**A Kernel Two-Sample Test**

**Definition–** non parametric statistical test for determining whether two samples are drawn from the same probability distribution. The method is based on kernel embeddings of distributions and it defines Maximum Mean Discrepancy (MMD) as the largest distance between the mean of two distributions in a reproducing kernel Hilbert space (RKHS). If it is large, the two samples are from different distributions. This is a computationally efficient and consistent test.

(could detect if the topic is still the same or if it split/merged into another)