

Propositions

- Without standardized evaluation metrics and shared data sets, it is impossible to assess the current state of the field of aspect-based sentiment analysis. (Chapter 2)
- For aspect classification, using just the co-occurrences between the aspects and the words in the text, can already give a solid performance. (Chapter 4)
- A small number of informative features is enough to account for the vast majority of the performance of a given sentiment classifier. (Chapter 6)
- The discourse structure is a valuable source of information when determining which part of the text is about which aspect. (Chapter 7)
- Proper use of domain ontologies in aspect-based sentiment analysis algorithms alleviates the need for large labeled training data (Chapter 8)
- Any experimental work that is not reproducible because it uses data sets that are not available to other researchers or because it has incomplete method or experiment descriptions, should not be considered a contribution to science.
- Ambiguity, while often a source of serious problems in natural language processing, is a necessary component of any natural language.
- Natural language is a lossy form of dimensionality reduction, where thoughts in n-dimensional space are projected onto a single dimension by speaking or writing one word or utterance after the other.
- Children should be taught at least the basics of programming as it is becoming an essential skill in the labor market.
- Artificial intelligence is dangerous, not because ‘it’ might decide to wipe out humanity, but because we do not understand well enough what happens inside the proverbial black box and an unexpected combination of input values could potentially lead to unexpected output values with massive implications.
- Becoming a parent while doing your PhD has an unknown net effect on work efficiency, as it is a source of both focus and distraction, and of both order and chaos.