

# 重庆某双一流自然语言处理 2024/6/6 期末考试真题

回忆版，答案不一定对，具体题目是啥也不知道，只是大致描述。

## 一、 填空题

1. \_\_\_\_\_ 技术是将词汇表转化为词向量。 词嵌入？
2. CRF 特征函数有\_\_\_\_\_和\_\_\_\_\_。 状态特征函数、转移特征函数
3. 机器学习损失函数\_\_\_\_\_。 交叉熵
4. 支持向量机通过\_\_\_\_\_分类。 找到最大类间界限的决策平面
5. 对非 0 事件无条件削减某一固定的出现次数值，节留出来的概率均分给 0 概率事件是\_\_\_\_\_数据平滑方法。 绝对减值法
6. LSTM 有\_\_\_门、输入门、\_\_\_门。 遗忘、输出
7. word2vec 有哪两种模型 \_\_\_\_\_和\_\_\_\_\_。 CBOW?、C&W?

## 二、 简答题

### 1. NLP 定义

- NLP, Natural Language Processing
- 用机器处理人类语言的理论和技术
- 将语言做为计算对象,并研究相应的算法、模型
- 目的是让人类可以用自然语言形式跟计算机系统  
统进行人机交互，从而更便捷、有效地进行信息  
管理

### 2. 最大似然估计原理

#### ◆最大似然估计(Maximization likelihood estimation, MLE)

如果一个实验的样本空间是  $\{s_1, s_2, \dots, s_n\}$ ，在相同情况下重复实验  $N$  次，观察到样本  $s_k$  ( $1 \leq k \leq n$ ) 的次数为  $n_N(s_k)$ ，则  $s_k$  的相对频率为：

$$q_N(s_k) = \frac{n_N(s_k)}{N} \quad (2)$$

由于  $\sum_{k=1}^n n_N(s_k) = N$ ，因此，  $\sum_{k=1}^n q_N(s_k) = 1$

当 $N$ 越来越大时，相对频率  $q_N(s_k)$  就越来越接近  $s_k$  的概率  $P(s_k)$ 。事实上，

$$\lim_{N \rightarrow \infty} q_N(s_k) = P(s_k) \quad (3)$$

因此，相对频率常被用作概率的估计值。这种概率值的估计方法称为最大似然估计。

### 3. 什么是梯度消失和梯度爆炸

#### ● 梯度消失或爆炸

梯度爆炸是指在模型训练过程中，参数的梯度值变得非常大，导致模型参数更新过大，从而影响模型的收敛稳定性。

梯度消失是指在模型训练过程中，参数的梯度值变得非常小，接近于零，导致模型参数更新缓慢，从而影响模型的训练效果。

### 4. 词性标注是什么，有什么作用？



## 7.5 词性标注

词性标注是在已经切分好的文本中，给每一个词标注其所属的词类，例如动词、名词、代词、形容词。

词性标注对后续的句子理解有重要的作用。

## 三、 计算题

1. 假设我们有一个句法分析器，它针对一个测试集进行了句法分析。该测试集包含 100 个句子。句法分析器正确地分析了 75 个句子的句法结构，但有 15 个句子被错误地分析了，还有 10 个句子没有被分析（即系统无法给出句法结构）。计算句法分析器的准确率（Accuracy）、召回率（Recall）或查全率、F1 分数（F1 Score）。

准确率：  $75 / (100 - 10) \approx 0.83$

查全率：  $75 / 100 = 0.75$

F1 分数：  $2 * 0.83 * 0.75 / (0.83 + 0.75) \approx 0.788$

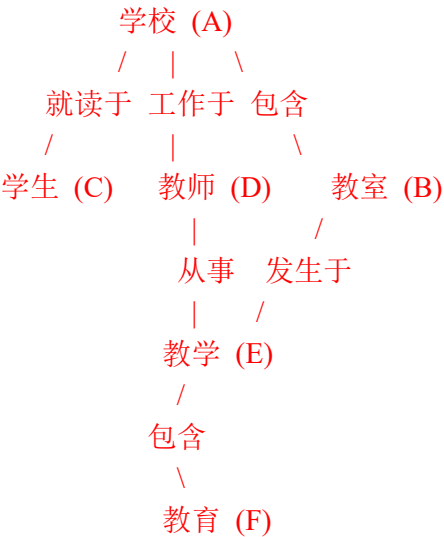
2. 使用语义网络分析以下概念之间的关系：

- 概念 A: 学校
- 概念 B: 教室
- 概念 C: 学生
- 概念 D: 教师
- 概念 E: 教学
- 概念 F: 教育

请根据以上概念描述它们之间的关系，并用语义网络图示出这些关系。

在语义网络中，可以描述以下概念之间的关系：

1. 学校 (A) 和 教室 (B) 之间存在“包含”关系，即学校包含多个教室。
2. 学校 (A) 和 学生 (C) 之间存在“就读于”关系，即学生就读于学校。
3. 学校 (A) 和 教师 (D) 之间存在“工作于”关系，即教师在学校工作。
4. 教室 (B) 和 教学 (E) 之间存在“发生于”关系，即教学活动发生在教室。
5. 教师 (D) 和 教学 (E) 之间存在“从事”关系，即教师从事教学活动。
6. 教学 (E) 和 教育 (F) 之间存在“包含”关系，即教学是教育的一部分。



3. （类似的题，改了参数，原题只需计算两天，标红处为原题答案）

有一个小城，这里的天气只有晴天和阴天，有一个小女孩平时经常做的就是三件事：散步、购物、打扫。有一个在外地的男朋友，想知道她的城市里最近三天天气如何，但是女孩只告诉他，第一天她去购物，第二天打扫，第三天在散步。

现在男生用这现有的信息要推断出小女孩城市里这三天天气最可能的状态。

已知：

隐含的天气状态 = { 晴 ， 阴 }

可观察的感觉状态 = { 购物， 散步， 打扫 }

初始概率分布 = { 晴： 0.6 ， 阴： 0.4 }

而且小女孩城市天气状态转换概率分布 =

	晴	阴
晴	0.7	0.3

	晴	阴
阴	0.4	0.6

而且男孩知道晴天和阴天状况下，小女孩活动情况的概率分布

	散步	购物	打扫
晴	0.5	0.4	0.1
阴	0.1	0.3	0.6

小女孩连续三天的活动依次是： 购物、散步、打扫 。

此处通过 Viterbi 算法进行计算即可。算法如下

## 6.5 Viterbi 搜索算法

递归计算： $\delta_t(j) = \max_{1 \leq i \leq N} [\delta_{t-1}(i) \cdot a_{ij}] \cdot b_j(O_t)$ ,   
 三选一

### ● 算法6.3: Viterbi 算法描述

(1) 初始化： $\delta_1(i) = \pi_i b_i(O_1)$ ,  $1 \leq i \leq N$

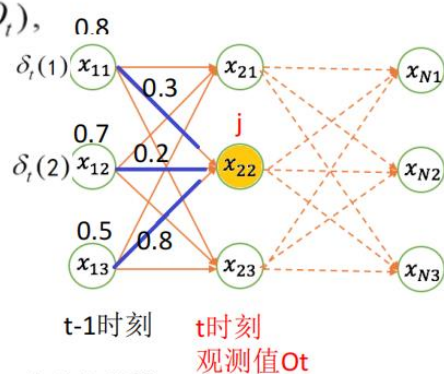
概率最大的路径变量： $\psi_1(i) = 0$

(2) 递推计算：

$$\delta_t(j) = \max_{1 \leq i \leq N} [\delta_{t-1}(i) \cdot a_{ij}] \cdot b_j(O_t), \quad 2 \leq t \leq T, \quad 1 \leq j \leq N$$

$$\psi_t(j) = \operatorname{argmax}_{1 \leq i \leq N} [\delta_{t-1}(i) \cdot a_{ij}] \cdot b_j(O_t), \quad 2 \leq t \leq T, \quad 1 \leq i \leq N$$

记录j的最优前驱





## 6.5 Viterbi 搜索算法

(3) 结束:

T时刻, 所有状态中 $\delta$ 最大的那个状态i

$$\hat{Q}_T = \underset{1 \leq i \leq N}{\operatorname{argmax}} [\delta_T(i)], \quad \hat{p}(\hat{Q}_T) = \max_{1 \leq i \leq N} \delta_T(i)$$

(4) 通过回溯得到路径 (状态序列) :

$$\hat{q}_t = \psi_{t+1}(\hat{q}_{t+1}), \quad t = T-1, T-2, \dots, 1$$

### 四、 综合题

简述三种 N-gram 模型的应用场景, 并描述其在对应场景下的主要作用。