

Propositions

accompanying the dissertation

PROXIMITY OF TERMS, TEXTS AND SEMANTIC VECTORS IN INFORMATION RETRIEVAL

by

Jeroen Bastiaan Pieter VUURENS

1. Proposition 1. The decay of the estimated relevance of a document is inversely proportional to the word distance between appearing query terms in text. (part I)
2. Proposition 2. Important news is published by multiple newspapers within a short timespan. (part II).
3. Proposition 3. To assess whether news articles discuss the same subtopic, the normalized information gain between news articles is indicative. (part II).
4. Proposition 4. Semantic item representations allow to generalize to patterns instead of reasoning towards specific items. (part III)
5. Proposition 5. In a semantic space, the perceived similarity between objects can be estimated by their proximity and a user dependent weighting of the encoded semantics. (Part III)
6. Proposition 6. Data understanding is a useful step that precedes hypothesis forming in the empirical cycle for research in data science.
7. Proposition 7. Generalization of results to other domains requires to consider a variety of corpora and a comparison against strong baselines.
8. Proposition 8. Objects or shapes that are in relative close proximity, regardless of shapes and sizes, appear to form groups. (from the Gestalt principle of Proximity)
9. Proposition 9. The higher tendency that people have to form interpersonal relationships with people in close proximity applies online as well. (from Social Psychology)
10. Proposition 10. When buying a house, humans tend to favor the size of their garden instead of the benefit of proximity to work, and regret that decision afterwards. (from the Unconscious Thought Theory)

These propositions are regarded as opposable and defensible, and have been approved as such by the promotor prof. dr. ir. A.P. de Vries.