## Weighted Jaccard Similarity to Checkpoint 5000 for pythia-160m-data-seed2

