

2022 年 NLP 考试回忆版

考试时间 2023.02.14

前两年回忆版可见 <https://github.com/VayneDuan/cqu-cs-learning-materials/>

填空题 10 * 2

1. K-L 散度又叫做 相对熵
2. LSTM 包含哪几个门 遗忘门、输入门、输出门
3. “喜欢乡下的孩子”包含哪两种意思_____、_____
4. 人工神经网络常用的两种激活函数_____、_____
5. CRF 模型中的两种特征函数是_____、_____
6. 句法分析任务分为短语结构分析和_____

简答题 4 * 5

1. 请简述什么是自然语言处理
2. 请简述什么是自动文摘技术
3. 请简述后退 (backoff) 平滑的原理
4. 互信息的作用

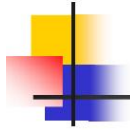
计算题 45 (过去两年考过 hmm 前向后向算法、维特比搜索、cnn 模型、n 元文法)

(CRF、RNN、PCFG 等比较复杂的题其实不用看)

1. 句向量为 (1, 1, 1, 0, 1) 和 (1, 1, 1, 1, 0), 求两个句向量的欧式距离和余弦相似度
2. 前向、逆向最大分词, 给出字典和最大字长和待分词句子
 - a) 求前向逆向的分词结果
 - b) 给出正确分词, 求两种结果的正确率和召回率
3. 线图句法分析 the man saw the dog
过程 + 句法分析树 (ppt 中有详细过程)

综合应用题 15 (可能因为疫情只考了一道题)

1. 现有一医院想对电子病历的症状进行识别, 采用将症状转换为序列标注的方法 (题目记不住了)。请设计一个症状识别系统 (即识别类型 + 对应含义), 以及对应识别的模型。



7.4 未登录词识别

- 除人名、地名、组织机构外，还有很多其它的命名体类别。
- 如医疗领域中，常见的命名体有**部位、疾病名、症状、药物、检查、手术**等。

患者4月前发现皮肤、巩膜黄染，伴食欲上降，晚餐后明显，时有阵发性腹痛、恶心，无腹泻、呕吐，时有胸闷、憋气、头晕，无头痛，无视物旋转，无发热、咳嗽，无胸痛、喘憋，大便颜色较前变浅。于我科住院，行**腹部增强MRI+MRCP**：胆总管末端占位，胆管癌可能性大，右肾下腺结节，考虑**腺瘤**，**右肾囊肿**，**两侧胸腔积液**。请肝胆外科会诊，认为患者无手术指征，于**2017-4-24**全麻上行**ERCP+胆总管金属支架植入**，术后患者恢复良好。并给以抗感染、抗肿瘤、免疫调节、平稳降糖、降压等对症治疗，全身及巩膜黄染逐渐减轻，化验指标逐渐改善。

CRF 和 HMM 应该都行