

Dimensionality reduction as infrastructure

Ben Schmidt: Northeastern University

August 2017

1. (a) i. Dimensionality Reduction as Infrastructure
ii. Social-technical systems.
2. 1. Current Methods.
3. 2. Stable Random Projection.
 - (a) i. A. Log transform
4. 3. Uses of the shared space.
 - (a) Visualizing the Hathi Trust.
 - (b) Prestige Classification Benchmark
 - (c) Duplicate detection
 - (d) Corpus alignment
 - (e) Transferrable classifiers
 - i. Predicting Language
 - ii. Predicting Year
 - iii. Predicting Copyright status
5. 4. Learning Library Classifications
 - (a) LC Classification
 - i. Success rate
 - A. By language
 - B. By date
 - C. Uses of a classification.
6. Classifying dates