**Sinkhorn Distances**

**Definition–** method for comparing probability distributions. It is based on the Optimal Transport (OP), which measures how much “mass” has to be moved in order to transform a distribution into another. The classic OT distances are computationally expensive. However, the Sinkhorn distances innovate by using Sinkhorn’s matrix scaling algorithm which results in computing orders of magnitude much faster.

(could be a solution for merge/split detection instead of cosine similarity)