



Sentiment Analysis in Tourism

Maria Cristina ENACHE*

ARTICLE INFO

Article history:
Accepted March 2020
Available online May 2020
JEL Classification
K22, M21

Keywords:
Sentiment analysis, Social media,
Machine learning

ABSTRACT

One of the most important parts in gathering information is what people think. Users express their opinions and opinions on tourist products and services, since the emergence of social networks. Thus opinions are subjective information that represent feelings, or subjective opinions about products and or services. Every day millions of people travel around the globe for business, vacations, sightseeing or other reasons. An astronomical amount of money is spent on tickets, accommodation, food, transportation and entertainment. Tourism is an information-based business where there are two types of information flows: an information flow is from suppliers to consumers or tourists. This is information about the goods that tourists consume, such as tickets, hotel rooms, entertainment and more. The other information flow consists of aggregated information about tourists to service providers. In this article we will address the information that comes from tourists. When aggregated data about tourists are presented correctly, analyzed by a correct algorithm and used on time, they lead to obtaining quality information for tourism service providers. The analysis of sentiments generated by tourism services on social networks can be a very useful tool for analyzing tourism related data.

© 2020 EAI. All rights reserved.

1. Introduction

On the fateful evening of April 9, 2017, United Airlines forced a passenger from an over-booked flight. The nightmare incident was filmed by other passengers on their smartphones and was posted immediately. Such a video, posted on Facebook, has been distributed more than 87,000 times and watched 6.8 million times until 6:00 pm on Monday, just 24 hours later. The response from the company increased the fiasco of the action, as Monday afternoon, they sent a statement from the CEO excusing himself that "he had to rearrange customers." This is exactly the kind of PR disaster that we would all like to miss. This is also a great example of why we care not only if people talk about a brand, but also how they talk about it. More online presence does not mean an increase in image capital.

Nowadays, business are of all shapes and sizes and have significant interactions with customers, especially on social networks. Most people involved in tourism promotion activities are overwhelmed by the huge volume of data found in the online environment. Nowadays, using sentiment analysis on social media, we can get incredible information about the quality of the conversation going on around a tourist location, hotel or tour package.

Nowadays, people share a lot of content on social media in the form of pictures - be they personal, or everyday scenes or their opinions in the form of cartoons or memories, photos or videos posted. On social media, people use natural language to express their joy, happiness, anger or frustration. The analysis of sentiments would be valuable for delivering information about the reputation of a travel agency, locations, tourist service etc. but it could deliver information about existing competitors. Many social networks offer a wealth of information that companies can use to analyze how people talk about services instead. Based on this analysis, companies can create personalized messages in the hope of generating specific reactions online and offline.

Travel companies can deliver personalized promotional messages based on business objectives. Also, I can strategically plan what to say on each channel and, more importantly, what not to post to keep the audience receptive. In addition to the fact that a more comprehensive picture of the target audience is being made, it could thus measure its reputation, or track customer service performance. When measuring people's reactions by interacting with the support team, a clearer picture of their level of satisfaction is obtained. This allows post-service customer issues to be addressed more effectively, based on a different type of feedback - unfiltered, less intrusive and more honest.

* Dunarea de Jos University of Galati, Romania. Email address: mpodoleanu@ugal.ro (M.C. Enache)

2. Concepts

Sentiment analysis is the contextual extraction of a text that identifies and extracts subjective information from a source material and helps an agency / hotel / board to understand the social sentiment of their products or services while monitoring online conversations. However, the analysis of social media flows is usually limited only to the analysis of basic sentiments and values based on number. This is similar to just scratching the surface and missing those high-value prospects waiting to be discovered. [1]

With sentiment analysis, one can develop the growth strategy in two parallel directions at the same time. On the one hand, this type of analysis makes it possible to quickly identify any negative comments on social media, so that it can be answered quickly and rectify the situation before it gets worse. On the other hand, information can be used to target customer segments and to provide them with a higher level of personalized service or experience.

Sentiment analysis is an imperative area of research of natural language processing, which involves extracting and identifying a person's attitude towards a particular subject. It is also known as extracting opinion. A "sentiment" is the expression of one's feelings. It can be an attitude, an opinion or an emotion. A framework for analyzing generic sentiments is shown in Fig. 1. Data can be extracted from various online resources such as e-mail, blogs and social media, such as Facebook, Twitter, etc. These data are processed in advance to improve the quality of the analysis. Multiple word extraction algorithms can then be applied to preprocessed algorithms to extract useful information. In the last stage, the analysis of sentiments is performed and the results are presented in the form of charts and graphs, which facilitate the understanding of the analyzed sentiments. [2]

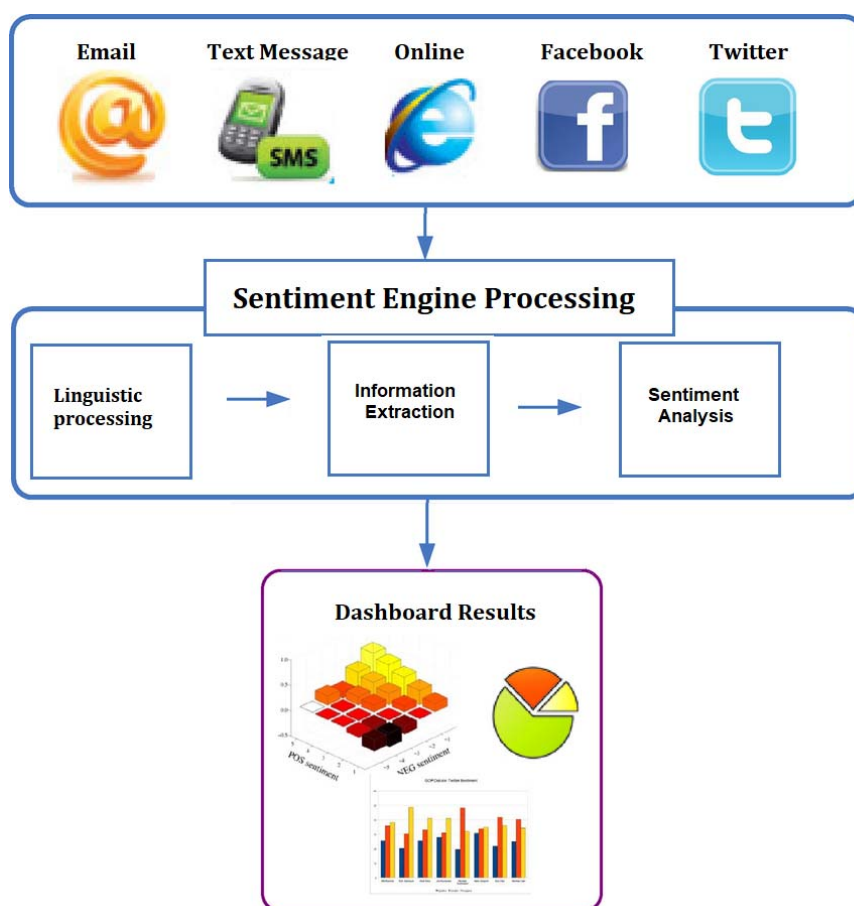


Figure 1. A generic framework for sentiment analysis

3. Analysis

How can sentiment analysis be used, in tourism

- Analyzing posts over a period of time to see the sentiment of a particular audience.
- Performing sentiment analysis on all mentions on the social networks for the agency, tour package
- Automating the analysis processes
- Using analytics to get a deep perspective on what is happening on social media channels.

The analysis of sentiments for tourism would be useful in tourism activities, because such functionalities would:

- Increases understanding of how a travel agency or tour operator evolves over time

- It would be a tool for researching the competition and would increase the understanding of how its reputation evolves over time.
- It would help to identify potential PR crises to take immediate action.

Whether you are exploring a new market, anticipating future trends or keeping an edge in competition, sentiment analysis can make all the difference.

This can be achieved by introducing functionalities for:

- To analyze the reviews of the tourist packages of an agency, pensions, hotel and perform comparative analyzes with those of the competition
- Generating weekly, monthly or daily reports - a kind of early warning system
- Comparison of the reputation of an agency or locations on international markets
- Generating formal market reports or business journals for long-term trends
- Post analysis on social networks for real-time events
- Post analysis in the form of reviews for non-filtered customer feedback
- Analysis of sentiments in social scenes or contexts to gain a rich perspective on details.

A component of the analysis of sentiments with functionalities in tourism would provide the opportunity to monitor the opinions regarding the services and their reputation (the so-called feedback), from social media in real time and to act accordingly, through functionalities of analysis of the sentiments.

A classification of the analyzed texts (sentences / phrases) could take into account the subjectivity criterion, having two distinct categories:

- objectives, which contain concrete information,
- topics which contain explicit opinions, beliefs and opinions regarding the specific entities.

4. Applications

The most relevant aspects appear through the analysis of the subjective category. For the analysis of texts based on the criterion of subjectivity, the specialized studies highlight the following types of analyzes:

- Posting sentiment analysis - is the simplest form of sentiment analysis and posting is supposed to contain an opinion on a single main object expressed by the author of the message. In the specialized literature, we found two approaches for analyzing sentiments at the post level: supervised and unsupervised. The supervised approach assumes that there is a finite set of classes. The posting must be classified, the training data being assigned to each class. The simplest case is when there are classes: positive, negative and neutral. Given the training data, the system learns a classification model, using a classification algorithm, such as SVM (Support Vector Machines) or KNN (K-nearest neighbors). This classification is then used to label new posts in sentiment classes.
- Sentence analysis - a single post can contain multiple views even on the same entity defined as a service or agency or location. When we want to have a clearer picture of the opinions expressed about an entity, we must move to the sentence / sentence level. Most methods use supervised methods to classify sentences into two positive or negative classes.

Understanding that people like certain services or tourist locations for different reasons, evaluating a client is not so challenging. Rather, interesting are the aspects that make tourist experiences attractive or not and how these aspects can be improved.

This is especially important for hotels or restaurants reviews, which might be interested in them. For example, you can plan a teambuilding in nice location. In this case, the service and atmosphere will be quite important for the experience. First, we can divide each review into sentences, then use a tool, such as spaCY or Gensim, to get the distinct topics that reviewers mentioned in each sentence (ie, quality of accommodation, service, hours waiting, atmosphere and variety of the menu). After getting the topics, we have to figure out if a reviewer had a positive or negative attitude about the experience itself. We can simulate a modeling of reviewer sentiment through StanfordCoreNLP. [8]

Stanford's Core NLP program was trained in movie reviews in which a reviewer could discuss both positive and negative aspects of the movie in the same sentence (for example, "the plot was slow, but the acting was great").

According to the site, the model actually builds a representation of whole sentences based on the sentence structure. Calculates the sentiment based on how words make up the meaning of longer sentences. In this way, the model is not so easy to use in other contexts.

But there's a Python wrapper that can be used to make calls to the NLP Server. To make the calls, you will need to pip install pycorenlp, and import StanfordCoreNLP from pycorenlp.

The text to be analyzed is `text_amb = "everything was like garbage to me but we keep coming back because my colleagues loves the place"`

```
sentiment_analyzer_scores(text_amb)
get_sentiment(text_pos)
```

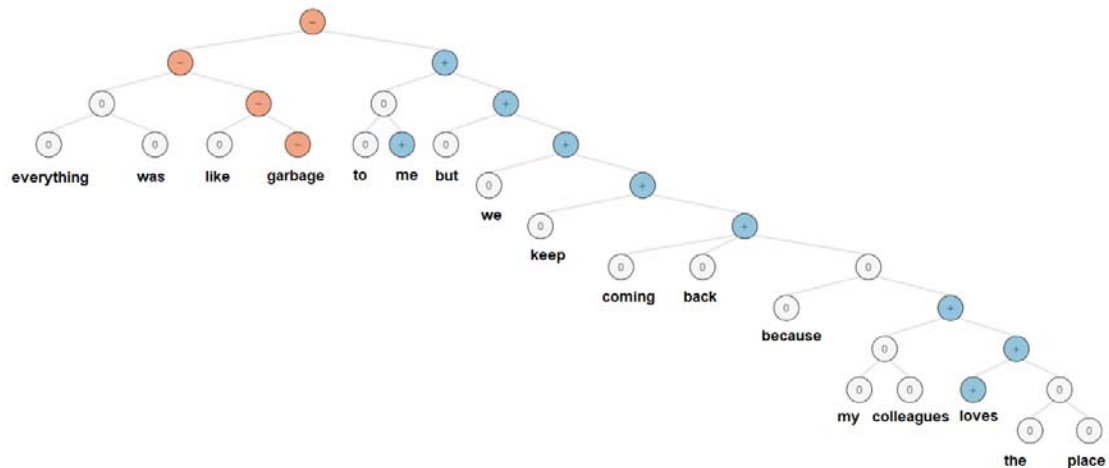


Figure 2. Sentiment Analysis Simulation with
<http://nlp.stanford.edu:8080/sentiment/rntnDemo.html>

Sentiment: Negative

Sentiment score: 1

Sentiment distribution (0-v.negative, 5-v.positive:[0.058, 0.39, 0.35, 0.16, 0.02]

The model classifies the general sentence as quite negative (0 is the most negative, 4 is the most positive). The distribution of sentiments shows that there are some neutral and even positive aspects of this sentence, but in general, this is not a good evaluation.

- Aspects based on sentiment analysis - Previous approaches work effectively when either the entire post or each piece of text analyzed refers to a single entity. However, in many cases, people talk about entities that have more features (attributes) and, of course, opinions differ. The analysis based on the emotional load is the research problem that focuses on the recognition of all forms of expression- sentiment in a post and the aspects to which it refers. The classic approach, increasingly used by the public relations department of tourist agencies, consists in identifying the emotional nature of the posts regarding the quality of services or the image of the agency.
- Comparative sentiment analysis - In many cases, people do not give a direct opinion about a tourist service or package, instead preferring comparable opinions, such as: "Bran is more beautiful than Moeciu". In this case, the purpose of such functionality for sentiment analysis is to identify sentences that contain comparative opinions, as well as to extract the preferred entity from them. Using a relatively small number of comparative words: more, less, light or more, at least, best or favor, big, prefers, than, superior, inferior, number one, against, can cover 98% of the total opinions comparative. For these words / phrases being used very often, but with low accuracy, a classifier can be used to filter sentences that do not contain comparative opinions. An algorithm could be used to identify preferred entities based on the type of comparisons used and the presence of the negation.

Most of the work in sentiment analysis in recent years has focused on developing more accurate sentiment classifiers, addressing some of the main challenges and limitations in the field, such as:

- Subjectivity and tone - Detecting subjective and objective texts is just as important as analyzing their tone. In fact, the so-called objective texts do not contain explicit sentiments. Say, for example, you intend to analyze the sentiment of the following two texts: The package is beautiful or The package is green. Most people would say that the sentiment is positive for the first and neutral for the second. But adjectives, verbs and some nouns should not be treated the same way as they create the sentiment. In the examples above, beauty is more subjective than green.
- Context and tone - All words are spoken at a certain time, in a certain place, by and for some people, in context. The analysis of sentiment without context becomes quite difficult. If the statements "Everything", "Absolutely nothing!" there would be answers to the question "What did you like about this product?", we would say that the first answer would be positive and the second negative. But if the answers come to the question "What didn't you like about this service?" The negative in the question will cause the sentiment analysis to change everything. A large amount of pre-processing or post-processing will be required if we consider at least part of the context in which the texts were produced. However, how to preprocess or post-process data to capture context bits will help to analyze sentiment.
- Irony and sarcasm - Differences between the literal and intended meaning (ie, irony) and the more insulting or ridiculous version of irony (ie sarcasm) usually change the positive sentiment

into the negative, while the negative or neutral sentiment can be changed into the positive. However, detecting irony or sarcasm requires a better analysis of the context in which the texts are produced and, therefore, are really difficult to detect automatically.

- Comparisons - There are some comparisons, which do not need contextual clues to be correctly classified. However, most are a little more difficult to classify. Context can make a difference.
- Emoticons - There are two types of emoji: Western ones (for example: D) are encoded in a single character or in a combination of characters, while the Eastern emoji (eg ㄟ _ (ツ) _ / ㄟ) are a longer character combination with a vertical character. However, no matter how much attention is given to each emoji, a lot of preprocessing might be needed. For example, they could be preprocessed as social media content and by transforming Western and Eastern emoticons into chips and interpreting them as features for classification purposes, to help improve the performance of sentiment analysis.

5. Conclusions

The age of getting relevant information from social media data has now come with the advancement of technology. It's time for any tourism company to take advantage of the power of data, but to get the most in-depth information, you must use the power of AI, deep learning and smart classifiers, such as semantic contextual search and sentiment analysis.

Acknowledgement

This work was supported by a grant of the Romanian Ministry of Research and Innovation, CCCDI – UEFISCDI, project number PN-III-P1-1.2-PCCDI-2017-0800 / 86PCCDI/2018 - FutureWeb, within PNCDI III.

References

1. Balahur, *Sentiment analysis in social media texts*, in: *4th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, 2013, pp. 120–128.
2. Cui, M. Zhang, Y. Liu, S. Ma, *Emotion tokens: bridging the gap among multilingual twitter sentiment analysis*, in: *Proceedings of the 7th Asia Conference on Information Retrieval Technology, AIRS*, 2011, pp. 238–249
3. <https://towardsdatascience.com/sentiment-analysis-concept-analysis-and-applications-6c94d6f58c17>.
4. <https://thatware.co/sentiment-analysis/>
5. https://www.researchgate.net/publication/321812828_Sentiment_Analysis_in_Tourism_Capitalizing_on_Big_Data
6. https://link.springer.com/chapter/10.1007/978-981-13-6339-9_6
7. http://www.ijiset.com/v1s9/IJISSET_V1_I9_69.pdf
8. <http://nlp.stanford.edu>