

Distribution of cosine similarities for different versions of the same dataset

