

1. 下列属于无监督学习的是：（ ）

- A. k-means
- B. SVM
- C. 最大熵
- D. CRF

正确答案：（A）

解析：

A 是聚类，BC 是分类，D 是序列化标注，也是有监督学习。

2. 下列哪个不属于 CRF 模型对于 HMM 和 MEMM 模型的优势（ ）

- A. 特征灵活
- B. 速度快
- C. 可容纳较多上下文信息
- D. 全局最优

答案为 B。

解析：

首先，CRF，HMM(隐马模型)，MEMM(最大熵隐马模型)都常用来做序列标注的建模。

隐马模型一个最大的缺点就是由于其输出独立性假设，导致其不能考虑上下文的特征，限制了特征的选择。

最大熵隐马模型则解决了隐马的问题，可以任意选择特征，但由于其在每一节点都要进行归一化，所以只能找到局部的最优值，同时也带来了标记偏见的问题，即凡是训练语料中未出现的情况全都忽略掉。

条件随机场则很好的解决了这一问题，他并不在每一个节点进行归一化，而是所有特征进行全局归一化，因此可以求得全局的最优值。

3. 在 HMM 中,如果已知观察序列和产生观察序列的状态序列,那么可用以下哪种方法直接进行参数估计()

- A. EM 算法
- B. 维特比算法
- C. 前向后向算法
- D. 极大似然估计

正确答案：D

解析：

EM 算法：只有观测序列，无状态序列时来学习模型参数，即 Baum-Welch 算法

维特比算法：用动态规划解决 HMM 的预测问题，不是参数估计

前向后向算法：用来算概率

极大似然估计：即观测序列和相应的状态序列都存在时的监督学习算法，用来估计参数

注意的是在给定观测序列和对应的状态序列估计模型参数，可以利用极大似然估计。如果给定观测序列，没有对应的状态序列，才用 EM，将状态序列看不可测的隐数据。

4. 下列哪个不属于 CRF 模型对于 HMM 和 MEMM 模型的优势（ ）

- A. 特征灵活
- B. 速度快
- C. 可容纳较多上下文信息
- D. 全局最优

正确答案：(B)

解析：

CRF 的优点：特征灵活，可以容纳较多的上下文信息，能够做到全局最优 CRF 的缺点：速度慢
CRF 没有 HMM 那样严格的独立性假设条件，因而可以容纳任意的上下文信息。特征设计灵活（与 ME 一样）——与 HMM 比较
同时，由于 CRF 计算全局最优输出节点的条件概率，它还克服了最大熵马尔可夫模型标记偏置(Label-bias) 的缺点。——与 MEMM 比较
CRF 是在给定需要标记的观察序列的条件下，使用维特比算法，计算整个标记序列的联合概率分布，而不是在给定当前状态条件下，定义下一个状态的状态分布。——与 ME 比较

5. 解决隐马模型中预测问题的算法是？()

- A. 前向算法
- B. 后向算法
- C. Baum-Welch 算法
- D. 维特比算法

正确答案：(D)

解析：

A、B：前向、后向算法解决的是一个评估问题，即给定一个模型，求某特定观测序列的概率，用于评估该序列最匹配模型。
C：Baum-Welch 算法解决的是一个模型训练问题，即参数估计，是一种无监督的训练方法，主要通过 EM 迭代实现；
D：维特比算法解决的是给定一个模型和某个特定的输出序列，求最可能产生这个输出的状态序列。如通过海藻变化（输出序列）来观测天气（状态序列），是预测问题，通信中的解码问题。

6. 请选择下面可以应用隐马尔科夫(HMM)模型的选项：()

- A. 基因序列数据集
- B. 电影浏览数据集
- C. 股票市场数据集
- D. 所有以上

答案：(D)

解析：

只要是和时间序列问题有关的，都可以试试 HMM。

7. 简述 EM 算法、HMM、CRF。

解析：

- (1) EM 算法

EM 算法是用于含有隐变量模型的极大似然估计或者极大后验估计，有两步组成：E 步，求期望（expectation）；M 步，求极大（maximization）。本质上 EM 算法还是一个迭代算法，通过不断用上一代参数对隐变量的估计来对当前变量进行计算，直到收敛。

注意：EM 算法是对初值敏感的，而且 EM 是不断求解下界的极大化逼近求解对数似然函数的极大化的算法，也就是说 EM 算法不能保证找到全局最优值。对于 EM 的导出方法也应该掌握。

（2） HMM 算法

隐马尔可夫模型是用于标注问题的生成模型。有几个参数（ π , A, B）：初始状态概率向量 π ，状态转移矩阵 A，观测概率矩阵 B。称为马尔科夫模型的三要素。

马尔科夫三个基本问题：

概率计算问题：给定模型和观测序列，计算模型下观测序列输出的概率。—》前向后向算法

学习问题：已知观测序列，估计模型参数，即用极大似然估计来估计参数。—》Baum-Welch(也就是 EM 算法)和极大似然估计。

预测问题：已知模型和观测序列，求解对应的状态序列。—》近似算法（贪心算法）和维比特算法（动态规划求最优路径）

（3） 条件随机场 CRF

给定一组输入随机变量的条件下另一组输出随机变量的条件概率分布密度。条件随机场假设输出变量构成马尔科夫随机场，而我们平时看到的大多是线性链条随机场，也就是由输入对输出进行预测的判别模型。求解方法为极大似然估计或正则化的极大似然估计。

之所以总把 HMM 和 CRF 进行比较，主要是因为 CRF 和 HMM 都利用了图的知识，但是 CRF 利用的是马尔科夫随机场（无向图），而 HMM 的基础是贝叶斯网络（有向图）。而且 CRF 也有：概率计算问题、学习问题和预测问题。大致计算方法和 HMM 类似，只不过不需要 EM 算法进行学习问题。

（4） HMM 和 CRF 对比

其根本还是在于基本的理念不同，一个是生成模型，一个是判别模型，这也就导致了求解方式的不同。

8. 采用 EM 算法求解的模型有哪些，为什么不用牛顿法或梯度下降法？

解析：

用 EM 算法求解的模型一般有 GMM 或者协同过滤，K-means 其实也属于 EM。EM 算法一定会收敛，但是可能收敛到局部最优。由于求和的项数将随着隐变量的数目指数上升，会给梯度计算带来麻烦。