



CSAIL



| | EN-IT | | EN-DE | | EN-FI | | EN-ES | |
|-------------------------------------|--------------|------|--------------|------|--------------|------|--------------|------|
| | P@1 | Time | P@1 | Time | P@1 | Time | P@1 | Time |
| (Zhang et al., 2017a) [†] | 0 | 46.6 | 0 | 46.0 | 0.07 | 44.9 | 0.07 | 43.0 |
| (Conneau et al., 2018) [†] | 45.40 | 46.1 | 47.27 | 45.4 | 1.62 | 44.4 | 36.20 | 45.3 |
| (Artetxe et al., 2018) [†] | 48.53 | 8.9 | 48.47 | 7.3 | 33.50 | 12.9 | 37.60 | 9.1 |
| G-W | 44.4 | 35.2 | 37.83 | 36.7 | 6.8 | 15.6 | 12.5 | 18.4 |
| G-W + NORMALIZE | 49.21 | 36 | 46.5 | 33.2 | 18.3 | 42.1 | 37.60 | 38.2 |

RESULTS

RESULTS ON DATASET OF DIVERSE.

- ▶ The MUSE dataset is unrealistically simple [Artetxe et al., 2018]
 - ▶ Most pairs of languages are somewhat related
 - ▶ Embeddings trained on strongly comparable corpora

▶ We also test our method on the (much harder) dataset of Dinu et al. [2015]



GPU-minutes



CPU-minutes



RESULTS ON DATASET OF DINU ET AL.

- ▶ The MUSE dataset is unrealistically simple [Artetxe et al., 2018]
 - ▶ Most pairs of languages are somewhat related
 - ▶ Embeddings trained on strongly comparable corpora
- ▶ We also test our method on the (much harder) dataset of Dinu et al. [2015]

| | EN-IT | | EN-DE | | EN-FI | | EN-ES | | |
|---------------------------------------|--------------|------|--------------|------|--------------|------|--------------|------|-------------|
| | P@1 | Time | P@1 | Time | P@1 | Time | P@1 | Time | |
| (Zhang et al., 2017a) [†] | 0 | 46.6 | 0 | 46.0 | 0.07 | 44.9 | 0.07 | 43.0 | GPU-minutes |
| (Conneau et al., 2018) [†] | 45.40 | 46.1 | 47.27 | 45.4 | 1.62 | 44.4 | 36.20 | 45.3 | |
| → (Artetxe et al., 2018) [†] | 48.53 | 8.9 | 48.47 | 7.3 | 33.50 | 12.9 | 37.60 | 9.1 | |
| G-W | 44.4 | 35.2 | 37.83 | 36.7 | 6.8 | 15.6 | 12.5 | 18.4 | CPU-minutes |
| → G-W + NORMALIZE | 49.21 | 36 | 46.5 | 33.2 | 18.3 | 42.1 | 37.60 | 38.2 | |

BONUS: THE GROMOV-WASSERSTEIN LINGUISTIC DISTANCE