



## 半监督学习

数据科学与计算机学院 17大数据与人工智能

17341015 陈鸿峥

**问题 1.** 分析生成式方法、半监督SVM、图半监督学习、半监督聚类等半监督方法的相同和不同之处。

**解答.** 这几种方法都属于半监督学习的范畴，也就是让学习器不依赖外界交互、自动利用未标记样本来提升学习性能。他们都需要外部的信息或是一些预设的模型进行构造。

生成式方法会假设所有数据（无论是否有标记）都是由同一个潜在的模型生成的，而不同的模型假设将会产生不同的方法。半监督SVM则是直接给定了“低密度分隔”的基本假设，试图找到能将两类有标记样本分开，且穿过数据低密度区域的划分超平面。图半监督学习则给数据集中的样本强加了相似性/相关性，并依此构造了带“颜色”标注的图，通过“颜色”在图上扩散或传播的过程来实现学习。半监督聚类充分利用了外部监督信息，包括同簇异簇的限制、少量的标记样本等，这些限制是与其他半监督方法的较大不同之处。