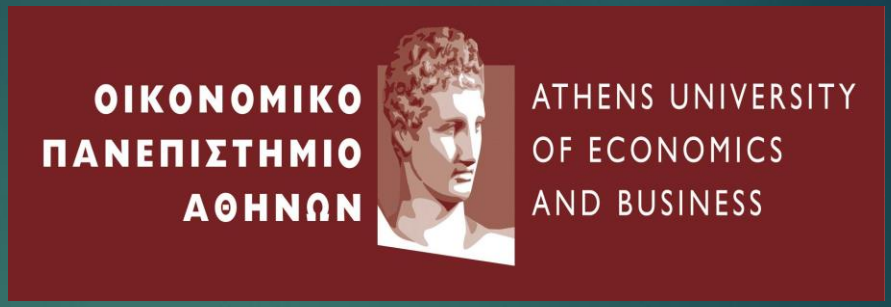

Utilising Fuzzy Sets Qualitative Comparative Analysis to Identify the necessary Conditions that a User Review expresses a Highly Positive Sentiment

Dimitris K. Kardaras
Assistant Professor in information Management
Business Informatics Lab
Dept. of Business Administration
School of Business
Athens University of Economics and Business
Email: dkkardaras@yahoo.co.uk
Kardaras@aueb.gr

Structure of presentation…

- Aims of our research
- Background
  - Importance of tourism industry
  - Introduction to Recommender Systems
  - Related work
- Methodology-Models
- Results

Aims of our research

- To develop an approach to user profiling drawing on the user reviews sentiment and the TF of selected features by **utilising the Fuzzy set Qualitative Comparative Analysis** (FsQCA).

- To identify the **causal combinations,** i.e. what are the necessary and sufficient combination of services features demographic characteristics that to lead to highly positive sentiment.

Background …

The tourism industry

- The World Travel and Tourism Council (WTTC) indicates that the contribution of travel and tourism to world GDP grew for the sixth consecutive year in 2015…
- rising to a total of 9.8% of world GDP (US$7.2 trillion).
- According to WTTC, the tourism sector employs 284 million people, which globally represents 1 in 11 jobs.

Recommender Systems…I

- A particular class of intelligent information assistants known as *recommender systems* has been used in the travel domain, to provide travellers with relevant recommendations regarding their trips,
- E.g. which places to visit, which hotel to choose, …etc.

Recommender Systems…II

- Some of the travel recommenders draw their knowledge from sources that describe travel and tourist locations, travel modes and other related aspects (called the ***content based recommendation*** approach).

Recommender Systems…III

- Other travel recommender systems employ the experience of fellow travelers in order to provide relevant recommendations (an approach known as ***collaborative filtering***).

Recommender Systems in tourism

- Current literature shows that recommendation is a common service in the tourism subdomains of travel and travel services such as accommodation (i.e., hotel recommendations).

Examples of Recommender Systems in tourism

CT-Planner, City Planner, etc

- Offers tour plans that can be refined gradually as the users express their preferences and characteristics (e.g., willingness to walk, walking speed, etc.).

Methodology steps I

1. **Select documents** published by user (using web crawler tool).
2. Perform **Sentiment Analysis**
3. **Fuzzify sentiment** (e.g. use Triangular Fuzzy Sets)
4. **Identify the features** that will constitute the causal combinations and specify the term that will represent the outcome set.
5. **Calculate** the Term Frequency (TF) for each identified term, (Using KNIME text mining tool).

Methodology steps II

4. **Calculate the weight of each term** using the following formula:

$$W_{tk} = TF_{tk} * log \frac{N_i}{d_{tk}}$$

Where W_{tk} represents the weight of term (tk).

TF is the term frequency;

N_i is the total number of documents published by user (i) and d_{tk} represents the number of documents that contain term tk.

Calibrate term important by review's publication time: W_{tk}***Review_Time**

Review Time: defined as a fuzzy set (e.g., recent, earlier, old)

Methodology steps III

5. **Apply the FsQCA and produce User Interests causal combinations**.

a. Produce the truth table of all possible permutations of the terms considered. Each permutation is a possible causal combination.

b. Calculate membership degrees for each combination. Its calculation is performed drawing on the fuzzy sets operations theory (Union, Intersection, Complement).

Methodology steps IV

6. **Calculate the consistency and the coverage** of the solutions
7. **Identify best combinations**, by selecting the combinations that exhibit a consistently rate above a threshold (in this paper is set at 0.8).

Data Analysis-Input Sets

- This paper analyses reviews collected reviews from a sample of (5) hotel customers.

- For simplicity reasons, (5) **Input Sets**, i.e. terms representing hotel services are selected from the total set of terms identified in the reviews, i.e.

Quietness	Sea View	Staff Friendliness	Cultural Activities	Restaurant

Data Analysis-Output Set

- The **Outcome Set** is the *High Positive Sentiment (HPS)* by each user during his/her hotel stay.

Data Analysis-Output set TFN

- The **Sentiment Outcome Set** is modelled as a TFN with the following linguistic scales, that indicate levels of Sentiment.

Sentiment Linguistic scale	Triangular fuzzy scale			Mean of fuzzy numbers
Very Positive	(0.75,	1.00,	1.00)	1.00
Positive	(0.50,	0.75,	1.00)	0.75
Neutral	(0.25,	0.50,	0.75)	0.50
Negative	(0.00,	0.25,	0.50)	0.25
Very Negative	(0.00,	0.00,	0.25)	0.00

Sentiment Analysis Example (MS Azure Machine Learning)

Nice. Brilliant location opposite the cathedral. Bed and linen ideal for a good nights sleep. Good combination of design in neo-classical building. Quiet. The roof terrace is currently very trendy for an early evening drink. Great view. The "8 hours before sunrise" cocktail is, incidentally, fun and delicious. Breakfast has a good choice and is good quality. Staff professional and friendly. We will definitely want to revisit.	positive	0.996195793151855
The upscale hotel Daios has much to offer its guests. It is located on the waterfront not far from the White Tower and nearby museums. The rooms are adequate in size and the bathrooms are well designed with both tub and separate shower. The staff provides excellent service. When we reported a drainage problem on our way to breakfast we found it completely fixed when we returned to our room. One thing it lacks are convenient USB and electrical outlets for charging cell phones and tablets.	negative	0.328726410865784
Nice hotel with friendly staff and free parking near the see shore. Central location. Hotel is more like a boutique hotel of 2 floors and the rooms are big. Even if you expect the hotel to be noisy, once you close the windows you can forget about the street and you can feel the holidays. Still no pool or sauna or nothing similar. I would say that the hotel is 5 stars in location, room facilities and air conditioning (actually the best one I meet in this area), breakfast, Greek coffee and hospitality but there is nothing to do on site except the coffee bar. So overall all if you compare with 5 stars in big seaside towns, I would say 4 stars. But if you compare with other 5 starts in old towns and city centers, the. It's really ok.	positive	0.986069917678833

Data Analysis-Input Dataset

		Terms' Weights (Wtk) for each Customer				
Outcome Set Highly Positive Sentiment (Y)	Customer	Quietness	Sea View	Staff Friendliness	Cultural Activities	Restaurant
0.50	1	0.30	0.50	0.40	0.70	0.70
0.70	2	0.30	0.70	0.60	0.70	0.90
0.1	3	0.10	0.30	0.20	0.60	0.50
0.7	4	0.50	0.70	0.40	0.50	0.70
0.9	5	0.30	0.70	0.60	0.70	0.70

Data Analysis-Truth Table (FsQCA)

Developing the truth table is developed.

Since there are 5 terms to consider the number of permutations is

$$2^5 = 32$$

The cells in the truth table take the value (1) or (0) representing true or false.

Data Analysis-The Truth Table

Causal Combination	Quietness	Sea View	Staff Friendliness	Cultural Activities	Restaurant
1	0	0	0	0	0
2	0	0	0	0	1
3	0	0	0	1	0
4	0	0	0	1	1
5	0	0	1	0	0
6	0	0	1	0	1
7	0	0	1	1	0
8	0	0	1	1	1
9	0	1	0	0	0
10	0	1	0	0	1
11	0	1	0	1	0
12	0	1	0	1	1
13	0	1	1	0	0
14	0	1	1	0	1
15	0	1	1	1	0
16	0	1	1	1	1
17	1	0	0	0	0

Data Analysis- Calculate Membership degrees

Calculate membership degrees for each causal combination, using fuzzy operations (union, intersection, complement)

Data Analysis- Calculate Membership degrees II

For example, combination number 3 for customer-1:

μ=(*Quietness=false ∩Sea View=false ∩ Staff Friendliness=false ∩ Cultural Activities=true ∩ Restaurant=false)* = (not *(Quietness), not (Sea View), not (Staff Friendliness), Cultural Activities, not (Restaurant))*.

The μ(Quietness=false) = ((1- (Quietness)) = (1-0.3)=0.7

Data Analysis

- The Membership degrees for combinations for each customer

Causal Combination	Customer 1	Customer 2	Customer 3	Customer 4	Customer 5
1	0.3	0.1	0.4	0.3	0.3
2	0.3	0.3	0.4	0.3	0.3
3	0.3	0.1	0.5	0.3	0.3
4	0.5	0.3	0.5	0.3	0.3
5	0.3	0.1	0.2	0.3	0.3
6	0.3	0.3	0.2	0.3	0.3
7	0.3	0.1	0.2	0.3	0.3
8	0.4	0.3	0.2	0.3	0.3
9	0.3	0.1	0.3	0.3	0.3
10	0.3	0.3	0.3	0.5	0.3
11	0.3	0.1	0.3	0.3	0.3
12	0.5	0.4	0.3	0.5	0.4
13	0.3	0.1	0.2	0.3	0.3
14	0.3	0.3	0.2	0.4	0.3
15	0.3	0.1	0.2	0.3	0.3
16	0.4	0.6	0.2	0.4	0.6
17	0.3	0.1	0.1	0.3	0.3

Data Analysis

- Calculate Consistency and Coverage for each Causal combination.

Causal Combination	Consistency	Coverage
1	0.785714286	0.379310345
2	0.8125	0.448275862
3	0.733333333	0.379310345
4	0.789473684	0.517241379
5	0.916666667	0.379310345
6	0.928571429	0.448275862
7	0.916666667	0.379310345
8	0.933333333	0.482758621
9	0.846153846	0.379310345
10	0.882352941	0.517241379
11	0.846153846	0.379310345
12	0.904761905	0.655172414
13	0.916666667	0.379310345
14	0.933333333	0.482758621
15	0.916666667	0.379310345
16	0.954545455	0.724137931
17	1	0.379310345

Data Analysis- Necessary and Sufficient Causal Combinations

- Assuming a threshold value of 0.8 for the consistency firstly,
- and then looking for the higher possible coverage,

- **the analysis results into** two causal combinations; the combinations number 12 and 16.

Data Analysis

Causal Combination	Quietness	Sea View	Staff Friendliness	Cultural Activities	Restaurant
16	0	1	1	1	1
12	0	1	0	1	1

Data Analysis Causal Combinations Final set

Customers who express Highly Positive Sentiment, show interest in

- (Sea View) AND (Staff friendliness) AND (Cultural activities) AND (Restaurant) OR
- (Sea View) AND (Cultural activities) AND (Restaurant).

Conclusions

- This study suggests that FsQCA can be applied in the context of web personalisation and recommender systems.
- Future research can focus on examining the applicability of the FsQCA.
- Interesting challenges regarding the method itself, such as its applicability with large datasets, multiple output sets.

Thank You!!!