

A Crowdsourcing Annotated Tourism Review Corpus of Emotion

Mansour Almansour, Paul Rayson and Scott Piao
School of Computing and Communications, Lancaster University, UK
{m.s.almansour; p.rayson; s.piao}@lancaster.ac.uk

We present a tourism review corpus, named TORCE (TOurism Review Corpus of Emotion), which was collected from a major online tourism website and annotated with Plutchik’s eight emotion categories (Plutchik, 1980) via crowdsourcing. Our survey shows there are very few such corpora with emotion annotation, and our corpus provides a valuable language resource for studying emotion expressions in tourism context.

The raw corpus data was collected using web crawling via the TripAdvisor tourism website (<https://www.tripadvisor.co.uk>), with their approval. From the raw data, we first selected reviews (some contain multiple sentences) related to attractions in the UK, from which we further selected those reviews that contain at least one emotion word. We also removed excessively long and short reviews, obtaining 2,500 sentences for the crowdsourcing annotation.

In regard to the crowdsourcing annotation, we paid MTurk (<https://www.mturk.com>) workers to classify each review into one or two most likely emotion category/ies from Plutchik’s eight emotion classes: “anger”, “anticipation”, “disgust”, “fear”, “joy”, “sadness”, “surprise”, and “trust”. We assigned the task only to the workers with proven high reputation and expertise. For each review, five workers were asked to annotate it with up to two most likely emotion categories. Inevitably, the crowdsourced annotation contained disagreements, such as incompatible emotion annotations from different workers, which we have analysed in order to reflect on how distinctive the emotion classes are and what information we can potentially learn where people disagree on a classification. To produce a final agreed annotation, the crowdsourced annotation was pruned by applying majority voting approach, i.e. only retaining those emotion classes for each review chosen by the majority of annotators.

We also analysed the inter annotator agreement (IAA), which shows a promising annotation consistency, obtaining a Plutchik Emotion Agreement (PEA) (Desai et al., 2020) score of 0.86. For example, in our corpus, 23% reviews received full agreement on emotion category from 5 workers, while 54% received the agreement from 3 workers. In terms of distribution of emotion categories, “Joy” is the dominant class that covers about 31% of the reviews. In contrast, only 2% of the reviews fall under the “Fear” category. This finding implies that tourism reviews

tend to be optimistic and positive. Moreover, our experiment shows that emotion annotation of travel reviews involves subjective judgement and therefore it is a highly challenging task.

References

Shrey Desai, Cornelia Caragea, and Junyi Jessy Li (2020). Detecting perceived emotions in hurricane disasters. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 5290–5305, Online. Association for Computational Linguistics.

Robert Plutchik (1980). A general psychoevolutionary theory of emotion. In Robert Plutchik and Henry Kellerman (eds), *Theories of Emotion*, pages 3–33. Academic Press.