Weekly Report

WU Zihan

January 18, 2024

Achieving Optimal Speed

- Current implementation has achieved the fastest performance to date.
- Beat the previous fastest by 10% on sparse datasets and 5x on non-sparse datasets.

Universality in Datasets and Models

- Algorithm shows universality across different datasets and models.
- Datasets: sparse and non-sparse.
- Models: SVD and NMTF (non-negative matrix tri-factorization).
- Note: Our partitioning and ensembling method can be applied to any matrix-factorization-based co-clustering algorithm.

Enhanced Performance on Diverse Datasets

- Outperforms current fastest by 10% on sparse datasets, which is achieved by applying our partitioning and ensembling method to the current fastest algorithm. (Reuter's Datasets: 272.2s over 302.5s, sparsity: 0.85)
- Outperforms current fastest by 5x on non-sparse datasets, which is achieved by applying our partitioning and ensembling method on an SVD-based co-clustering algorithm. (Amazon Datasets: 66.7s over 333.2s, sparsity: 0.01)