MM6761: Take-home Assignment 2

Dai Yao (dai@{yaod.ai, mle.bi})

March 2, 2024

1 Packard and Berger (JCR 2021) and Wang et al. (JCR 2021)

In Study 1 of Packard and Berger (2021), they show that concrete language helps improve customer satisfaction, without controlling for vocal patterns. Study 2 further shows that when vocal information is not present, concrete language still helps improve customer satisfaction.

On the contrary, Wang et al. (2021) demonstrates in various studies that vocal tones play an important role in persuading listeners to take actions.

Do you feel that the results in the two papers are consistent, or not? If so, why? If not, why?

Answer:

Regarding Packard and Berger (2021), it is tempting to conclude, based on their study 2 (showing that concrete language helps improve customer satisfaction **absent of** any vocal elements in the communication) and study 1 (showing that using more concrete language in customer service **calls** improves customer ratings), that vocal information does not contribute to customer satisfaction. But this is not true. With their study 2 and 1, we can only conclude that using concrete language helps with increasing customer satisfaction, and we cannot say anything about the effect of vocal information on customer satisfaction.

Hence, it is difficult to conclude whether the results in the papers are consistent, or not.

2 Jedidi et al. (JM 2021) and Ryoo et al. (JM 2021)

Both Jedidi et al. (2021) and Ryoo et al. (JM 2021) measure the correlation or relevance between two text corpus. Please specify the main models in both papers, and discuss the main differences (two to three aspects).

Answer:

Jedidi et al. (JM 2021)

The two main text corpus in Jedidi et al. (2021) are: academic papers (corpus 1) and practitioner-oriented articles.

To assess the practical relevance of academic papers (i.e., how corpus 1 are related to corpus 2), the authors further assemble a set of 1,154 marketing terms, and use surveys to calibrate the practical relevance of each term (denoted as r_w). They then run LDA on all the documents in corpus 1 to discover the underlying topic structures, and calculate two important metrics: to what extent a topic is practically-relevant (Topical relevance), and to what extent a topic is discussed in each year (Timely relevance).

- Topical relevance: the practice relevance of topic t is $M_t = \sum_{w \in M} p_{wt} \times r_w$, where $M \subset W$ is the set of the 1,154 terms that are present in corpus 1, and p_{wt} is w's contribution to t.

- Timely relevance: the estimated LDA is used to predict the topic distributions of all the practitioner articles in corpus 2. For a particular year y, suppose the total number of documents in corpus 2 is D_y, and let $\hat{q}_{dty}$ be the predicted probability that the practitioner article d in year y covers topic t. Then, the timely relevance of topic t is calculated as $C_{ty} = \frac{1}{D_y} \sum_{d=1}^{D_y} \hat{q}_{dty}$.

With the two metrics, the relevance to practice of article d in year y can be calculated as $\mathrm{R2M}_{dy} = 100 \sum_{t=1}^{T} q_{dt} C_{ty} M_t$, where q_{dt} is the contribution of topic t to article d.

Ryoo et al. (JM 2021)

The main text data in Ryoo et al. (2021) are the spoiler reviews and non-spoiler reviews from IMDB.

The authors use the correlated topic modelling (CTM) to discover the main topics in the entire corpus (i.e., spoilers and non-spoilers). Different from LDA, the underlying topic weights η are modeled as $\eta|(\mu, \sigma) \sim \mathrm{MultiNomial}(\mu, \sigma)$, rather than $\eta|\mu \sim \mathrm{Dir}(\mu)$.

They then use a logistic regression (equation 2 on page 77) to identify the contribution of each topic to a review being a spoiler. They identify 23 (out of 61) topics that have significant prediction power. For each of these 23 topics, they calculate the "marginal" contribution (c_{ij1}, equation 3 on page 78) the topic in a review, and normalize these marginal contribution using all the reviews for each movie i (α_{ij}, equation 4 on page 78). They then calculate the spoiler INTENSITY using the α_{ij} and the number of words from each topic w_{j1} (equation 5 on page 78).

Main differences

- The former (Jedidi et al. 2021) applies a LDA on only one corpus, while the latter (Ryoo et al. 2021) combines both corpus (they virtually have only one corpus), and applies a CTM.

- The former employs experts to assess the practical relevance of each term, while the latter counts on the institutional background which forces the reviewers to disclose if a review is a spoiler (or not).

- The former directly uses the labels generated by experts in the calculation of practical relevance of academic papers, while the latter conduct further analysis to identify the determinants of spoiler reviews.
- Many more if you will.