DATAQUEST

Amsterdam, 2017

Analyzing hotel reviews
# using Latent Aspect Rating Analysis

Analyzing hotel reviews using Latent Aspect Rating Analysis

Thousands of hotel reviews are available (TripAdvisor, Booking.com, Expedia), typically containing an overall (numerical) rating and a written review. Public ratings are an important factor in the demand of rooms in the hotel industry. Therefore, getting a good picture of the motivation behind a rating can be of great value.

RiskQuest developed a customized web scraper to aggregate review data from Tripadvisor.com:

•   Hotels in the Netherlands, Germany and the US were considered

•   2250 hotels were scraped

•   Over 1 million reviews were gathered

The written hotel reviews are analyzed using sentiment analysis to reveal the opinion on hidden aspects such as the hotel's service and location. Aspect ratings are often not (directly) visible, because the reviewers are only asked to give an overall rating. However aspect ratings can provide a lot of information. Using latent aspect rating analysis (LARA), the hidden aspects in the written review are extracted together with the relative weights the reviewer has placed on each aspect. [1].
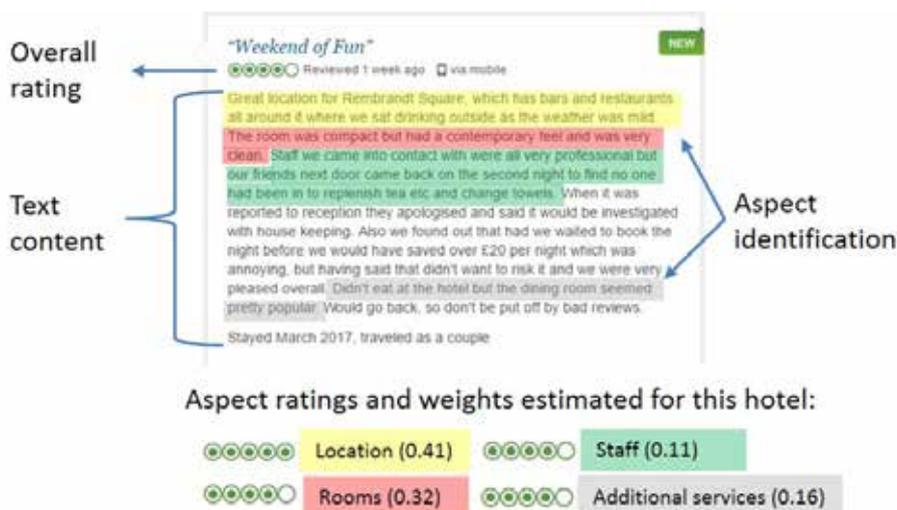

The model provides a flexible framework for exploring any aspect of customer satisfaction. Any aspect/topic can be analyzed without supervision; after defining a new aspect the model is self-learning. Because of this, we can answer questions about any kind of aspect.

---

1   http://times.cs.uiuc.edu/~wang296/paper/p618.pdf

Does it work? Location ratings of hotels in Amsterdam

For many hotel visitors location is an important aspect. RiskQuest extracted the location scores of all hotels in Amsterdam using the LARA model. The picture below shows the percentile scores (1.00 has highest location rating, 0.00 worst). We see that in most cases the hotels near each other have similar scores. The exceptions can be explained by sample reading of the reviews. The nice thing about location scores is that we can also compare these with the reputation of the locations in Amsterdam.
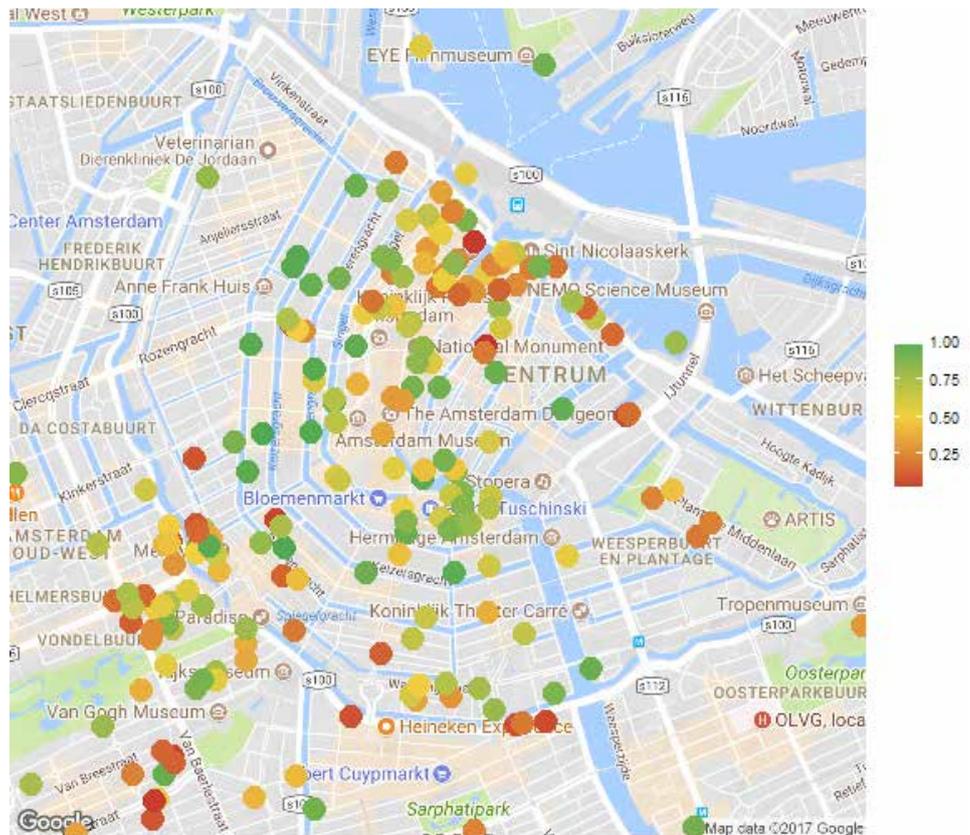
For example, we can see that the hotels around the canals score relatively high on location, which is no surprise, since the canals are one of the main reasons people visit Amsterdam. Furthermore, the hotels near the Central Station score relatively low, which makes sense, since a lot of people experience this location as much too crowded.

There are more examples that confirm the correctness of the location scores. We can for example look at the public transport facilities of the different locations. In conclusion we can say that these results seem consistent with common sense and sample readings of reviews.

Possible applications

As mentioned before, any topic can be analyzed. We can come up with any aspect that interests the hotel, like the bar/restaurant and noise disturbance in the area, and extract the ratings of these aspect. Using these aspect ratings it can, for example:

- Provide clarity about the cause of low/high ratings.
    - ° Identify aspects with low/high ratings.
    - ° Can be used to formulate clear objectives to improve ratings and thereby revenue.
- Help hotel visitors find an appropriate hotel given their preferences:
    - ° If they attach great importance to a certain aspect, advice a hotel with a high rating on that particular aspect. This can be implemented in a hotel search engine.

Moreover, it can be used for all kind of reviews, like restaurant reviews and product reviews on websites like Amazon and Bol.com. This makes it a powerful and widely applicable framework.

Key features

Flexible:

- Can be used for any type of review containing numerical rating and text content.
- Any kind of topic/aspect can be investigated.

Self-learning:

- Analysis can be done without supervision, just the defining of aspects is needed.

Plausible:

- Results seem consistent with common sense and sample readings of reviews.

Started as a subsidiary of RiskQuest, a leading Dutch consultancy firm in the financial sector, DataQuest is an Amsterdam based consultancy firm with a broad expertise in modelling any type of data, big or small.

In a data driven society, DataQuest understands how data can be used to gain insight and add value to companies in finance, retail and industry. At DataQuest we can help you with the exploratory phase of determining where value can be added, but we are equally at home with a clearly formulated problem.

You are more than welcome for a cup of coffee at our office to see how DataQuest can help your organization.

DATAQUEST

Herengracht 495, Amsterdam

+31 20 693 29 48

info@dataquest.com