Chapter 6
Sentiment Analysis for Tourism

Mike Thelwall

Abstract Sentiment analysis software is a key component of tourism big data research for its ability to detect positive and negative opinions in text. This supports large-scale analyses of the key affective dimension of reviews and social web posts about tourism venues and experiences. Sentiment analysis is fast and reasonably accurate, enabling patterns to be mined from large numbers of texts that would not be evident to experts reading those texts, such as differences between genders or venues in the aspects of destinations that are liked. This chapter reviews the main sentiment analysis approaches with a focus on practical descriptions of how the methods work and how they can be applied. The chapter also illustrates the value of sentiment analysis for tourism research.

Keywords Sentiment analysis · Tourism research · Social web posts · Online reviews · Tourism experiences

6.1 Introduction

Customer feedback and sentiment can help individual tourist attractions, hotels and restaurants to gain word of mouth recommendations and repeat visitors (e.g., Chen 2003) as well as to improve services (Schweidel and Moe 2014). Sentiment is at the heart of tourism because people expect to enjoy a holiday or visit. A critical part of the customer experience is therefore satisfaction: are they happy with the service that they received? When writing a review, the overall level of satisfaction may be flagged by an accompanying rating (e.g. 1–5 stars). Much feedback will not be explicitly rated, however, such as holidaymakers' tweets or Facebook posts (Philander and Zhong 2016). For big data analysis it is therefore essential to be able to detect the sentiment of this informal feedback automatically. This would reveal, for example, aspects of attractions that tend to appear in positive or negative comments.

M. Thelwall (✉)
Statistical Cybermetrics Research Group, University of Wolverhampton, Wolverhampton, UK
e-mail: m.thelwall@wlv.ac.uk

© Springer Nature Singapore Pte Ltd. 2019

M. Sigala et al. (eds.), *Big Data and Innovation in Tourism, Travel, and Hospitality*,
https://doi.org/10.1007/978-981-13-6339-9_6

Sentiment analysis software has the task of detecting and classifying opinions and feelings expressed in text. Whilst there are many challenges to be solved before sentiment analysis programs attain expert human-level performance (Cambria et al. 2017), they give results that correlate positively with user ratings and so it is reasonable to use sentiment analysis scores when user ratings are absent (López-Barbosa et al. 2015). In addition, aspect-based sentiment analysis (see below) can give finer grained evidence than overall review ratings by identifying aspects of hotels or attractions that are singled out for praise or criticism. Sentiment analysis can also give additional insights into the important components of an offering. For example, an analysis of restaurant reviews found that context was important to diners in addition to those more widely recognised: food, service, price and ambience (Gan et al. 2017).

Sentiment analysis has become a standard component of the social media analysis toolkit of marketers and customer relation managers in large organisations (Hofer-Shall 2010). These may purchase relevant data (e.g., Facebook posts mentioning a hotel) and gain access to a suite of sentiment analysis and other tools to analyse it in a platform like Pulsar (pulsarplatform.com). Understanding how sentiment analysis works is therefore important for those working in tourism and the wider marketing sector.

This chapter reviews the main sentiment analysis approaches, describes how they may be exploited to analyse customer feedback and gives some examples from tourism research. It does not cover consumer sentiment indexes (Dragouni et al. 2016), but focuses on sentiment extracted from social media content, such as online reviews and comments (e.g., Tian et al. 2016). The chapter introduces the main current issues in sentiment analysis to the non-specialist and then reviews tourism-related sentiment analysis research. More technical details that are more relevant to computer scientists are avoided but can be found in previous reviews (Liu 2012; Pang and Lee 2008).

6.2 Core Sentiment Analysis Methods

6.2.1 Lexical Sentiment Analysis

The most transparent type of sentiment analysis is known as the lexical approach (Taboada et al. 2011). This uses a dictionary of sentiment-related terms, often together with estimates of their strength. The program SentiStrength (Thelwall et al. 2010; Thelwall et al. 2012), for example, includes 2846 positive or negative words or word stems (e.g., shabby, vex*). Each is accompanied by an estimate of its polarity and average strength in common usage. SentiStrength's positive terms are given a strength rating between 2 (mildly positive) and 5 (strongly positive). Negative terms are scored from −2 (mildly negative) to −5 (strongly negative). Thus, since *tremendous* scores +4 and *overheated* scores −3, SentiStrength would classify the text below as follows (i.e., containing both moderately negative sentiment −3 and strongly positive sentiment 4):

- Tremendous hotel but the rooms were overheated. → **−3**, **4**

As this shows, SentiStrength gives a separate positive and negative score. It can be tried online at http://sentistrength.wlv.ac.uk. Other programs give a single output (e.g., positive or negative).

Sentiment can be expressed in linguistically complex ways that a lexical sentiment analysis program might try to detect with additional rules. Most importantly, sentiment can be neutralised or flipped by negation and sentiment can be strengthened or weakened with booster terms like "very". For example, with SentiStrength

- I was not happy with the room but very satisfied with the view. → **−2**, **4**

Here, *happy* scores +3 but it is negated by "not", changing its value to −2. In contrast, "satisfied" normally scores +3 but is boosted by "very" to +4.

Sentiment can be expressed typographically with emoticons or with spellings that suggest enthusiasm. A lexical program might have a list of emoticons with sentiment scores and apply rules to detect enthusiastic spelling. SentiStrength would score "haaaapy" as +4 after detecting that (a) the underlying word is "happy" with score +3 and (b) the extra "aaa" increases the strength of this word.

- At the top of the Eiffel Tower:) → **−1**, **2**
- Haaaaapy to be in the Tivoli gardens! → **−1**, **4**

Given the relative simplicity of the above procedures it is important to check whether they work well enough in practice to be useful. This can be assessed by comparing computer and human scores for the same texts that have been implicitly or explicitly coded for sentiment by humans. An explicit coding might be derived from the sentiment judgements of expert coders whereas an implicit coding might be a rating associated with a review on the TripAdvisor website. Because computer accuracy varies by text type it is best to test the text type that will be analysed (e.g., hotel reviews) or multiple text types. SentiStrength has been tested on six social web sites, correlating with human judgements at a rate of between 0.30 and 0.65 for positive sentiment strength and between 0.50 and 0.60 for negative sentiment strength. Although these correlations are moderate, humans also tend to agree only moderately with each other. The agreement rate between SentiStrength and expert coders for negative sentiment strength is lower than the agreement rate between the expert coders but higher than for typical people doing the same task. The same is true for positive sentiment strength (except that it is less accurate on political discussions due to sarcasm).

As the above might suggest, sentiment analysis software can be expected to be less accurate than experts but more accurate than random people for the task of classifying a single isolated text. Sentiment analysis is fundamentally subjective, so humans often disagree. The primary advantage of software is speed: SentiStrength can classify 14,000 tweets per second but careful humans may require a minute to judge each tweet.

The above discussion described SentiStrength but all lexical sentiment analysis programs have their own lexicon and set of additional rules. Some, like Socal (Taboada et al. 2011), use advanced linguistic text processing to identify deeper structures in sentences, such as distinguishing adjectives from other words or disambiguating sentiment terms (Baccianella et al. 2010). This can improve their accuracy, especially on grammatically correct text, but reduces their speed.

6.2.2 Machine Learning Sentiment Analysis

Sentiment analysis can be achieved without a lexicon using machine learning (Liu 2012). For this, a set of human classified texts is needed. The algorithm learns how to identify positivity and negativity by examining many examples of both. To illustrate this, the program might notice that many negative reviews contain the term "dirty" and formulate the rule that future reviews are negative if they contain this term. Machine learning sentiment analysis typically focuses on the words, bigrams (consecutive word pairs) and trigrams (consecutive word triples) in each review, rather than complete sentences. Thus, a machine learning algorithm might use words and phrases like "love the room", "good service" and "rats" to help decide whether a hotel review is positive or negative overall. In practice, state of the art machine learning algorithms use complex pattern recognition methods, such as support vector machines (SVM) or deep learning, that are not understandable by humans.

- I love the room! We arrived early and took advantage of it. → **positive**
- We had good service from all the staff—prompt and polite. Thank you! → **positive**
- Could not believe we saw rats outside the restaurant kitchen. → **negative**

An advantage of machine learning sentiment analysis is that it does not require the human effort of creating a lexicon, in contrast to lexical sentiment analysis. In addition, its sentiment classifications can be more accurate than for lexical sentiment analysis if it has a large volume of training data (human classified texts) to learn from. An important disadvantage is that the machine learning sentiment analysis rules are opaque and so the causes of incorrect classifications cannot easily be deduced and corrected. Machine learning requires a substantial set of pre-classified texts (usually at least 1000) to achieve a reasonable performance. The results tend to be tailored to the type of texts used for training. Thus, a broad variety of pre-classified texts are needed to generate a generally useful algorithm.

For tourism-related reviews, all public automatic sentiment analysis tools seem to be less accurate than human judges for three different genres of tourism related text, with machine learning methods outperforming lexical software (Kirilenko et al. in press).

6.3 Universal Sentiment Analysis Tasks and Considerations

6.3.1 The Impact of Topic Domain on Algorithm Accuracy

A sentiment analysis algorithm designed for one language would not work well for another because it would not recognise the sentiment words. The same is true to a lesser extent for text types. A machine learning algorithm that learns how to identify sentiment from a set of book reviews might identify words like *interesting* and *readable* as positive. If it is then fed with the following hotel review it is likely to classify it incorrectly. In the illustrations below, the square brackets describe the type of sentiment algorithm applied to the data.

- Finding mouse poo under the bed was interesting. [books] → **positive**

In contrast, a machine learning algorithm training on hotel reviews is likely to have worked out that mentions of vermin or excrement indicate a bad experience and realise that the review is very negative.

- Finding mouse poo under the bed was interesting. [hotels] → **negative**

A machine learning sentiment analysis system trained on tourism related reviews has achieved an accuracy of 80% for the task of deciding whether they are positive or negative overall (Ye et al. 2009), presumably by learning a range of tourism-related sentiment expressions. Another system subsequently achieved an accuracy of 90% with a domain-specific lexicon (Gräbner et al. 2012). Thus, high levels of accuracy are possible for this task.

A general-purpose sentiment analysis program should be reasonably accurate on all types of text but perhaps not as accurate as one designed for the type of text being analysed. In contrast, a domain-specific program that is designed for one type of text is likely to be the most accurate for its domain. It is possible to develop methods to translate a program from one domain to another or to detect the domain of a text when processing it (Glorot et al. 2011). This has been shown to work well in the context of Tourism in the popular Chinese review site Ctrip.com (Li et al. 2015).

6.3.2 Language

Sentiment analysis algorithms tend to be language-specific because they learn from human coded texts from one language. Thus, for example, learning that "nice room" is positive does not help with identifying sentiment in Chinese hotel reviews.

- We had a nice room, which was a relief. [English] → **positive**
- We had a nice room, which was a relief. [Chinese] → **neutral**
- 我们有一个漂亮的房间, 这是一种解脱. [English]- → **neutral**

Although most sentiment analysis research has been conducted in English, there are many algorithms in other languages. For example, sentiment systems have been designed for Spanish TripAdvisor reviews (Salas-Zárate et al. 2017), Arabic TripAdvisor reviews (Cherif et al. 2016), and Russian Olympics-related tweets (Kirilenko and Stepchenkova 2017). A generic system has also been applied to Spanish TripAdvisor and Booking.com reviews (Fondevila-Gascón et al. 2016). More generally, there are multilingual toolkits, such as OpeNER, that have different language variants (García-Pablos et al. 2016). There are also methods to generate multi-lingual applications or to translate sentiment detection algorithm from one language to another (Balahur and Turchi 2014). These can be expected to be less accurate than language-specific versions.

6.3.3 Image Sentiment Analysis

Sentiment is often expressed explicitly in images (smiling for the camera) or implicitly (photograph of blue sky and golden sand) rather than in text. Images are increasingly used for informal communication between friends (Thelwall et al. 2016) due to the affordability and availability of smartphone image sharing apps as well as simple pathways to post images in social network profiles. It is much more difficult for computers to identify sentiment in images than in text because there is more data to process and it is more complex. Nevertheless, recent advances in image processing with convolutional neural networks (e.g., Oquab et al. 2014), which mimic the human brain to some extent, have led to breakthroughs with visual sentiment analysis. It is now possible to detect sentiment in some types of image with a high degree of reliability (You et al. 2015) but the process is much slower than for text processing.

Pictures have a central role in tourism (Chalfen 1979), from promotional images to the postcard home, souvenir snapshots and selfies at attractions posted to social media (Lyu 2016). Image-based sentiment analysis therefore has the potential to open an underexplored dimension of the leisure industry.

6.3.4 Universal Sentiment Analysis Tasks

Sentiment analysis programs have different output types. The most common is probably **trinary**: each text is classified as positive, negative or neutral overall. **Sentiment strength** or intensity is also sometimes estimated, as by SentiStrength (positive and negative separately) and Socal (combined positive-negative scale). Finally, **fine grained emotion detectors** attempt to distinguish between a set of emotions, such as happiness, sadness and anger (Neviarouskaya et al. 2009). Fine grained emotion detection seems to be not common due to the difficulty to obtain reasonably accurate results.

- Builders again at 5.30am! Aaaaargggggh!!!! [trinary] → **negative**
- Builders again at 5.30am! Aaaaargggggh!!!! [SentiStrength] → **−4**, **1**
- Builders again at 5.30am! Aaaaargggggh!!!! [SoCal] → **−3**: **very negative**
- Builders again at 5.30am! Aaaaargggggh!!!! [fine] → **anger**

Aspect-based sentiment analysis goes one step further by detecting both sentiment and the sentiment object (Jo and Oh 2011). Thus, it may detect in "cold shower but comfy bed" that the shower was bad [cold = negative] and the bed was good (comfy = positive).

- Cold shower but comfy bed → **shower**: **negative**; **bed**: **positive**

An aspect-based sentiment summarisation might list the aspects of a product or service that attract positive or negative feedback, generating useful market intelligence (Gamon et al. 2005; Reis et al. 2014). For tourism, this could take the form of identifying the positive and negative terms that are commonly associated with the key aspects of a hotel, such as its view (Marrese-Taylor et al. 2013a, b).

- [lots of reviews] → **view (positive)**: **sea**, **beach**; **view (negative)**: **distance**, **blocked**

In another case, a sentiment analysis of TripAdvisor reviews of California State Parks extracted opinions about aspects including the shops, camping, rangers, roads, shops and trails (Farhadloo et al. 2016).

One clever application of aspect-based analysis is to separate the text of a review into the different aspects of an attraction discussed in the review and then to apply different algorithms to the text associated with each aspect. This combines domain adaptation to improve the overall accuracy of the sentiment analysis results (Sharma et al. 2017).

6.3.5 Accuracy and Bias

As discussed above, automatic sentiment analysis is imperfect and not always correct. Ideally, these imperfections will average out and disappear when processing many texts. For example, a program might incorrectly classify the first text below as positive but over a large set of texts might work out that the hotel was upsetting residents by losing their post.

- The hotel lost my happy birthday card. → **positive**
- They lost my parcel and it is a disgrace. → **negative**
- I don't know how they lost my letter. → **negative**

It is impossible to eliminate all sources of bias but analysts should be aware that they exist and might influence the results.

Bias occurs in sentiment analysis when the errors are systematic in a way that would influence the conclusions drawn from the results. If a program persistently classified reviews of Happy Eater restaurant food as positive due to its name then an analyst might conclude that Happy Eater food was universally loved.

- We all had fish and chips at the Happy Eater. → **positive**
- I ate at the Happy Eater. → **positive**

Bias can also be subtler. For example, there are gender differences in the expression of sentiment (Teso et al. 2018) and sentiment analysis software is more accurate for female-authorised texts (Thelwall 2018a, b; Volkova et al. 2013). This is because there is a small but statistically significant tendency for females to express sentiment more clearly, as in the first of the two reviews below.

- Thanks so much to Sheila and Keith for keeping the bar open until 2am for us! All the girls had a wonderful time! → **positive**
- The hotel kept the bar open until 2am for the lads. What more can I say. → **neutral**

A big data analysis might give more weight to the opinions of females because it can detect them better. There can be important gender differences in the opinions of tourists (Yan et al. 2018), and so gender biases in detecting sentiment may affect the overall results. In contrast, an analysis of over 20,000 reviews of restaurants in two US cities found no gender differences in average sentiment (Micu et al. 2017), so gender bias may not always be relevant.

Bias can occur when one type of customer is more likely to post, such as younger users. Hotel customers might also be more likely to post if their positive or negative experience could be attributed to a named member of staff. For this reason, services might be commented on more than bed comfort even if both were equally important. In addition, positive experiences are more likely to lead to sharing on general social media sites whereas negative experiences are more likely to lead to posting a review to an integrated tourist site, such as TripAdvisor (Yan et al. 2018).

6.4 Special Considerations for Tourism-Based Sentiment Analysis

When applying or interpreting sentiment analysis to tourist reviews there are some important generic issues.

6.4.1 Limitations of Social Media Analysis for Tourism

As discussed above, an important limitation of any form of social media analysis is that the people that post to the social web form a self-selected subset of all customers and may be a highly biases subset. Customers that have had a very good or bad experience are more likely to post a review to a website and perhaps most likely if the experience was bad. Older people and young children may also be less likely to post an online review or share their experiences in social media because they are less likely to be web users. Conversely, busy parents may be frequently online but too busy to post reviews.

There may be more subtle biases in the demographics of users, such as in favour of some ethnicities or social classes. Some reviews may be fake—perhaps malicious reviews posted by competitors or paid positive reviews to boost a new or unpopular destination. Of course, self-selection bias also exists for most survey based research but this can be minimised by effective research designs that ensure high response rates. There is no equivalent for the social web and no good internet remedy for types of people that are not well represented online. Thus, researchers should be careful to identify sources of bias and their likely effect and interpret the results in the light of these.

Another limitation of online research is that it lacks evidence of direct connections with relevant actions, such as online bookings or purchases made because of the opinions expressed (Schuckert et al. 2015). Thus, whilst it seems reasonable to assume that negative reviews would not be good for an attraction, it is hard to quantify their impact in terms of reduced visits.

6.4.2 Tourism Domains

As discussed above, sentiment analysis programs can be general purpose or domain (i.e., topic) specific, with domain-specific applications likely to be the most accurate. In the context of tourism, a machine learning algorithm that had trained on mobile phone data might know that "expensive" is bad without learning whether more hotel-specific terms or phrases, such as "clean", "unfriendly" and "welcoming receptionist" expressed sentiment. Whilst a general-purpose sentiment analysis algorithm should give reasonably accurate results if enough customers use common generic sentiment terms (e.g., "excellent", "good value"), learning domain-specific terminology can improve performance. This even applies between different types of tourism offering. Whilst common sentiment terms for a restaurant would also apply to a hotel because hotels offer food, the reverse is true to a lesser extent. For other attractions, such as beaches, arcades and cultural activities, the key sentiment terms might be quite different.

- A sunny day but had expensive <u>mouldy</u> bread rolls and an <u>overheated</u> room. [hotels] → **negative**
- A sunny day but had expensive <u>mouldy</u> bread rolls and an overheated room. [restaurants] → **negative**
- A <u>sunny</u> day but had expensive mouldy bread rolls and an overheated room. [beaches] → **positive**

The key characteristics of reviews vary substantially between websites and so methods and findings from TripAdvisor or any other site should not be assumed to apply to all other review sites. For example, a comparison of TripAdvisor (439,000 reviews), Expedia and Yelp reviews in English of hotels found substantial differences in length (Expedia reviews were much shorter), themes (Yelp focused more on basic service) and sentiment (TripAdvisor and Expedia were dominated by positive reviews) (Xiang et al. 2017).

6.5 Applications of Sentiment Analysis to Tourism

TripAdvisor and other customer review sites seem to be taken seriously by many hotels and restaurants, who encourage customers to make positive comments and perhaps also respond individually to negative feedback. Positive reviews are recognised as important for future customers, many of whom book online with sites that feature customer reviews. Tourism managers may also analyse comments in non-review general social web sites, such as Twitter and Facebook.

6.5.1 Customer Relations Management Applications

Most national destination marketing organisations for the top ten destination countries mentioned social media analysis as a potential investigating tool in surveys conducted in 2010 (Hays et al. 2013) and today it seems likely to be an accepted tool. Providers of specialist services, such as development tourism (Jurowski 1998) may need to pay attention to online perceptions of their offering since negative sentiments could make it unviable. Of course, tourism organisations, attractions and hotels may run their own social media marketing campaigns (e.g., Hays et al. 2013) and may use sentiment analysis to help evaluate the success of individual initiatives. There isn't a definitive list of the most common applications of sentiment analysis to tourism but the following are likely candidates for inclusion.

- **Brand/business/service monitoring over time**: For example, a hotel chain might monitor its brand image in Twitter and Facebook daily and to identify long term trends in popularity. An early basic lexical sentiment analysis of an unspecified number of tweets from seven months in 2010 containing the words 'Phuket' or 'Bangkok' found some evidence of decreasing positivity (Claster et al. 2010).
- **Competitive intelligence**: A business can benchmark itself or investigate its competitors by analysing their social media presence in parallel with their own. For instance, a set of destinations could be rated and compared for sentiment (Valdivia et al. 2017a). This competitive intelligence might also identify market opportunities created by failures or threats posed by others' successful innovations.
- **Macro trends**: Tourism researchers can mine customer feedback and comments to identify broader trends than evident from individual hotels, restaurants, and attractions. For example, big data collections of tourist attraction reviews or other

feedback can help to identify which features are important for visitors (Alcoba et al. 2017). These may reveal patterns that are not evident from local information, such as the types of experience valued by different genders or nationalities.

- **Managing reputational risk**: Reputational risk management can be aided through the early online identification of surges in customer negativity (e.g., Das and Das 2016). The Hong Kong tourism boycott following the "occupy central" social movement was spread on social media (Sina Weibo), and could have been identified by routine monitoring (Luo and Zhai 2017).
- **Answer questions**: Managers may have specific problems or questions that can be investigated with social media data. An analysis of about 20,000 reviews about Paris in 14 categories (e.g., restaurants, shopping, off the beaten path) from virtualtourist.com. It used sentiment analysis to discover the main reasons why tourists were critical of the city transportation system (Kim et al. 2017).

6.5.2 Computer Systems to Support Tourists

Tourists can benefit from sentiment analysis through computing systems that provide recommendations. Software can do this by extracting the key aspects of potential destinations (e.g., hotels), extracting reviews about these destinations, processing the reviews with aspect-based sentiment analysis and then summarising the overall sentiment expressed about the customers potential choices. A web user wishing to decide between three hotels might be presented with a graphic showing their key aspects (e.g., breakfast quality, cleanliness, view) and then be shown the average rating of customers about those aspects (Schmunk et al. 2013). In theory, this would save customers the task of reading reviews if, for example, they were only concerned with one aspect of a hotel (e.g., children's entertainment, bar, beach) and so overall ratings that considered all aspects of a hotel would not be relevant. Whilst this would seem to be a useful service, there does not seem to be a successful commercial example so far. A promising but so far also apparently unsuccessful application is to overlay a map with icons representing tourism services, colour coded by sentiment (Cresci et al. 2014). Other applications have also attempted to exploit sentiment in tourist reviews to generate improved recommendations for them (e.g., Dong and Smyth 2016).

An unusual research application of sentiment analysis analysed the sentiment of tourism-related press coverage for countries and cities, generating an overall map of destinations by sentiment (Scharl et al. 2008). This study used a range of tools to gather the data and linguistic processing to identify relevant content and classify it for sentiment. The sentiment map may help customers to select destinations and plan itineraries.

An algorithm has been written to extract key phases from large sets of TripAdvisor hotel reviews to automatically summarise consumer feedback about them. Sentences were judged to be more relevant if they had been written by authoritative reviewers, were in reviews voted as helpful, were recent or contained a word or phrase indicating a thoughtful review (e.g., “all in all”, “nevertheless”). Similar sentences were rejected so that the final set of key phrases represented differing perspectives (Hu et al. 2017).

Computer software can also recommend activities to users based on their interests. Using reviews of 1036 activities in the USA, one system attempted to cluster these activities by type and match them to user interests. Sentiment was harnessed to help ensure that poor attractions were not recommended (Mittal and Sinha 2017). This idea points to the benefits of using big data to help personalise the experience of individual users, although care must be taken to avoid confusing travellers with strange suggestions, however.

6.5.3 Research Insights from Review Sentiment

Some research studies have focused on the characteristics of reviews on the basis that they are important for businesses and that information about how they work may help businesses to interpret or react to them. A comparison of TripAdvisor review sentiment and ratings found them to be consistent for both budget and high-end hotels (Geetha et al. 2017). It also pointed to aspects of hotels that most affect reviewer sentiments. A statistical analysis of factors influencing the helpfulness of English language TripAdvisor reviews found that sentiment expressed in a review tended to make it more helpful to users, as judged by the number of helpful votes per month received by the review in the site (Hu et al. 2017). This study analysed all types of entity in three regions in the USA, with a total of over 700,000 reviews.

From a different perspective, an investigation of TripAdvisor reviews found that the overall rating of a review was frequently not a good guide to the sentiments expressed in the review, at least as judged by automatic sentiment analysis (Valdivia et al. 2017b).

An aspect-based sentiment analysis of over half a million TripAdvisor reviews of US Hilton hotels detected sentiment related to key aspects of a hotel stay (value, location, sleep, rooms, cleanliness, service, check-in, business service). The study also characterised reviewers by type and extracted review dates. From this volume of data, it was possible to compare average opinions about the different aspects of the hotels over time, between hotels, and between customer types, giving a rich dataset for exploration. The analysis also showed that the different types of customer gave different average ratings and therefore it could be misleading to compare average scores between hotels that attract different types of customer. Business users tended to be the least positive (Chang et al. in press).

Based on half a million TripAdvisor hotel reviews from New York City, visitors found negative sentiment to be more helpful in a review than positive sentiment (Lee et al. 2017). Potential customers may be more willing to discount positive reviews as potentially fake, but find negative reviews helpful to check whether there is an aspect of a hotel that would be unacceptable for them. Similar results were found in reviews of 10 Las Vegas hotels (Chang 2015).

Related to the above, an investigation of 676,000 tourism-related tweets about Milan from early 2011 found negative tweets to be more influential in the sense of being more likely to be retweeted (Barbagallo et al. 2012). This aligns with previous studies that have found negativity to be more powerful than positivity in provoking discussions (Chmiel et al. 2011) and so companies should be concerned to react to negative online comments.

Social media analysis can at least partially replace the traditional questionnaire strategy since it can exploit feedback already posted to social media. Whilst survey research has revealed that cleanliness is probably the single most important factor in hotel satisfaction and can analyse this issue in more detail (Lockyer 2003; Zemke et al. 2015), social media analysis has the potential to give even more fine-grained information through access to much larger samples of feedback.

An investigation into whether guided tours improve customer satisfaction at Spanish ports visited by Mediterranean cruises employed both questionnaires and TripAdvisor reviews, with sentiment extracted by the lexical algorithm Rapidminer 6.3. The sentiment analysis component showed that a guided tour helped to make each port visit a more explicitly positive experience, at least in terms of TripAdvisor comments. Alternatively, however, it is possible that more positive tourists are more likely to choose a guided tour (Sanz-Blas and Buzova 2016).

Many innovative applications of sentiment analysis have been proposed that are relevant to specific types of attraction. One application exploited geographic location information on Twitter to investigate the areas of Disneyland, California that attracted strong emotions, finding three areas with many positive tweets (Park et al. 2018).

6.6 Data Sources

There are many sources of tourism customer feedback. Structured websites like TripAdvisor are a good source of information because they contain explicit reviews and are categorised by attraction. This makes it easy to identify and interpret relevant data. General social media sites like Twitter, Weibo and Facebook can also be useful because they are probably used by a wider segment of the population. These sites are more difficult to exploit because heuristics are needed to identify relevant posts and the posts will not have explicit ratings. Businesses may also have their own data, such as responses to online feedback forms from their website and customer emails.

Twitter is an obvious general social web site for big data sentiment analysis because it is public and widely used to share information and updates. Although in the early days of Twitter, its data was distributed free to researchers via Spinn3r, this offering was withdrawn in 2010. Researchers can now either buy tweets from resellers or use the real-time free data collection option, the Twitter API. This allows tweets to be collected with a set of keyword queries but the data is restricted to the previous week. Large scale (millions) and/or long-term Twitter data can still be collected free for research projects that are planned well in advance. This is possible with free software like Mozdeh (mozdeh.wlv.ac.uk), COSMOS (socialdatalab.net/COSMOS) or Chorus (chorusanalytics.co.uk) to monitor collections of queries over a long period (typically months). Twitter's terms of service prohibit data sharing between researchers but analysis results can be published.

An investigation of sentiment related to Iizuka City in Japan is of interest for its method to identify relevant tweets. It used a manually curated list of 200 queries describing key attractions (e.g., "Kaho performing theatre" in Japanese) and an automatic query expansion method so that the resulting large set of Twitter queries could identify a substantial percentage of all relevant tourism-related tweets (Shimada et al. 2011).

Social media posts from Facebook and other widely used general purpose websites are probably more valuable than Twitter for businesses because they focus on sharing information and experiences between friends, giving a natural environment for posting about holidays. This data is rarely used by researchers because of the cost to access it, however.

6.7 Summary

This chapter introduced sentiment analysis, a mainstream commercial technique for analysing customer feedback and comments as part of a social media analysis strategy. Although sentiment analysis is imperfect, it is accurate enough to deliver useful information, such as identifying sentiment patterns and trends, when applied to big data. Standard sentiment analysis tasks include detecting the polarity of a text, the sentiment strength or fine-grained emotions expressed in it. Some software can also detect the aspect of a text that relates to sentiment expressed in it. A range of sentiment analysis software is available for applications but users should be aware of the two different types (lexical and machine learning) and the factors that can affect its accuracy, such as whether it is tailored for the tourism context.

There are many research-based and customer relations management applications of sentiment analysis. These exploit the sentiment expressed by consumers in their posts about an attraction, restaurant or hotel to mine insights that can inform decision makers and customer relations personnel. These vary from small scale issues like identifying the aspects of an individual hotel that trouble guests most to large scale identifying trends in customer reactions to a country or holiday destination region. Sentiment analysis can also be used to research theoretical issues related to customer satisfaction (e.g., whether guided tours improve the experience of visitors).

All users of sentiment analysis for big social media data should be aware of the self-selection bias limitations inherent in the medium. This is critical for researchers but also important for commercial users. A logical way to exploit social media data is to use it to identify issues or to get initial evidence to test hypotheses or explore data, with traditional methods, such as surveys, used for more robust checking of the most important points.

In terms of future research, a key attraction of sentiment analysis is that the free rich datasets currently extractable from websites like TripAdvisor allow big data approaches to be applied relatively cheaply to run larger scale studies than previously possible. This may give deeper and more general insights into tourism than previously possible. If recent attempts to extract sentiment from visual non-textual data, such as facial expressions (Soleymani et al. 2017), can be harnessed then future tourism research may also have access to new sources of implicit customer feedback that will support even more fine-grained analyses.

References

Alcoba J, Mostajo S, Paras R, Ebron RA (2017) Beyond quality of service: exploring what tourists really value. International conference on exploring services science. Springer, Berlin, Germany, pp 261–271
Baccianella S, Esuli A, Sebastiani F (2010) SentiWordNet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. In: LREC2010, pp 2200–2204
Balahur A, Turchi M (2014) Comparative experiments using supervised learning and machine translation for multilingual sentiment analysis. Comput Speech Lang 28(1):56–75
Barbagallo D, Bruni L, Francalanci C, Giacomazzi P (2012) An empirical study on the relationship between Twitter sentiment and influence in the tourism domain. In: Information and communication technologies in tourism 2012. Springer, Vienna, Austria, pp 506–516
Cambria E, Poria S, Gelbukh A, Thelwall M (2017) Sentiment analysis is a big suitcase. IEEE Intell Syst 32(6):74–80
Chalfen RM (1979) Photograph's role in tourism: some unexplored relationships. Ann Tourism Res 6(4):435–447
Chang WL (2015) Discovering the voice from travelers: a sentiment analysis for online reviews. Workshop on E-Business. Springer, Berlin, Germany, pp 15–26
Chang YC, Ku CH, Chen CH (in press) Social media analytics: extracting and visualizing Hilton hotel ratings and reviews from TripAdvisor. Int J Inf Manage
Chen JS (2003) Market segmentation by tourists' sentiments. Ann Tourism Res 30(1):178–193
Cherif W, Madani A, Kissi M (2016) A combination of low-level light stemming and support vector machines for the classification of Arabic opinions. 11th international conference on Intelligent Systems: Theories and Applications (SITA). IEEE Press, Los Alamitos, pp 1–5
Chmiel A, Sobkowicz P, Sienkiewicz J, Paltoglou G, Buckley K, Thelwall M, Hołyst JA (2011) Negative emotions boost user activity at BBC forum. Physica A: Stat Mech Appl 390(16):2936–2944
Claster WB, Cooper M, Sallis P (2010) Thailand—tourism and conflict: modeling sentiment from Twitter tweets using naïve Bayes and unsupervised artificial neural nets. In: Second international conference on Computational Intelligence, Modelling and Simulation (CIMSiM). IEEE Press, Los Alamitos, A, pp 89–94
Cresci S, D'Errico A, Gazzé D, Lo Duca A, Marchetti A, Tesconi M (2014) Tourpedia: a web application for sentiment visualization in tourism domain. In: Proceedings of the 9th Language Resources and Evaluation Conference (LREC 2014), pp 18–21
Das S, Das A (2016) Fusion with sentiment scores for market research. 19th international conference on Information Fusion (FUSION). IEEE Press, Los Alamitos, CA, pp 1003–1010
Dong R, Smyth B (2016) Personalized opinion-based recommendation. International conference on case-based reasoning. Springer, Berlin, Germany, pp 93–107
Dragouni M, Filis G, Gavriilidis K, Santamaria D (2016) Sentiment, mood and outbound tourism demand. Ann Tourism Res 60:80–96
Farhadloo M, Patterson RA, Rolland E (2016) Modeling customer satisfaction from unstructured data using a Bayesian approach. Decis Support Syst 90(1):1–11

Fondevila-Gascón JF, Mir-Bernal P, Puiggròs-Román E, Muñoz-González M, Berbel-Giménez G, Gutiérrez-Aragón Ó, Crespo JL (2016) Semantic fields to improve business: the hotels case. 5(2):47–59
Gamon M, Aue A, Corston-Oliver S, Ringger E (2005) Pulse: mining customer opinions from free text. Lect Notes Comput Sci 3646:121–132
Gan Q, Ferns BH, Yu Y, Jin L (2017) A text mining and multidimensional sentiment analysis of online restaurant reviews. J Qual Assur Hospitality & Tourism 18(4):465–492
García-Pablos A, Cuadros M, Linaza MT (2016) Automatic analysis of textual hotel reviews. Inf Technol Tourism 16(1):45–69
Geetha M, Singha P, Sinha S (2017) Relationship between customer sentiment and online customer ratings for hotels–an empirical analysis. Tour Manage 61(1):43–54
Glorot X, Bordes A, Bengio Y (2011) Domain adaptation for large-scale sentiment classification: a deep learning approach. In: Proceedings of the 28th International Conference on Machine Learning (ICML-11). ICML Press, Bellevue, WA, pp 513–520
Gräbner D, Zanker M, Fliedl G, Fuchs M (2012) Classification of customer reviews based on sentiment analysis. Springer, Helsingborg, Sweden, pp 460–470
Hays S, Page SJ, Buhalis D (2013) Social media as a destination marketing tool: its use by national tourism organisations. Curr Issues Tourism 16(3):211–239
Hofer-Shall Z (2010) The forrester wave: listening platforms, Q3 2010. Forrester Res. http://www.demainlaveille.fr/wp-content/uploads/2010/07/forrester.pdf
Hu YH, Chen YL, Chou HL (2017a) Opinion mining from online hotel reviews–a text summarization approach. Inf Process Manage 53(2):436–449
Hu YH, Chen K, Lee PJ (2017b) The effect of user-controllable filters on the prediction of online hotel reviews. Inf Manage 54(6):728–744
Jo Y, Oh AH (2011) Aspect and sentiment unification model for online review analysis. In: Proceedings of the fourth ACM international conference on web search and data mining. ACM Press, New York, NY, pp 815–824
Jurowski C (1998) A study of community sentiments in relation to attitudes toward tourism development. Tourism Anal 3(1):17–24
Kim K, Park OJ, Yun S, Yun H (2017) What makes tourists feel negatively about tourism destinations? application of hybrid text mining methodology to smart destination management. Technol Forecast Soc Chang 123:362–369
Kirilenko AP, Stepchenkova SO (2017) Sochi 2014 olympics on Twitter: perspectives of hosts and guests. Tour Manag 63:54–65
Kirilenko AP, Stepchenkova SO, Kim H, Li X (in press) Automated sentiment analysis in tourism: comparison of approaches. J Travel Res
Lee M, Jeong M, Lee J (2017) Roles of negative emotions in customers' perceived helpfulness of hotel reviews on a user-generated review website: a text mining approach. Int J Contemp Hospitality Manag 29(2):762–783
Li X, Li J, Wu Y (2015) A global optimization approach to multi-polarity sentiment analysis. PloS ONE 10(4):e0124672
Liu B (2012) Sentiment analysis and opinion mining. Morgan and Claypool, San Rafael, CA
Lockyer T (2003) Hotel cleanliness—how do guests view it? let us get specific. A New Zealand study. Int J Hospitality Manag 22(3):297–305
López Barbosa RR, Sánchez-Alonso S, Sicilia-Urban MA (2015) Evaluating hotels rating prediction based on sentiment analysis services. Aslib J Inf Manage 67(4):392–407
Luo Q, Zhai X (2017) "I will never go to Hong Kong again!" How the secondary crisis communication of "Occupy Central" on Weibo shifted to a tourism boycott. Tour Manag 62:159–172
Lyu SO (2016) Travel selfies on social media as objectified self-presentation. Tourism Manag 54(1):185–195
Marrese-Taylor E, Velásquez JD, Bravo-Marquez F, Matsuo Y (2013a) Identifying customer preferences about tourism products using an aspect-based opinion mining approach. Procedia Comput Sci 22:182–191

Marrese-Taylor E, Velásquez JD, Bravo-Marquez F (2013b) Opinion zoom: a modular tool to explore tourism opinions on the web. In: Proceedings of the 2013 IEEE/WIC/ACM international joint conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT), vol 3. IEEE Computer Society, pp 261–264

Micu A, Micu AE, Geru M, Lixandroiu RC (2017) Analyzing user sentiment in social media: implications for online marketing strategy. Psychol Mark 34(12):1094–1100

Mittal R, Sinha V (2017) A personalized time-bound activity recommendation system. IEEE 7th annual Computing and Communication Workshop and Conference (CCWC). IEEE Press, Menlo Park, CA, pp 1–7

Neviarouskaya A, Prendinger H, Ishizuka M (2009) Compositionality principle in recognition of fine-grained emotions from text. In: Proceedings of the Third International ICWSM Conference (ICWSM2009). IEEE Press, Menlo Park, CA, pp 278–281

Oquab M, Bottou L, Laptev I, Sivic J (2014) Learning and transferring mid-level image representations using convolutional neural networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. IEEE Press, Los Alamitos, CA, pp 1717–1724

Pang B, Lee L (2008) Opinion mining and sentiment analysis. Found Trends Inf Retrieval 2:1–135

Park SB, Kim HJ, Ok CM (2018) Linking emotion and place on Twitter at Disneyland. J Travel Tourism Mark 35(5):664–677

Philander K, Zhong Y (2016) Twitter sentiment analysis: capturing sentiment from integrated resort tweets. Int J Hospitality Manage 55(1):16–24

Reis G, Blair-Goldensohn S, McDonald RT (2014) U.S. Patent No. 8,799,773. U.S. Patent and Trademark Office, Washington, DC

Salas-Zárate MP, Paredes-Valverde MA, Rodríguez-García MÁ, Valencia-García R, Alor-Hernández G (2017) Sentiment analysis based on psychological and linguistic features for Spanish language. Current trends on knowledge-based systems. Springer, Berlin, Germany, pp 73–92

Sanz-Blas S, Buzova D (2016) Guided tour influence on cruise tourist experience in a port of call: an eWOM and questionnaire-based approach. Int J Tourism Res 18(6):558–566

Scharl A, Dickinger A, Weichselbraun A (2008) Analyzing news media coverage to acquire and structure tourism knowledge. Inf Technol Tourism 10(1):3–17

Schmunk S, Höpken W, Fuchs M, Lexhagen M (2013) Sentiment analysis: extracting decision-relevant knowledge from UGC. In: Information and communication technologies in tourism 2014. Springer, Cham, pp 253–265

Schuckert M, Liu X, Law R (2015) Hospitality and tourism online reviews: recent trends and future directions. J Travel Tourism Mark 32(5):608–621

Schweidel DA, Moe WW (2014) Listening in on social media: a joint model of sentiment and venue format choice. J Mark Res 51(4):387–402

Sharma S, Chakraverty S, Jauhari A (2017) Leveraging polarity switches and domain ontologies for sentiment analysis in text. In: International conference on computational intelligence, communications, and business analytics. Springer, Singapore, pp 84–92

Shimada K, Inoue S, Maeda H, Endo T (2011) Analyzing tourism information on Twitter for a local city. First ACIS International Symposium on Software and Network Engineering (SSNE2011). IEEE Press, Los Alamitos, CA, pp 61–66

Soleymani M, Garcia D, Jou B, Schuller B, Chang SF, Pantic M (2017) A survey of multimodal sentiment analysis. Image Vision Comput 65(1):3–14

Taboada M, Brooke J, Tofiloski M, Voll K, Stede M (2011) Lexicon-based methods for sentiment analysis. Comput Linguist 37(2):267–307

Teso E, Olmedilla M, Martínez-Torres MR, Toral SL (2018) Application of text mining techniques to the analysis of discourse in eWOM communications from a gender perspective. Technol Forecast Soc Chang 129(1):131–142

Thelwall M (2018a) Gender bias in sentiment analysis. Online Inf Rev 42(1):45–57

Thelwall M (2018b) Gender bias in machine learning for sentiment analysis. Online Inf Rev 42(3):343–354

Thelwall M, Buckley K, Paltoglou G, Cai D, Kappas A (2010) Sentiment strength detection in short informal text. J Am Soc Inf Sci Technol 61(12):2544–2558
Thelwall M, Buckley K, Paltoglou G (2012) Sentiment strength detection for the social Web. J Am Soc Inf Sci Technol 63(1):163–173
Thelwall M, Goriunova O, Vis F, Faulkner S, Burns A, Aulich J, Mas-Bleda A, Stuart E, D'Orazio F (2016) Chatting through pictures? A classification of images tweeted in one week in the UK and USA. J Assoc Inf Sci Technol 67(11):2575–2586
Tian X, He W, Tao R, Akula V (2016) Mining online hotel reviews: a case study from hotels in China. In: AMCIS 2016: surfing the IT innovation wave—22nd Americas conference on information systems, pp 1–8
Valdivia A, Luzón MV, Herrera F (2017a) Sentiment analysis in TripAdvisor. IEEE Intell Syst 32(4):72–77
Valdivia A, Luzón MV, Herrera F (2017b) Sentiment analysis on TripAdvisor: are there inconsistencies in user reviews? International conference on hybrid artificial intelligence systems. Springer, Berlin, Germany, pp 15–25
Volkova S, Wilson T, Yarowsky D (2013) Exploring demographic language variations to improve multilingual sentiment analysis in social media. In: EMNLP2013, pp 1815–1827
Xiang Z, Du Q, Ma Y, Fan W (2017) A comparative analysis of major online review platforms: implications for social media analytics in hospitality and tourism. Tour Manag 58(1):51–65
Yan Q, Zhou S, Wu S (2018) The influences of tourists' emotions on the selection of electronic word of mouth platforms. Tour Manag 66:348–363
Ye Q, Zhang Z, Law R (2009) Sentiment classification of online reviews to travel destinations by supervised machine learning approaches. Expert Syst Appl 36(3):6527–6535
You Q, Luo J, Jin H, Yang J (2015) Robust image sentiment analysis using progressively trained and domain transferred deep networks. In: Proceedings of the twenty-ninth AAAI conference on artificial intelligence. AAAI Press, Menlo Park, pp 381–388
Zemke DMV, Neal J, Shoemaker S, Kirsch K (2015) Hotel cleanliness: will guests pay for enhanced disinfection? Int J Contemp Hospitality Manage 27(4):690–710