

2023NKUFall IR Lab2 论文阅读 实验课讲稿

第二次作业为顶会论文阅读与讲解，随机分配每位同学一篇论文。如果大家觉得指定论文实在阅读有困难的话，请联系助教替换论文并说明情况。

这次作业需要撰写**论文笔记**并录制**讲解视频**。论文笔记的模板分享给了大家，大家以此为基础框架记录笔记，但不一定要用表格形式，内容也可以进行扩展。讲解视频需要**逻辑清晰地讲述论文解决的问题以及方法**等，时间控制在**15分钟以内**。

提交时，请将所有作业内容打包成压缩包，命名格式为：学号_姓名_hw2，例如：1911440_刘旭博_hw2.zip。截止日期**10.11** 提交邮箱：nkuir2023fall@163.com

• Q1: 我该去哪里找论文？

- GoogleScholar (作者信息、引用数、h-index)
- dblp (计算机领域，可以按会议浏览)
- arxiv (作者自己上传的预印版)
- Sci-Hub (知识无价)

• Q2: 我该怎么阅读一篇论文？

- 题外话：喜欢在Pad上手写做批注的同学可以用Note APP，电脑端用AcrobatReader也挺好；喜欢纸质阅读/需要护眼可以打印下来。
- 第一遍通读的时候，不要陷入某一个不明确的概念或方法名称，可以暂时当成黑盒理解，即不深究细节（当然如果整篇文章都是黑盒就读不了了）；如和文章主体内容有关，再详细了解。
- Title: 发表期刊/会议，发表时间，一作单位，
- Abstract、Conclusion: 了解文章idea，以及到底做了什么。
- Introduction: 该文章现有研究背景，为什么要做这个工作(Motivation)，作者的主要Claim。
- Related Work: 有哪些工作可以作为baseline，在哪些benchmark上完成。
- 文章主体内容: idea是怎么想到的？可能包括：模型结构是怎么连接的；如果提了一个很小很trick的算法，这个算法的伪代码、流程图；如果提出了新的evaluation metric或新的方法，给出数学公式？总之，这个部分是需要精读的内容，你需要完整搞懂文章到底说了什么，做了什么？如有余力，可以对文章的部分进行复现（有source code复现起来会节省很多事情）。
- Experiments: 明确benchmark，即用了哪些数据集？明确benchmark的evaluation metrics，用哪些指标评价？这些指标怎么证明了论文笔者的方法比baseline的方法要好？可能会有Ablation Study，或许和对比实验也没什么区别？
- References: 对文章感兴趣的话，进一步了解。

• Q3: 我该怎么写一篇阅读报告？

- 如按照模板，可边读边做一些文章中有明确答案的填空【如模板中，实验内容】；完整理解后再填写一些需要主观总结的内容【如模板中，思考内容】。
- 也可按照自己的方式完成。