JDA，联合分布适配，假设两种分布同等重要，但在真实环境中，两种分布不一定同样重

$$D(D_s,D_t)\approx D(P(X_s),P(X_t))+D(P(y_s|X_s),P(y_t|X_t))$$

⇒ 为了考虑和衡量两种分布的重要程度，引入一个平衡因子 μ，来动态调整两个分布之间的距离

$$D(D_s,D_t)\approx(1-\mu)D(P(X_s),P(X_t))+\mu D(P(y_s|X_s),P(y_t|X_t)) \quad \mu\in[0,1]$$

⇒ 为获得估计条件分布 { 充分统计量：用类条件概率近似条件概率
用一个弱分类器生成目标域的初始软标签 $\hat{y}_t$

$$D(D_s,D_t)\approx(1-\mu)D(P(X_s),P(X_t))+\mu D(P(X_s|y_s),P(X_t|\hat{y}_t))$$

当 $\mu\to 0$，表示源域和目标域存在较大差异，边缘分布适配更重要（TCA）

当 $\mu\to 1$，———— 有较高的相似性，条件概率分布适配更重要

当 $\mu=0.5$，JDA

⇒ 使用MMD衡量不同分布中两个域之间的距离

$$D(D_s,D_t)\approx(1-\mu)\left\|\frac{1}{n}\sum_{i=1}^{n}X_{si}-\frac{1}{m}\sum_{j=1}^{m}X_{tj}\right\|_H^2+\mu\sum_{c=1}^{C}\left\|\frac{1}{n_c}\sum_{X_{si}\in D_s^{(c)}}X_{si}-\frac{1}{m_c}\sum_{X_{tj}\in D_t^{(c)}}X_{tj}\right\|_H^2$$

⇒ 特征转换，PCA降维 A，边缘、条件分布适配，优化，得到

$$\min\ tr\left(A^TX\left((1-\mu)M_0+\mu\sum_{c=1}^{C}M_c\right)X^TA\right)+\lambda\|A\|_F^2$$

$$s.t.\quad A^TXHX^TA=I,\quad 0\le\mu\le1$$

△ 关于平衡因子 μ 的求解与估计，使用A-distance来估计（计算两个领域数据整体和局部的A-distance）

A-distance：用来估计不同分布之间的差异性，被定义为建立一个线性分类器来区分两个领域数据的hinge损失(e

即在源域和目标域上训练一个二分类器h，使其可以区分样本来自哪个领域

用err(h)表示分类器h的损失，则 A-distance 定义为

$$A(D_s,D_t)=2(1-2err(h))$$

在BDA中 ①求解源域S和目标域T整体的A-distance，A_1

② 对目标域聚类，计算源域和目标域每个类的A-distance，A_2（求和，取平均）

③ 计算上述两个距离的比值，即为平衡因子 $\mu=\frac{A_1}{A_1+A_2}$