

Combining Text Summarization and Aspect-based Sentiment Analysis of Users' Reviews to Justify Recommendations

Cataldo Musto, Gaetano Rossiello, Marco de Gemmis, Pasquale Lops, Giovanni Semeraro



Università degli Studi di Bari – Italy

[\[name.surname\]@uniba.it](mailto:[name.surname]@uniba.it)

[cataldomusto](#) - [@SWAP_research](#)



Background

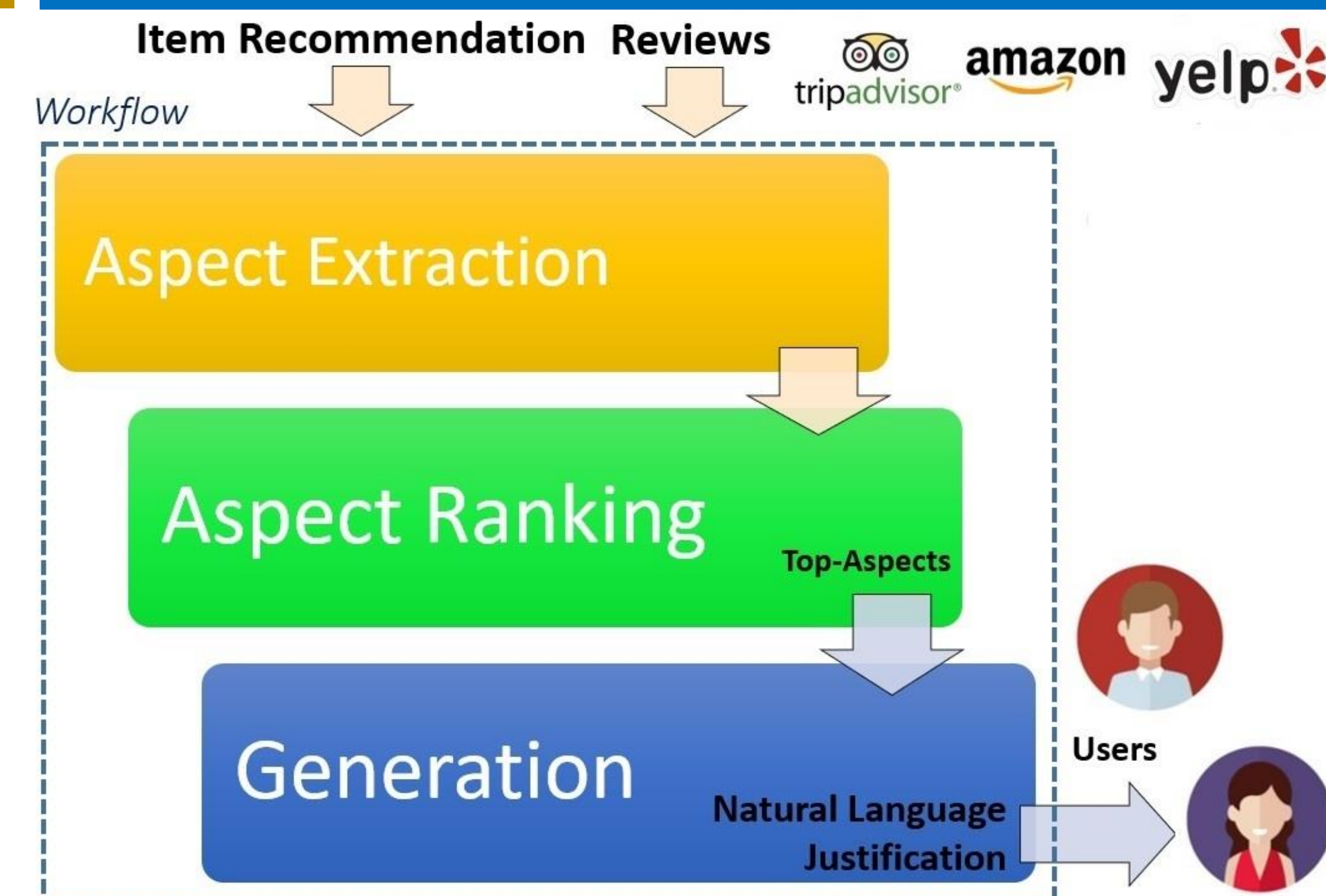
Richness of Users' Reviews
+
Effectiveness of Text Summarization Techniques



Research Question

Is it possible to exploit automatic text summarization techniques to generate a natural language justification based on users' reviews?
Is it effective?

Framework



Methodology

ASPECT EXTRACTION– Exploits Kullback-Leibler divergence. Nouns having a KL-divergence higher than a threshold ϵ are labeled as 'aspects'.

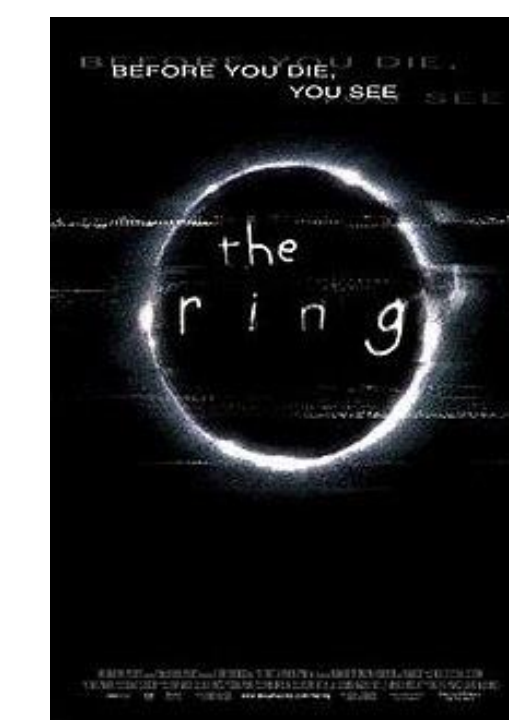
$$\delta_t(c_a || c_b) = p(t, c_a) \log \frac{p(t, c_a)}{p(t, c_b)}$$

ASPECT RANKING– Combines KL-divergence and average sentiment of the aspect. Aspects that are frequently discussed with a positive sentiment are identified and exploited.

$$score(a_j) = \frac{\sum_{i=1}^N n_{a_j, r_i} * rel(a_j, r_i) * sent(a_j, r_i)}{N}$$

GENERATION– Based on two steps. **Sentence Filtering** and **Text Summarization**. **Sentence Filtering** identifies compliant sentences (i.e., expressing a positive sentiment, containing a main aspect, having an impersonal style). **Text Summarization** uses compliant sentences to generate a *summary of the reviews*. *The algorithm builds a centroid vector of all the reviews and identifies sentences that are similar to the centroid and maximize coverage and diversity, as well.*

Use Case



MOVIE: The Ring

MAIN ASPECTS

actor - thriller - effects
horror - character

OUTPUT: "If you like or love the blood and gore kinds of films, this movie will certainly disappoint you as the focus is on **character**, story, mood and unique **special effects**. The Ring is a story about supernatural evil therefore, it is a **horror** film, done very much in the style of the psychological **thriller**."

Experimental Design

Protocol: User Study in the MOVIE domain. 141 people. 300 movies. ~150k reviews.

Parameters: Justification Length (Short=50 words, Long=100) and #Aspects (10 and 30).

Experiments: (1) Between-subject experiment on varying of the parameters – (2) Within-subject experiment. Comparison to a baseline exploiting users' reviews that did not employ text summarization.

Metrics: Transparency, Persuasion, Engagement, Trust, Effectiveness. Acquired through questionnaire.

Technical Details: CoreNLP + Stanford Sentiment Analysis + Word2Vec for Text Summarization.

Experimental Results

Experiment 1

Configuration		Metrics				
Aspects	Length	TRA	PER	ENG	TRU	EFF
Top-10	Short	2.83	3.06	3.06	2.83	0.89
Top-30	Short	3.16	3.06	2.69	3.19	0.94
Top-10	Long	3.95	3.64	3.37	3.55	0.55
Top-30	Long	3.24	3.18	3.12	3.22	0.38

Experiment 2

MOVIES	ASPECTS+SUMMAR.	ASPECTS	INDIFFERENT
Transparency	54.55%	40.91%	4.55%
Persuasion	77.27%	13.64%	9.09%
Engagement	63.63%	27.27%	9.09%
Trust	68.18%	4.55%	27.27%

Findings

- Long justifications > short justifications, on average
- Top-10 > Top 30 aspects
- Top-10+Long best configuration

Findings

- Justifications based on text summarization **better than** review-based baseline