UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Measuring the Semantic Consistency of Ordinal Annotations via Text Embedding Spaces and Its Applications

Permalink

https://escholarship.org/uc/item/9hm9z41w

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 47(0)

Author

Ehara, Yo

Publication Date

2025

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Measuring the Semantic Consistency of Ordinal Annotations via Text Embedding Spaces and Its Applications

Yo Ehara
Tokyo Gakugei University, Koganei, Tokyo, Japan

Abstract

We propose a method for measuring the consistency of ordinal annotations based on a pre-trained embedding vector space. Intuitively, our method finds a direction in the embedding space along which data points align as closely as possible to their annotated ranks. The proposed approach guarantees a globally optimal solution that is free from approximation errors. Thus, it yields a unique consistency measure given a dataset with human-provided ordinal annotations and a pre-trained embedding model. This feature facilitates a wide range of applications, including not only ordinal prediction but also the unsupervised detection of annotation errors within datasets, as well as consistency assessment of stage-based scales (e.g., whether the transitions "beginner to intermediate" and "intermediate to advanced" form linear progressions in the embedding space) during dataset construction. We evaluate our method using real-world datasets with ordinal annotations to demonstrate its effectiveness.

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society* ©2025 the author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).