fang_2021_criteria_determination_of_analytic_hierarchy_process_using_a_topic_model

Year

2021

Author(s)

Jin Fang and Fariborz Y. Partovi

Title

Criteria determination of analytic hierarchy process using a topic model

Venue

Expert Systems with Applications

Topic labeling

Partially automated

Focus

Secondary

Type of contribution

Established approach

Underlying technique

Manual assignment of pre-existing aspect labels

Topic labeling parameters

\

Label generation

To deliver the contents of topics more precisely, human judgments and intervention are required to assign a topic label/title, and is done based on the semantic similarity of significant words in the topic matrix

The TripAdvisor.com has also designed several aspects for both hotels and restaurants. For hotels, it provides the aspects 'Location', 'Cleanliness', 'Service', 'Value', which allow guests to evaluate their experience.
It also displays the list of 'Property Amenities' and 'Room Features' for each hotel.

For restaurants, it provides the aspects of 'Food', 'Service', 'Value', 'Atmosphere', which allow guests to evaluate their experience.
These factors are predetermined labels proposed by website experts, which identify the main issues of concern by guests.

Additionally, Parasuraman, Zeithaml, and Berry (1988) described the development of a 22-item instrument (SERVQUAL) for assessing customer perceptions of service quality in service and retailing organizations. This paper suggested the following labels for the five dimensions of service quality, and we also considered these five dimensions to determine the sub-criteria.

- Tangibles: Physical facilities, equipment, and appearance of personnel.
- Reliability: Ability to perform the promised service dependably and accurately.
- Responsiveness: Willingness to help customers and provide prompt service.
- Assurance: Knowledge and courtesy of employees and their ability to inspire trust and confidence.
- Empathy: Caring, individualized attention the firm provides its customers.

Considering the criteria/sub-criteria mentioned above and the se- mantic similarity of significant words in topics, we labeled topics and put them in the third column of Tables 4 and 5.

Table 4
Extracted topics for DS_1 using LDA.

ID	Topics (β_k)	Label	Proportions (θ_d)
1	Stay hotel room property good service philadelphia price marriott location	Price	0.0546
2	Staff front_desk make great service helpful friendly time check experience	Responsiveness	0.0816
3	Bed bathroom large comfortable small nice shower suite room area	Room Features	0.0547
4	Bathroom dirty bed clean floor carpet shower smell bad stain	Cleanliness	0.0671
5	Staff great clean good location nice room friendly comfortable restaurant	Empathy	0.1800
6	Breakfast food good order coffee service restaurant eat day bar	Restaurant & Bar	0.0282
7	Nice free good lobby coffee breakfast pool clean water small	Breakfast & Pool	0.0373
8	Stay great staff clean location philly time friendly night family	Reliability	0.0765
9	Airport shuttle night flight clean check stay good free minute	Airport Shuttle	0.0368
10	Street market location reading_terminal convention_center walk great close restaurant block	Close to Attractions	0.0455
11	Parking car park valet lot night check pay street nice	Valet Parking	0.0408
12	Breakfast stay home house philadelphia inn comfortable penn lovely wonderful	Assurance	0.0336
13	Great location restaurant walk city distance philadelphia view staff within_walk	Close to Business	0.0768
14	Great service love staff beautiful stay lobby view bar restaurant	Tangibles	0.0608
15	Check call front_desk arrive day book time night reservation wait	Front desk	0.0809
16	Night elevator noise sleep door room hear loud floor people	Quiet	0.0448

Table 5
Extracted topics for DS_2 using LDA.

ID	Topics (β_k)	Label	Proportions (θ_d)
1	City_tavern history philadelphia tavern great visit experience lunch time historic	History	0.0804
2	Review enjoy glad hope experience hear time visit feedback service	Assurance	0.0959
3	Hotel good stay walk philadelphia street lunch philly find decide	Location	0.0498
4	Pizza good order chicken great delicious dish sushi fresh noodle	Food Variety	0.1093
5	Good service price order bad average give time quality bit	Price	0.0245
6	Table reservation wait seat dinner arrive order time bar sit	Responsiveness	0.0627
7	Service great friendly staff good excellent atmosphere dinner delicious attentive	Empathy	0.0428
8	Good delicious salad order appetizer excellent dinner great chicken dish	Food Quality	0.0603
9	Sandwich good philly cheesesteak order meat cheese_steak great steak cheese	Cheesesteak Sandwich	0.1507
10	Visit philly philadelphia year good time dinner friend wife eat	Visit	0.0831
11	Menu good dish great service excellent delicious choice option item	Menu Options	0.0726
12	Breakfast good brunch great coffee delicious service lunch egg order	Brunch	0.0402
13	Great good beer bar drink nice service selection atmosphere friendly	Bar	0.0677
14	table great bar atmosphere decor nice area room view beautiful	Decoration	0.0600

Motivation

"To better summarize the topics, we had to condense and label them instead of showing all ten words"

Topic modeling

LDA

Topic modeling parameters

Hyperparameters optimised every ten iterations
α: 5.0
η: 0.01
(α controls the division of documents into topics and η controls division of topics into words)
Training iterations: 5000
Nr of topics (K): 2 to 60
Random seed: 4, 5, 6

Nr. of topics

16 (DS1) and 14 (DS2)

Label

Pre-existing aspect label manually assigned to the topic

Label selection

\

Label quality evaluation

\

Assessors

\

Domain

Paper:
Dataset: Hotels & Restaurants (Reviews)

Problem statement

The purpose of this paper is to develop a technology-based model for identifying various criteria in a decision-making situation.
We used topic modeling to discover critical criteria and their corresponding weights in the Analytic Hierarchy Process (AHP).
Approximately 100,000 hotel reviews and 100,000 restaurant reviews were scraped from TripAdvisor.com for criteria determination.
Next, an AHP model with criteria and 12 hotels / restaurants as alternatives were compared and ranked. The results compared favorably with more than 1000 re- views of these hotels/ restaurants in TripAdvisor.com, thus validating the methodology.

Corpus

Origin: TripAdvisor
Nr. of documents: 230,599 (101,744 + 128,855)
Details:

- hotel reviews (DS1) and restaurant reviews (DS2)
- customer reviews of various hotels and restaurants in Philadelphia

Document

Hotel review including the hotel name, reviewer name, submission date, and review description
Restaurant review including the restaurant name, reviewer name, submission date, and review description

Pre-processing

- Remove repeated records
- Tokenize (into a list of words with special characters and punctuations removed) and remove special characters
- create an n-gram model (union of unigrams and bigrams. Each word is a unigram, and every two adjacent words create a bigram)
- remove stop words
- make n-grams
- lemmatization
- create a term-document matrix

```
@article{fang_2021_criteria_determination_of_analytic_hierarchy_process_using_a_topic_model,
  abstract = {The purpose of this paper is to develop a technology-based model for identifying various criteria in a decision-making situation. We used topic modeling to discover critical criteria and their corresponding weights in the Analytic Hierarchy Process (AHP). Approximately 100,000 hotel reviews and 100,000 restaurant reviews were scraped from TripAdvisor.com for criteria determination. Next, an AHP model with criteria and 12 hotels/restaurants as alternatives were compared and ranked. The results compared favorably with more than 1000 reviews of these hotels/restaurants in TripAdvisor.com, thus validating the methodology.},
  author = {Jin Fang and Fariborz Y. Partovi},
  date-added = {2023-03-21 16:58:08 +0100},
  date-modified = {2023-03-21 16:58:08 +0100},
  doi = {https://doi.org/10.1016/j.eswa.2020.114306},
  issn = {0957-4174},
  journal = {Expert Systems with Applications},
  keywords = {Analytic hierarchy process, Topic model, Latent Dirichlet allocation, Hotel selection, Restaurant selection, Group decision},
  pages = {114306},
  title = {Criteria determination of analytic hierarchy process using a topic model},
  url = {https://www.sciencedirect.com/science/article/pii/S0957417420310046},
  volume = {169},
  year = {2021}}
```

#Thesis/Papers/Initial