

From-scratch Name Disambiguation

Validation Data Format Descriptions :

sna_valid_author_raw.json :

The two-level dictionary, the key is the author's name, and value is the list of paperIDs with the same author name. The competitors need to cluster the paper under the same author name, and then return a group of papers clustering, so that the papers within a cluster belong to the same author, and the papers between different clusters do not belong to the same author.

sna_valid_example_evaluation_scratch.json :

The submission example file, organized as a two-level dictionary. The key is the author's name. The value is a two-dimensional list. The length of the first dimension represents the number of clusters(authors), and the second dimension represents the paperIDs belonging to the same cluster(author).

sna_valid_pub.json:

The file contains the specific paper information which belongs to sna_valid_author_raw.json (The same format with train_pub.json).

sna_valid_author_ground_truth.json :

The ground truth of clustering result about sna_valid_author_raw.json.