1991), Woodside and Sherrell (1977), and Sirakaya and Woodside (2005). All models show that destinations will probably not be considered at all for the next holiday if the media reporting is too negative. The result of this opinion-forming process, based on the media, is that there will be no social media posts and no active search process as regards the destinations.

Apart from the methods that engage big data and search engine query volumes, there is a large number of studies in the domain of tourism research that use sentiment analysis methods. In contrast to this study, they usually focus on customer care and microeconomic data like automated mail analysis or the analysis of travel reviews. Ye, Zhang and Law (2009) use machine learning-based sentiment analysis to analyze the reviews of travel destinations. In addition, they evaluate the quality of different approaches. González-Rodríguez, Martinez-Torres and Toral (2016) analyze the image of travel destinations after the visit and the influence of the then-created online reviews. Ren and Hong (2017) do not just analyze the sentiment of entire online reviews, but implement an algorithm with a deeper textual understanding that analyzes the sentiment toward different aspects of each review. Chaabani, Toujani and Akaichi (2017) provide one of the rare sentiment analysis studies that focus on macroeconomic structures of tourism. They analyze the perceptions of Tunisia in the media after the Arab spring and the impact on tourism. Instead of analyzing the sentiment in the media, Yuan, Liu and Wei (2016) use the existing dataset "The Global Data on Events, Location, and Tone" to model the tourism demand and the image of China. Recently also a few studies in the tourism domain employ LSTM networks for their sentiment analysis and demand forecasting. Law et al. (2019) could show that LSTM networks can help at the feature selection for tourism demand forecasting and how this could lead to improve forecast in general. Li and Cao (2018) used an LSTM network for pure time series forecasts and could outperform Auto Regressive Integrated Moving Average (ARIMA) models and less sophisticated neural networks. Rizal, Soraya and Tajuddin (2019) confirmed parts of these findings while trying different LSTM models. Martin et al. (2018) could show that LSTM networks are the most robust and accurate estimators for classifying the sentiment in comments made on booking.com and tripadvisor.com for hotels located on the island of Tenerife.

In addition to the sentiment analysis studies, there are further studies that analyze the impact of terrorism, war, instability, and shocks on tourism. These are important complementary studies, since these events are reflected in the media before they surface in any other indicator. Sentiment indices are especially useful to nowcast and forecast the impact of these events. Sönmez and Graefe (1998) provide a comprehensive review of several studies on terrorism and the impact on tourism. A more recent interesting study that analyzes the image of a country from a slightly different viewpoint is from Kotler and Gertner (2002), who propose treating a country as a brand and maintaining and analyzing the image in a manner similar to that of a product.

The foundations for the sentiment analysis and the LSTM neural networks are provided by the domain of computer science. Pang and Lee (2008) created the first comprehensive survey of sentiment analysis studies. There are many approaches to analyze the sentiment in different types of text, starting from simple approaches, such as counting the occurrence of words and n-grams that are characteristic for positive and negative texts and Naïve Bayes, and then going to more advanced machine learning techniques like logistic regression, support vector machines, conditional random fields, and artificial neural networks. Even simple algorithms like Naïve Bayes perform well enough in many cases to distinguish positive from negative texts. Domingos and Pazzani (1997), Endres (2003), and Potts (2011) even show that 85% of correct sentiment predictions could be reached when just negative and positive texts need to be distinguished. Liu (2015) provides an excellent summary of current sentiment analysis methodologies. Zhang, Wang and Liu (2018) created a survey of the application of deep learning for sentiment analysis. Their study shows that the approach of this study is state of the art.

This study uses a LSTM artificial neural network, as proposed by Hochreiter and Schmidhuber (1997) and Gers, Schmidhuber, and Cummins (2000). This is a state-of-the-art supervised learning algorithm. To enhance the performance of the sentiment analysis algorithm, the data vectorization is also done with state-of-the-art word vectorization models, as proposed by Mikolov, Chen, Corrado and Dean (2013). The methodology that is used, therefore, leads to up-to-date results, as the latest findings in computer science are applied to this study.

Other studies in the domain of sentiment analysis and the impact on stock and financial markets are often from the field of applied computer science. A more comprehensive review of these studies is given in Starosta et al. (2019).

The methodology of some studies from the fields of economics provides further foundations for this study. As this study focuses on macroeconomic phenomena and not on marketing or microeconomic tourism specific questions, some economic studies are more similar to this study than any research in the tourism domain. The studies from Daas and Puts (2014), Förschler and Alfano (2017), and Shapiro, Sudhof and Wilson (2017) try to forecast or nowcast economic indicators with the help of sentiments in the media. Daas and Puts (2014) forecast the Dutch consumer confidence with the help of messages in social media. Förschler and Alfano (2017) analyze the correlations between financial news and the leading German indices. Shapiro et al. (2017) from the Federal Reserve Bank of San Francisco show that some business cycle indicators strongly correlate with their proprietary sentiment indices. The methodology in this research is similar to the study of Daas and Puts (2014), Shapiro et al. (2017), Förschler and Alfano (2017), and Starosta et al. (2019), but uses a more recent methodology and algorithms and has a different focus. While these studies try to find correlations between media sentiments and economic indicators or economic phenomena, this study is a multi-language study that tries to measure the differences in the perceptions between the UK and Germany.

Research question

Most differences in the tourism flows between Germany and the UK and the differences pertaining to their favorite tourist destinations can be explained by classical tourism demand models and existing explanatory variables in the current literature like disposable income, distance to destination, and many others (summarized by Song and Li (2008)). Further differences can be explained by historical assumptions like past immigration flows. This study tries to clarify whether there are further differences in the perceptions of German and British travelers that contribute to the divergences in the tourist flows and the favorite tourist destinations.

The null hypothesis is that there is no big difference in the perceptions of British and German travelers and that the differences in tourism flows can be explained by the existing assumptions.

H0: The perceptions of popular destinations do not differ between British and German travelers and the differences in tourism flows are explained exhaustively by the existing models and assumptions.

The media in the UK and Germany state that they are objective, no news is deliberately kept secret, and that the most important news is delivered in a timely manner.

Hence, the hypothesis seems reasonable, and news reporting in UK and Germany should be similar, and, so, the perceptions that are conveyed by such news reporting should not differ much.

Methodology

The study is divided into several steps of data acquisition and data analysis. These are explained in the following sections.

Data sources

To analyze the differences in tourist arrivals, we used the tourism statistics data of the national statistics offices GfK Marktforschung Germany and Office for National Statistics UK. Table 1 lists the most important tourist destinations for travelers from the UK and Germany. Apart from the most important tourist destinations, we analyzed the different perceptions of some more popular European tourist destinations, as shown in Table 3. To conduct the analysis of the differences between the favorite destinations, we accounted for overnight stays, rather than the trips or expenditures. We expected that overnights were the most important factor for our study, as they define the interest of travelers to stay for a longer period. In contrast, the number of trips might be biased by business trips, and expenditures are dependent on the price level in the destination country.

Table 1
Favorite tourist destinations (nights) 2018

Country	1	2	3	4	5	Non-domestic nights (thousand) (%)
UK	Spain (11.5)	France (6.6)	USA (4.7)	India (2.3)	Italy (2.3)	563,945 (63.8)
Germany	Spain (12.7)	Italy (9.6)	Austria (4.7)	Turkey (3.7)	Greece (3.6)	798,017 (59.1)

Source: For Germany: GfK Marktforschung (2019); For UK: Office for National Statistics (2019).

To analyze the media sentiments, we sourced news reports from publicly available news streams in the UK and Germany. In the period between January 2010 and July 2017, there were 1,464,689 news items available in the German news stream and 1,003,265 in the UK news stream. The news streams were broadly based, and covered various topics similar to the daily press. Table 2 displays the analysed sources and their number of articles.

Table 2
Analyzed news sources and number of articles

Country	Source	Website	Number of articles
UK	Daily mail	www.dailymail.co.uk	571,412
UK	Financial times (UK version)	www.ft.com	349,957
UK	Guardian	www.theguardian.com	81,896
Germany	FAZ	www.faz.net	430,958
Germany	Zeit	www.zeit.de	250,383
Germany	Finanzen.net	www.finanzen.net	783,348

Data selection

All news articles that contained the name of the destination country at least three times were considered as being somehow related to the destination country. Table 3 shows the details and the number of retrieved news items for each of the tourist destinations.

Table 3
Available news for favourite tourist destinations

Destination	Number of articles in stream Jan. 2010 – Jul. 2017 (Total 1,003,265) UK	Number of articles in stream Jan. 2010 – Jul. 2017 (Total 1,464,689) Germany
Austria	630	7,545
Croatia	152	253
Egypt	912	2,358
France	6,388	6,709
Greece	5,259	18,802
India	8,323	2,300
Italy	4,677	9,841
Portugal	1,416	2,419
Spain	4,060	5,766
Tunisia	544	229
Turkey	2,949	4,585
USA	80,505	78,175

Analysis of perceptions in the media

The analysis of the perceptions in respect of every news item in the news corpus of the study was conducted by using a supervised learning algorithm, a Long Short Term Memory (LSTM) artificial neural network. The LSTM network learns the characteristics of positive and negative news in English and German media, based on training samples.

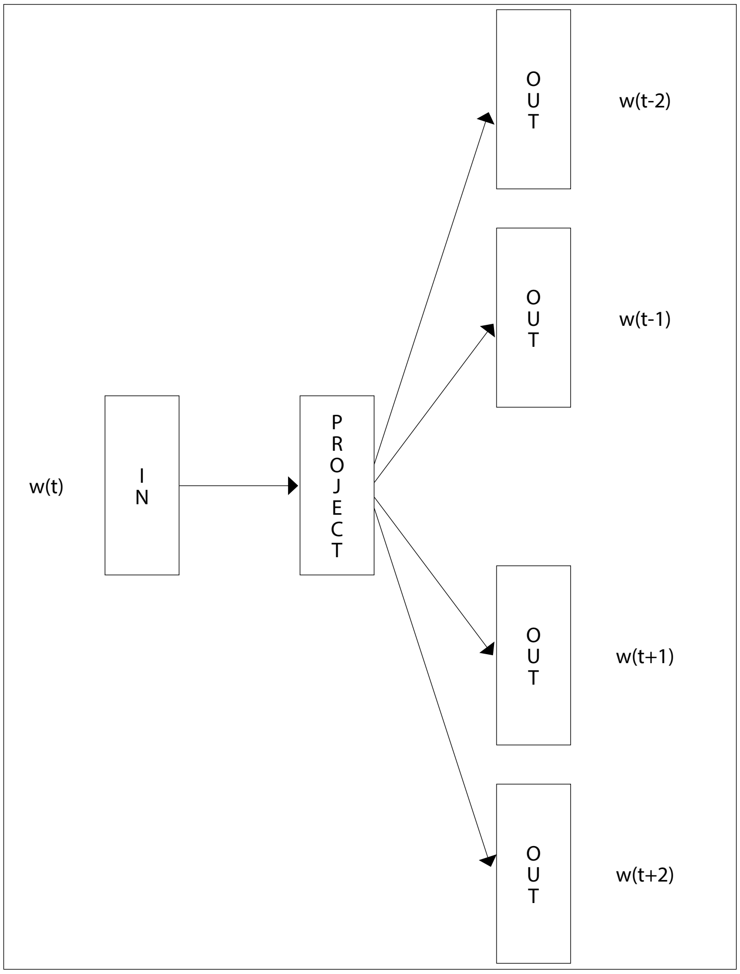
To learn the characteristics of positive and negative English and German news, we used a gold standard training corpus of 3,000 positive and 3,000 negative news items each in English and German. These were created with Amazon MTurk and manual sampling inspections by the authors.

The sentiment analysis pipeline consists of two major steps — vectorization of the input and training data and the machine-learning itself — which will be explained in detail in the following sections.

To generate machine-readable text that can be utilized by machine learning algorithms, a vectorization of the words and n-grams of a text is necessary. We vectorize our data with the help of the Word2Vec model, as proposed by Mikolov et al. (2013). The Word2Vec algorithm is a two-layer neural network that transforms words into numerical vectors of 300 dimensions. The purpose and usefulness of Word2vec is to group the vectors of similar words together in a vector space. It mathematically detects similarities, based on the distance of the words in the training corpus. To detect the distance of the words, we use the Skip-gram algorithm, which uses a word to predict the context around it (see Mikolov et al., 2013). Figure 1 shows the projection from the current word (w(t)) to the words in the context (w(t±n)). The words with their corresponding context are modeled as time-series because they are arranged by the time of their appearance.

If a feature vector that represents a word cannot accurately predict its context, the components of the vector are optimized by the neural network to better represent the context. While 300 dimensions seems to be a large space, it is still much more compact than most sentiment algorithms, which use an even more sparse space, where every word has its own dimension.

Figure 1
Skip-gram algorithm



To train the Word2Vec model, we use the English and German Wikipedia dumps and the available news in our corpus (also news articles without any country name, see Table 2).

Manual sampling inspections of the model show that similar or complementing words are close to each other in the trained vector space and that popular tests published in Mikolov et al. (2013) work. The words oak, elm, and birch cluster in one corner, while war, conflict, and strife huddle together in another. As a test of the calculation, the closest vector to vector ("Berlin") - vector ("Germany") + vector ("France") is vector ("Paris"). We now use the Word2Vec thus created to vectorize the news articles. The word vectorization using the Word2Vec algorithm has several advantages. The vectors of the Word2Vec algorithm already provide a deeper textual understanding than most other vectorizations as they group the word vectors in meaningful ways. No further stemming of the words is needed, since similar words get assigned similar vectors even in a more reasonable way than what the Porter stemmer (Porter 2008) can provide. Word2Vec averts many vocabulary quirks and is the most recent recommended way of vocabulary creation and word vectorization. It is also recommended by the Eclipse Deeplearning4j Development Team (2018). In the LSTM neural network for the sentiment analysis, we use the Word2Vec word vectors as input feature vectors.

For the sentiment analysis, we use a standard LSTM neural network. A LSTM network is a recurrent neural network that can take sequential data — such as text in the course of this study — as input data; therefore, we treat the news articles in the further steps as time-series of words. At each word, the neural network ingests the features — word vectors created in previous step — and the output is passed to the next word. While the output could also be a time-series of words that represents the sentiments of each word in the sequence of a whole article, we are only interested in the sentiments at the end of the articles.

The advantage of sentiment analysis of LSTMs compared to other neural network architectures is that it can capture long dependencies in sentences and documents and, so, provide a deep textual understanding of the sentiment analysis. A LSTM neuron can selectively choose what it should remember and what it can forget. It is, for instance, possible for the LSTM to take the previous weight states unchanged in the occurrence of a word that carries no sentiment information. This might happen if the current word is just a filler word or a stop word. This is how the long-term memory is realized. With the forget gate, the LSTM can also decide to forget the previous state. The LSTM is also able to combine the previous state and the current state and interpolate a new state.

Figure 2
LSTM cell

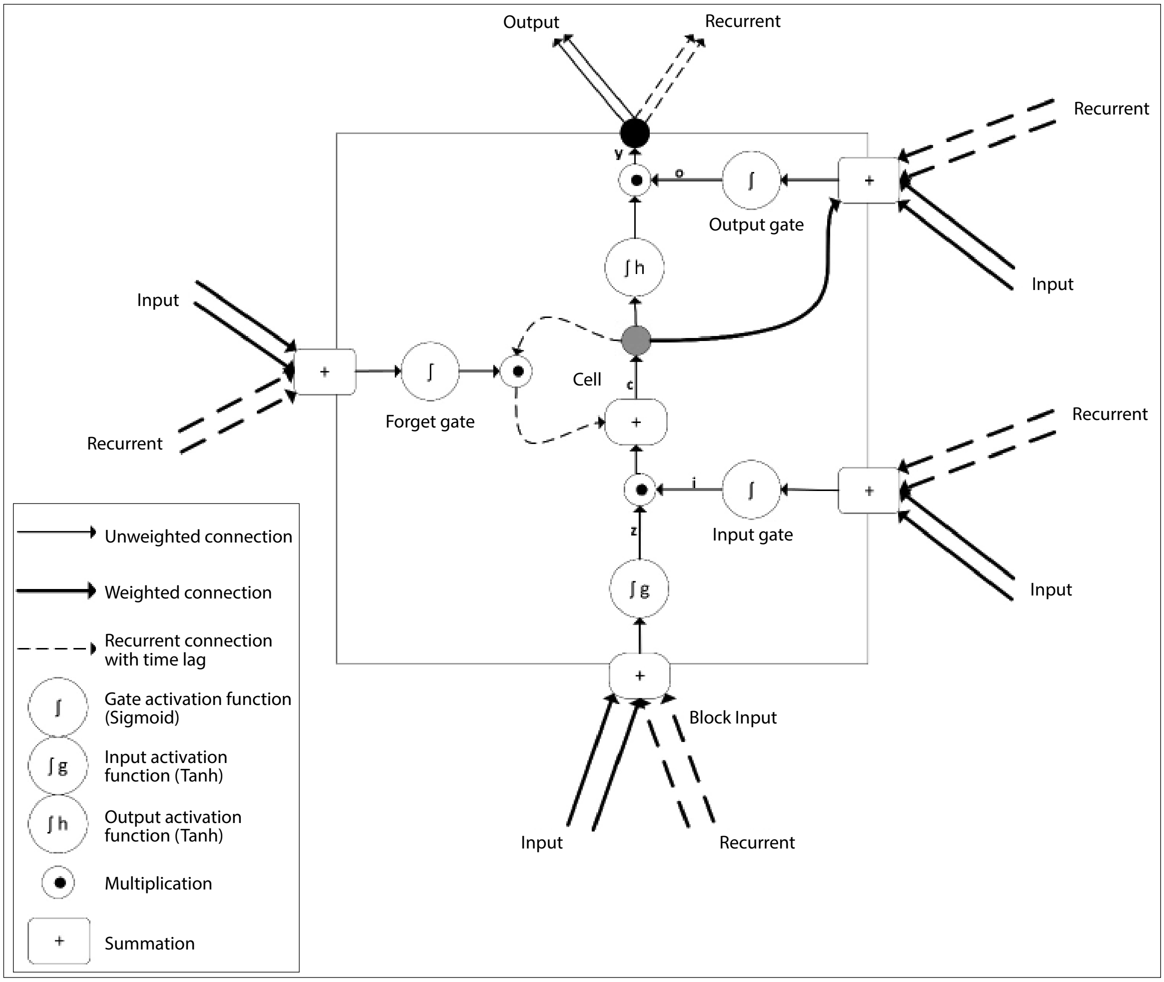


Figure 2 displays a standard LSTM cell or neuron. In order to understand how the state of the network is updated (or not), the following steps will explain the process in depth.

First, the network computes the information for the input gate i_t, which defines how much information can pass the input gate (see Equation (1)). How much information can pass the input gate is defined by the input weight matrices W_i and U_i, the current input x_t, the recurrent input y_{t-1}, and the input bias vector b_i.

$$i_t = \sigma (W_i x_t + U_i y_{t-1} + b_i) \qquad (1)$$

What information should pass the input gate is defined by the memory cell input z_t (see Equation (2)), which depends on the cell input weight matrices W_c and U_c, the current input x_t, the recurrent input y_{t-1}, and the cell input bias vector b_c.

$$z_t = \tanh(W_c x_t + U_c y_{t-1} + b_c) \qquad (2)$$

Apart from the calculation as to how much information from the input gate should be used to update the network, another question is how much of the current cell state can be discarded or forgotten because it carries no information for our problem. The forget gate defines how much information can be discarded (see Equation (3)). This is again influenced by the forget input weight matrices W_f and U_f, the current input x_t, the recurrent input y_{t-1}, and the forget input bias vector b_f.

$$f_t = \sigma(W_f x_t + U_f y_{t-1} + b_f) \qquad (3)$$

With the above parameters, the LSTM network can calculate the new cell state c_t with Equation (4).

$$c_t = i_t * z_t + f_t * c_{t-1} \qquad (4)$$

How much information can pass the output gate is defined by the output weight matrices W_o, U_o, and V_o, the current input x_t, the recurrent input y_{t-1}, the current cell state c_t, and the input bias vector b_o (see Equation (5).

$$o_t = \sigma(W_o x_t + U_o y_{t-1} + V_0 c_t + b_o) \qquad (5)$$

Finally, the new output is defined by Equation (6).

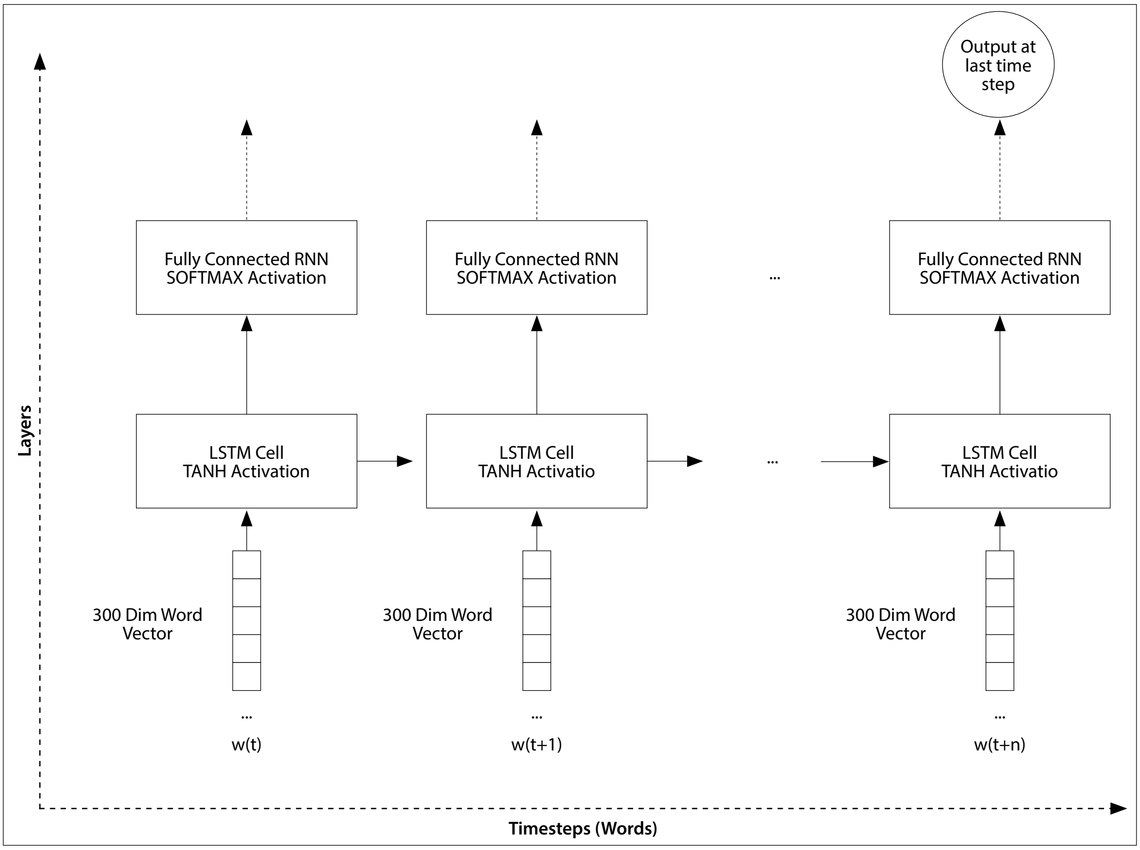
$$y_t = o_t * \tanh(c_t) \qquad (6)$$

After determining the weight matrices for all neurons with the help of the gold standard training corpus and the truncated backpropagation through time (BPTT) algorithm — as described by Williams and Peng (1990) — any news article can be classified with the LSTM network.

Figure 3 shows the network architecture unrolled over all time steps (words). The LSTM takes an input vector of 300 dimensions which is the word vector created by the Word2Vec algorithm. The LSTM step reduce the dimensions to 256, which are passed to a recurrent neural network (RNN) that calculates the output on the two output gates – one for each category. The maximum number of time steps (words) used in this study is 1000. News articles longer than 1000 words were truncated. We derived this architecture from Patterson and Gibson (2017) and Martin et al. (2018) and we could confirm the very good characteristics of it with our evaluation methodology.

We performed the evaluation of the algorithm with the help of a 6-fold cross-validation. For the cross-validation we split up the gold standard training corpus of 3000 articles into six chunks of 500 articles. We then used 5 chunks to train the network and the remaining chunk to test the trained algorithm. This procedure was done was done until all chunks were tested. The aggregated results for both languages show that this advanced sentiment analysis methodology leads to 89.4% correct classifications while all confusion matrices of the cross-validation show a very low dispersion. Continuous sampling inspections confirm that this performance is also reached outside the training set.

Figure 3
Network architecture



Building perception index

To directly compare the perceptions in the German and British media over a longer period, we build a perception index that shows the media perceptions over time.

With the help of the algorithm from the previous section, we were able to classify all the news items that we found in Table 3, either with the English model or with the German model. With this information, we were able to build four time-series for each destination country — one for the number of positive and one for the number of negative news items for each language. To create comparable data, we aggregated the positive and negative news series of each destination country to an overall sentiment series of the country.

$$Idx_t^{Fraction} = Idx_{t-1}^{Fraction} + \left(\frac{N_t^{positive}}{N_t^{total}} - \frac{N_t^{negative}}{N_t^{total}} \right) \qquad (7)$$

Equation 7 shows how we calculated the perception time series for each country.

Correlation and Regression Analysis

To verify or falsify the hypothesis, we conducted a correlation analysis between the sentiment indices of the UK and Germany for each destination.

To carry out this analysis, we used the ordinary least squares estimator [OLS], as displayed in Equation (8).

$$y = \beta_0 + Idx_t\beta_1 + \varepsilon \qquad (8)$$

where Idx_t is the index data of Equation (7). However, as there are heteroscedasticity and autocorrelations in our time series, and it is not a reasonable approach to create different models for different countries, we used Newey-West standard errors to address the problems that arise with OLS estimators because of the existence of these properties.

To measure the goodness of fit between the British and German sentiment indices for each destination, we used the coefficient of determination adjusted by the degrees of freedom [$Adj\ R^2$].

To verify the null hypothesis, the indices should correlate strongly, and we reject the null hypothesis H0 if the $Adj\ R^2 < 0.5$.

Results

Table 4 shows the results for the popular tourist destinations under observation. The table displays the strength of the correlation between the British and German sentiment indices for each destination country. In most cases, there is a strong correlation between the perception indices, which shows that there is a strong agreement on the topics and their interpretation in the British and German media. This is the behavior our null hypothesis expects and what should happen if the news reporting is neutral and objective. In contrast, the verification of the hypothesis failed for Austria, Croatia, Egypt, and Turkey. (Note that we reject the null hypothesis when the correlation is weak). Figure 4 and Figure 5 illustrate the detailed behavior of the sentiment time-series in comparison to the tourist arrivals and the number of news items. The green and purple lines display the sentiment in the German and the UK news reports for which we conducted the correlation analysis. The black lines are only for information and show the tourist arrivals as published by the Ministry of Tourism and Culture, Turkey (2019) and Statistics Austria (2019). The bar charts display the number of relevant news items in the news corpus. These exemplary figures clearly show the differences in British and German media reporting for Austria and Turkey.

Table 4
Sentiment correlation between British and German media

Destination	*Adj* r^2
Austria	0.351
Croatia	0.362
Egypt	0.489
France	0.814
Greece	0.993
India	0.952
Italy	0.841
Portugal	0.981
Spain	0.647
Tunisia	0.861
Turkey	0.184
USA	0.994

Figure 4
Time-series of different perceptions of Austria in Germany and the UK

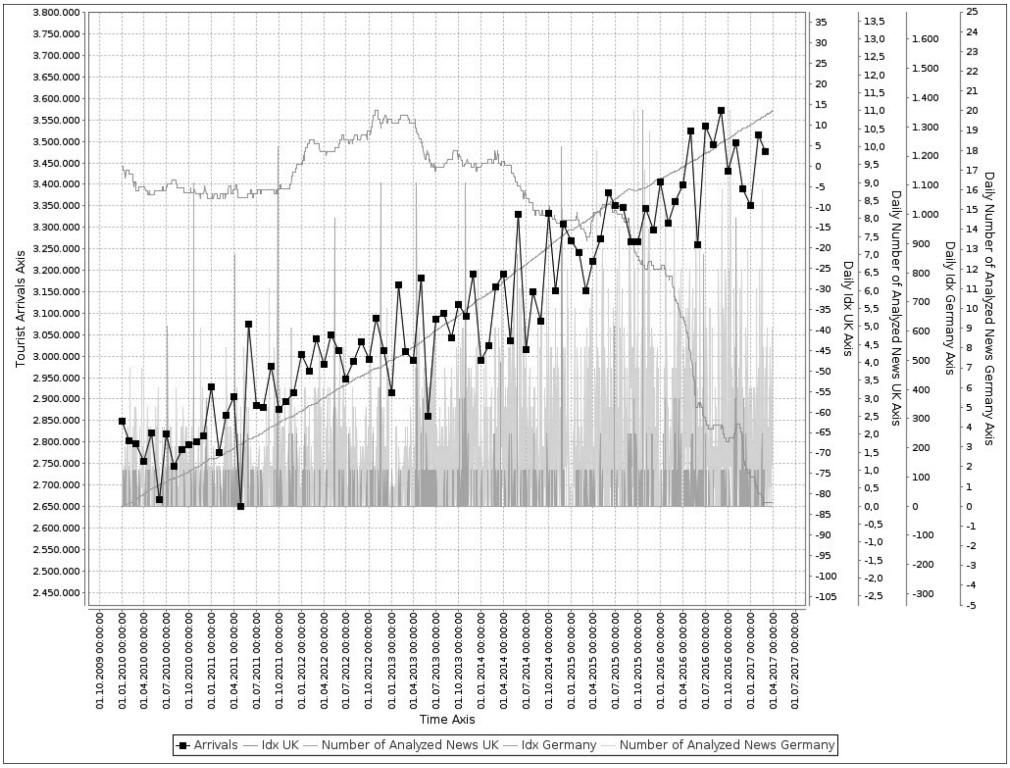
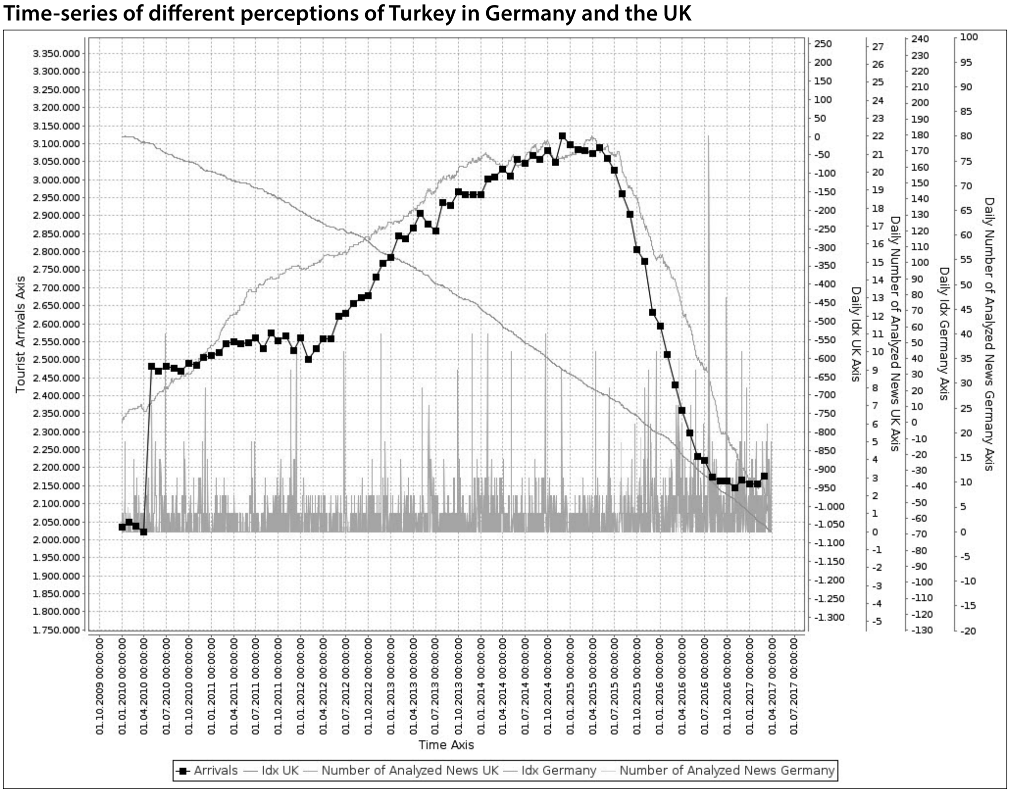


Figure 5:
Time-series of different perceptions of Turkey in Germany and the UK



Discussion

The null hypothesis could indeed be rejected for some countries and there are indeed considerable differences in the media reporting on some countries in the British and German media. We could show that there is a different sentiment and different perceptions in the media reporting for the same events in Austria, Turkey, Croatia, and Egypt.

The different perceptions do not just appear in the media reporting, but are reflected in the different popularities of destinations between British and German travelers as well (see Table 1). For Austria and Turkey — which are the third and fourth most popular destinations for German travelers — the null hypothesis can be rejected. The popularity of these countries is much lower for British travelers and they are not in the top destinations at all. While Austria only has a few British travelers, Turkey is an important tourist destination for British travelers as well and is in the top 10 list of top destinations in the UK. Apart from these very important destinations, our analysis also rejects the null hypotheses for the less important destinations, Croatia and Egypt. The following sections will elaborate the reasons for the rejection of the null hypothesis.

The sentiment in the news coverage of Austria in the UK news was relatively neutral between 2009 and 2015, but then it took a negative turn. This started with the beginning of the European migrant crisis. Austria — as one location on the "Balkan" route for refugees from the Middle East and also a target country of asylum seekers — was heavily affected by the European migrant crisis. While the news reporting in the UK was negative and the sentiment indices persistently deteriorated until the end of the observation, the sentiment in the German media did not react at all by the decisions and the actions of the Austrian government during the migration crisis. Instead, the German sentiment indices toward Austria kept increasing. As a closer look shows, the opinions drifted apart because of different judgments in the British and German media with regard to the refugee policies implemented by Austria. Interestingly, the sentiment in the media changed in a profound way and the negative news reporting continued till the end of the observation, it is possible that further factors contributed to the negativity in the news in the UK. Among others, there might be Brexit-related factors that further divided the perceptions.

The movements of the sentiment indices for Turkey are different. The sentiment index that reflects the sentiment in the German online media perfectly fits the Turkish tourist arrivals (0.899), which is an interesting fact, considering the political turmoil in Turkey in the period of observation. In contrast, the news reporting in the UK media was very negative even before the political turmoil. The autocratic shift that happened in Turkey before — and with an accelerated pace after the coup attempt in July 2016 — had a strong impact on the perceptions in the German news, with heavily deteriorating sentiment indices. The UK media had bad news reporting about the autocratic shift as well, but it was not worse than the bad news reporting before, so the sentiment indices for Turkey in the UK were decreasing at the same pace as before the autocratic shift; no change in the direction or the slope of the indices could be identified. Even if one explanation might be that Germany has a large number of citizens of Turkish origin (2,851,000 [2015], 4% of the population) — which will lead to some bias in the analysis — the results are still interesting. In a Europe with unbiased and objective news, the news reporting in a country should not depend on the origin of its citizens. Vice versa, too, the number and strength of prejudices and stereotypes in respect of a destination country should not be smaller or stronger, depending on the country of origin.

The behavior of the sentiment indices is very similar for Croatia. While the sentiment in the German media reports was positive most of the time in the observation period and only worsened when

Croatia closed its borders to Serbia — because of the European migrant crisis and the refugee flows on the "Balkan" route — the sentiment in the British media reports was negative during the entire observation period.

The last interesting case is Egypt, where many events occurred during the observation period, which also contained the Arab Spring: The overthrow of President Mubarak in February 2011, the overthrow of President Mursi in July 2013, and the bombing of a flight in October 2015. The sentiment in the German media reacted very strongly to every event and the media presence of the events was very high. In comparison, the media in the UK had a negative news reporting long before the events and the worsening of the sentiment indices did not speed up with any of these events. Interestingly, in other cases, the sentiment of the British media (and the corresponding sentiment indices) did also react to single events, so, in general, this behavior is not unique to the German media. Rather, this is a phenomenon for this specific Egyptian case. The reasons might be a much lower news presence of Egypt in British news and, so, a much lower interest or a different judgment of the events in the British media.

In general, the UK and Germany have the same perceptions about most of the popular tourist destinations for Europeans and the null hypothesis holds for most countries. In particular, the very high correlation for the USA and Greece (> 0.99) shows that a high media coverage and many data points might lead to accurate perception data and a strong consensus. It might even show that the greater the number of news articles, the more accurate the overall perception of a country. Also, probably, stereotypes and prejudices might be wiped away in the presence of the much detailed information of many news items.

Interestingly, there are a few destinations with regard to which the perceptions differ greatly. While some differences can be attributed to different judgments of the same events — as the worsening sentiment for Austria in the British media in the aftermath of the refugee policy changes showed — other root causes remain unclear, and cannot be elaborated in the course of this study.

Furthermore, the causal relationship between tourist arrivals and different perceptions does not seem to be clear. Are there different perceptions caused by fewer travelers or are there fewer travelers because of different perceptions? The study shows that more tourist arrivals come along with more news in the countries of origin. Still, the different perceptions cannot be explained with the different number of news and the differences in the media coverage, as a comparison with Table 3 shows.

Conclusions

We are able to demonstrate that the perceptions of two of the most popular tourist destinations of the UK and Germany — Austria and Turkey — are different for British and German travelers, and that there are also further differences for other countries. While this is an interesting finding, it is equally interesting that there is a strong consensus in perceptions for most of the countries that were analyzed. This shows, among other things, the common values and norms of the UK and Germany, since the same perception also requires the same judgments.

While we are able to show that there is a clear relationship between media perceptions and tourist arrivals, the direction of the causal relationship is less clear, and needs to be elaborated in future studies. We can show that the media reports that set the agenda for the citizens and forms their perceptions of the world have a strong relationship with many aspects of the tourism business and other sectors.

The strong correlations between most of the country indices analyzed for the UK and Germany demonstrated that we found a robust approach of measuring perceptions in a cost-effective way. This case study can be extended to many other geographical areas and business sectors in order to further understand how perceptions drive opportunities and create threats for businesses, politics, and policymakers, and also probably how to mitigate these risks. The comparison of the different perceptions among different countries helps us understand the origins of these different perceptions. Are they based on prejudices or stereotypes? Are they based on a too-optimistic or a too-pessimistic picture? Or are they just based on too little, or wrong information? The strong consensus of the sentiment in the media about the USA and Greece (> 0.99), which were the countries with the most number of news items, can even be an indication that more information leads to less divergent perceptions. Answering these questions might help businesses obtain better insights into their global operations as well as enable policymakers to query the international perceptions of their policies. Finally, and importantly, our approach of comparing perceptions might bring new insights into the fake news and post-truth debate.

Declaration of conflicting interests

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

Brewer, P. R., Graf, J. & Willnat, L. (2003). *Priming or Framing: Media Influence on Attitudes toward Foreign Countries*. DOI: https://doi.org/10.1177/0016549203065006005.

Chaabani, Y., Toujani, R. & Akaichi, J. (2017). Sentiment Analysis Method for Tracking Touristic Reviews in Social Media Network. In G. De Pietro, L. Gallo, R. Howlett & L. Jain (eds), *Intelligent Interactive Multimedia Systems and Services 2017. KES-IIMSS 18, Vol. 76* (pp. 299-310). Cham: Springer. DOI: https://doi.org/10.1007/978-3-319-59480-4_30.

Daas, P. & Puts, M. (2014). Social Media Sentiment and Consumer Confidence. In *ECB Statistics Paper 5*. Frankfurt am Main: European Central Bank.

Dinis, M. G. F., Costa, C. M. M. & Pacheco, O. M. R. (2017). Similarities and Correlation between Resident Tourist Overnights and Google Trends Information in Portugal and Its Tourism Regions. Tourism & Management Studies, 13(3), 15-22. Retrieved October 30, 2017, from http://www.tmstudies.net/index.php/ectms/article/view/940.

Dogru, T., Sirakaya-Turk, E. & Crouch, G. I. (2017). Remodeling International Tourism Demand: Old Theory and New Evidence. *Tourism Management, 60*, 47–55. DOI: https://doi.org/10.1016/j.tourman.2016.11.010.

Domingos, P. & Pazzani, M. J. (1997). On the Optimality of the Simple Bayesian Classifier under Zero-one Loss. *Machine Learning, 29*, 103–130. Retrieved October 12, 2015, from http://www.ics.uci.edu/~pazzani/Publications/mlj97-pedro.pdf.

Eclipse Deeplearning4j Development Team. (2018). *Deeplearning4j: Open-source distributed deep learning for the JVM, Apache Software Foundation License 2.0.* Retrieved from http://deeplearning4j.org.

Endres, J. (2003). *Spam oder nicht Spam? E-Mail sortieren mit Bayes-Filtern*. c't 17/2003, p. 150.

Förschler, F. & Alfano, S. (2017). Reading between the Lines: The Effect of Language Sentiment on Economic Indicators. In *Enterprise Applications, Markets and Services in the Finance Industry* (pp. 89–104). Cham: Springer Press.

Gers, F., Schmidhuber, J. & Cummins, F. (2000). Learning to Forget: Continual Prediction with LSTM. *Neural computation, 12*(10), 2451–2471. DOI: https://doi.org/10.1162/089976600300015015.

GfK Marktforschung. (2019). *Tourismusanalyse 2019*. Retrieved July 23, 2019, from http://www.tourismusanalyse.de/zahlen/daten/statistik/tourismus-urlaub-reisen/2019/reiseabsichten-2019/.

González-Rodríguez, M. R., Martinez-Torres, R. & Toral, S. (2016). Post-visit and Pre-visit Tourist Destination Image through eWOM Sentiment Analysis and Perceived Helpfulness. *International Journal of Contemporary Hospitality Management, 28*(11) 2609–2627. DOI: https://doi.org/10.1108/IJCHM-02-2015-0057.

Hochreiter, S. & Schmidhuber, J. (1997). Long short-term memory. *Neural computation, 9*(8), 1735–1780. DOI: https://doi.org/10.1162/neco.1997.9.8.1735.

Howard, J. A. (1963). *Marketing Management* (2nd ed). Homewood, Illinois: Irwing Publishing Company.

Kotler, P. & Gertner, D. (2002). Country as Brand, Product, and Beyond: A Place Marketing and Brand Management Perspective. *Journal of Brand Management, 9*, 249–261. DOI: https://doi.org/10.1057/palgrave.bm.2540076.

Lattin, J. H. & Roberts, J. M. (1991). Development and Testing of a Model of Consideration Set Composition, in Journal of Marketing Research, Vol. XXVIII, November 1991, pp. 429–440.

Law, R., Li, G., Fong, D. K. C. & Han, X. (2019). Tourism demand forecasting: A deep learning approach. *Annals of Tourism Research, 75*, 410-423. DOI: https://doi.org/10.1016/j.annals.2019.01.014.

Li, X., Pan, B., Law, R. & Huang, X. (2017). Forecasting Tourism Demand with Composite Search Index. *Tourism Management, 59*, 57–66. DOI: https://doi.org/10.1016/j.tourman.2016.07.005

Li, Y. & Cao, H. (2018). Prediction for Tourism Flow based on LSTM Neural Network. *Procedia Computer Science, 129*, 277-283. DOI: https://doi.org/10.1016/j.procs.2018.03.076.

Liu, B. (2015). *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions* (1st ed.). Cambridge University Press.

Martin, C. A., Torres, J. M., Aguilar, R. M. & Diaz, S. (2018). Using Deep Learning to Predict Sentiments: Case Study in Tourism. *Complexity, 2018*. DOI: https://doi.org/10.1155/2018/7408431.

Mathieson, A. & Wall, G. (1982). Tourism: Economic, Physical, and Social Impacts. Harlow: Longman Group Limited, Longman House, Burnt Mill.

McCombs, M. & Reynolds, A. (2002). News Influence on our Pictures of the World. In J. Bryant & D. Zillmann (Eds.), *Media Effects: Advances in Theory and Research* (2nd ed., pp. 1-18). Mahwah: LEA.

Mikolov, T., Chen, K., Corrado, G. & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. *Proceedings of Workshop at ICLR. arXiv:1301.3781v1*. Retrieved from https://arxiv.org/abs/1301.3781.

Ministry of Tourism and Culture, Turkey (2019). *Tourism Statistics*. Retrieved July 16, 2019, from https://www.ktb.gov.tr/EN-153017/tourism-statistics.html.

Moufakkir, O., 2014. What's Immigration Got to Do with It? Immigrant Animosity and Its Effects on Tourism. *Annals of Tourism Research, 49*, 108-121. DOI: https://doi.org/10.1016/j.annals.2014.08.008.

Office for National Statistics (2019). *Leisure and Tourism Statistics 2019*. Retrieved July 24, 2019, from https://www.ons.gov.uk/peoplepopulationandcommunity/leisureandtourism

Zeitzoff, T. (2016). Does Social Media Influence Conflict? Evidence from the 2012 Gaza Conflict. *Journal of Conflict Resolution*, *62*(1) 29-63. DOI: https://doi.org/10.1177/0022002716650925.

Patterson, J. & Gibson, A. (2017). *Deep Learning* (1st Ed.). Sebastopol: O'Reilly Media.

Pang, B. & Lee, L. (2008). Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval, 2*(1–2), 1–135. DOI: http://dx.doi.org/10.1561/1500000011.

Potts, C. (2011). *Sentiment Symposium Tutorial: Classifiers*. Retrieved February 3, 2015, from http://sentiment.christopherpotts.net/classifiers.html.

Ren, G. & Hong, T. (2017). Investigating Online Destination Images Using a Topic-Based Sentiment Analysis Approach. *Sustainability 2017*. DOI: http://dx.doi.org/10.3390/su9101765.

Rizal, A., Soraya, S. & Tajuddin, M. (2019). Sequence to sequence analysis with long short term memory for tourist arrivals prediction. *Journal of Physics: Conf. Series, 1211*. DOI: http://dx.doi.org/10.1088/1742-6596/1211/1/012024.

Seetaram, N. (2012). *Immigration and International Inbound Tourism: Empirical Evidence from Australia*. DOI: https://doi.org/10.1016/j.tourman.2012.02.010.

Shapiro, A., Sudhof, M. & Wilson, D. J. (2017). Measuring News Sentiment. *Federal Reserve Bank of San Francisco Working Paper 2017-01*. Retrieved February 16, 2017, from http://www.frbsf.org/economic-research/publications/working-papers/wp2017-01.pdf.

Sirakaya, E. & Woodside, A. G. (2005). Building and Testing Theories of Decision Making By Travellers. *Tourism Management, 26*(6), 815–832.

Sönmez, S. F. & Graefe, A.R. (1998). Influence of Terrorism Risk on Foreign Tourism Decisions. Annals of Tourism Research, 25(1), 112–144. DOI: https://doi.org/10.1016/S0160-7383(97)00072-8.

Starosta, K., Budz, S. & Krutwig, M. (2019). The impact of German-speaking online media on tourist arrivals in popular tourist destinations for Europeans. *Applied Economics, 51*(14), 1558-1573. DOI: 10.1080/00036846.2018.1527463.

Starosta, K., Budz, S. & Krutwig, M. (2018). Artificial Neural-Network-Based Emotion Classification in the Online Media for Tourism Businesses. *Proceedings of the 5th European Conference on Social Media*. Limerick 2018.

Statistics Austria. (2019). *Tourismus 2019*. Retrieved August 2, 2019, from https://www.statistik.at/web_de/statistiken/wirtschaft/tourismus/index.html.

Song, H. & Li, G. (2008). Tourism Demand Modelling and Forecasting — A Review of Recent Research. Tourism Management, 29(2), 203–220. DOI: https://doi.org/10.1016/j.tourman.2007.07.016.

Wanta, W., Golan, G. & Lee, C. (2004). *Agenda Setting and International News: Media Influence on Public Perceptions of Foreign Nations*. DOI: https://doi.org/10.1177/107769900408100209.

Williams, R. J. & Peng, J. (1990). An efficient gradient-based algorithm for online training of recurrent network trajectories. *Neural Computation, 2*, 490–501.

Woodside, A.G. & Sherrell, D.L. (1977). Traveler Evoked, Inept, and Inert Sets of Vacation Destinations. Journal of Travel Research, 16(1), 14-18.

Ye, Q., Zhang, Z. & Law, R. (2009). Sentiment Classification of Online Reviews to Travel Destinations by Supervised Machine-Learning Approaches. *Expert Systems with Applications, 36*(3), Part 2, 6527–6535. DOI: https://doi.org/10.1016/j.eswa.2008.07.035.

Yuan, Y., Liu, Y. & Wei, G. (2017). Exploring Inter-country Connection in Mass Media: A Case Study of China. *Computer, Environment and Urban Systems, 62*, 86–96. DOI: https://doi.org/10.1016/j.compenvurbsys.2016.10.012.

Zhang, L., Wang, S. & Liu, B. (2018). Deep learning for sentiment analysis: A survey. WIREs Data Mining Knowledge Discovery, 8(4). DOI: https://doi.org/10.1002/widm.1253.

Received: 20/11/2018
Accepted: 25/11/2019