

# 科研经历分享



**Reporter:** Ye Liu

**Research area:** NLP (Information Extraction, KG)

**Supervisor:** Prof. En-Hong Chen

**E-mail:** liuyer@mail.ustc.edu.cn

**Date:** 2023-09-26

# 自我介绍



## □ 刘烨

- 博士三年级
- 大数据学院2021级博士班
- 主要研究方向：自然语言处理，知识图谱，LLM
- 一作论文：ICDM2020, TKDD2023, ACL2023
- 合作论文：IJCAI2021, SIGIR2021, DASFAA2023, IJCAI2023, ACL2023等

# Outline



- **怎么开始科研?**
- 我没有 idea , 该怎么办?
- 有了 idea 之后, 如何实现?
- 怎么把 idea 转换成一篇科研论文?

# 怎么开始科研?



## □ 优先和导师沟通

- 基于实验室的科研项目（有导师push）
- 基于实验室过往的方向（有实验室之前的积累）

## □ 听取实验室师兄师姐的意见

- 师兄师姐在更具体、细节上的讨论上也很重要

## □ 要有主观能动性

- 从上课学习 → 自主学习要有心态的转变
- 就算导师很忙，不要害怕因为科研讨论打扰老师

# Outline



- 怎么开始科研?
- **我没有 *idea* , 该怎么办?**
- 有了 *idea* 之后, 如何实现?
- 怎么把 *idea* 转换成一篇科研论文?

# 我没有 idea，该怎么办？



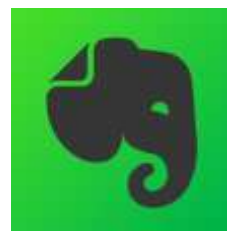
## □ 养成日常阅读论文的习惯

- 广泛阅读大领域的文章
- 在找到一个兴趣点之后，重点阅读其中的一个小领域



## □ 在阅读论文时候，学会进行文献整理

- One-note、印象笔记、PPT等整理软件
- 并且要对已经读过的论文有归类整理

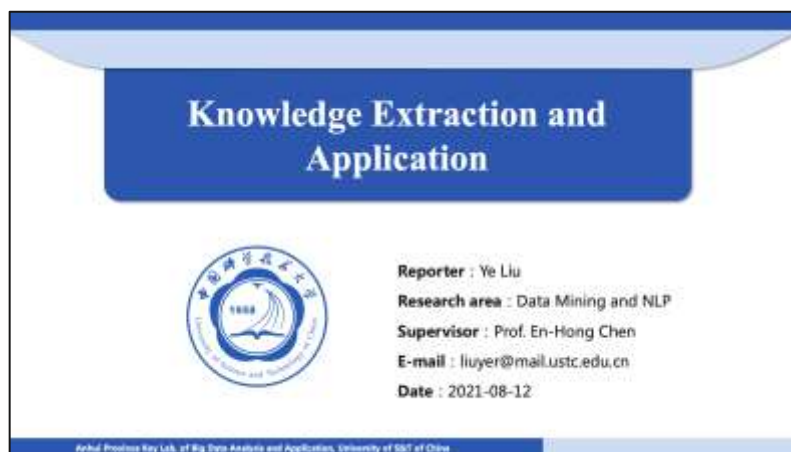


# 我没有 idea , 该怎么办?



## □ 在阅读论文时候，学会进行文献整理

- 学会在阅读时候对论文做批注，可以阅读完一篇文章后用几句话总结、评价一下这篇文章。
- 在完成一个小领域20篇以上的论文阅读后，可以自己整理出一个Survey类型的PPT
- Important! ! !



# 我没有 idea，该怎么办？



## □ 关注一些科研公众号

- 可以利用碎片化的时间了解一些科研热点
- 现在文章爆炸式增长，尤其是和深度学习相关的领域
- 尤其是很多没必要深入阅读的文献，整理好的推文能极大的节省时间



深度学习与NLP

...

专注深度学习、NLP相关技术、资讯，追求纯粹的技术，享受学习、分享的快乐。



PaperWeekly

...

PaperWeekly是一个推荐、解读、讨论和报道人工智能前沿论文成果的学术平台，致力于让国内外优秀科研工作得到...



专知 专知

...

专知，为人工智能从业者服务，提供专业可信的人工智能知识与技术服务，让认知协作更快更好！



夕小瑶科技说

...

更快的AI前沿，更深的行业洞见。聚集25万AI应用开发者、算法工程师和研究人员。一线作者均来自清北、国外...



# 我没有 idea，该怎么办？



- 大家都告诉我要多看，可是我看了很多还是没有idea？
  - 不要做单纯的阅读者，要学会在阅读的时候思考
  - 如果在充分的阅读后，还没有具体的想法。可以考虑动起手来，复现一些方法。看论文和写实验有时候完全是两个世界的感受
  - 多和同门讨论，头脑风暴是产生 idea 的最好时机



# 我没有 idea，该怎么办？



- 要有一定程度的“摆烂”心态
  - 不要压力过大，神经过于紧张时候更难想到 idea
  - 很多想法都是在不经意间涌入脑海的
  - “摆烂”不是躺平，不多阅读、多思考，idea肯定不会凭空出现



# Outline



- 怎么开始科研?
- 我没有 idea , 该怎么办?
- **有了 *idea* 之后, 如何实现?**
- 怎么把 idea 转换成一篇科研论文?

# 有了 idea 之后，如何实现？



## □ 按部就班，先确定做实验的数据集

- 最简单的方式，别人用什么数据集，自己就用什么，如果模型的基础设置与别人的一样，还可以直接把别人的结果拿过来，不用重新跑baseline
- 数据集比较少见，需要自己收集，处理

Dataset	$ Y $	Depth	$\text{Avg}( y_i )$	Train	Dev	Test
WOS	141	2	2.0	30,070	7,518	9,397
NYT	166	8	7.6	23,345	5,834	7,292
RCV1-V2	103	4	3.24	20,833	2,316	781,265

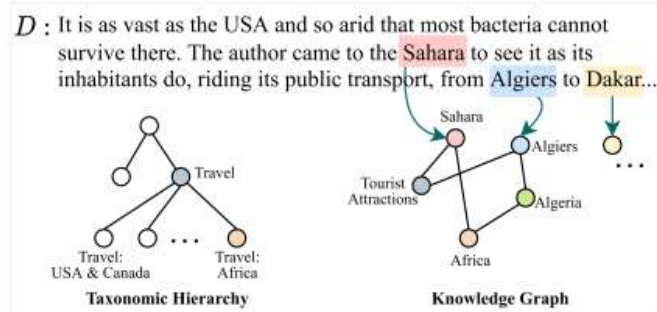
Table 1: Data Statistics.  $|Y|$  is the number of classes. Depth is the maximum level of hierarchy.  $\text{Avg}(|y_i|)$  is the average number of classes per sample.

# 有了 idea 之后，如何实现？



## □ 确定模型的架构

- 自己设计模型？在别人的架构上修改？
- 切忌一口想吃成个胖子，学会做预实验，验证自己的motivation是否合理
- 要学会代码复用，避免重复的无意义劳动



Hierarchical Level	BGC	WOS
L-1	4.29	5.82
L-2	4.93	8.00
L-3	5.96	—
L-4	5.94	—
Total	3.12	4.87

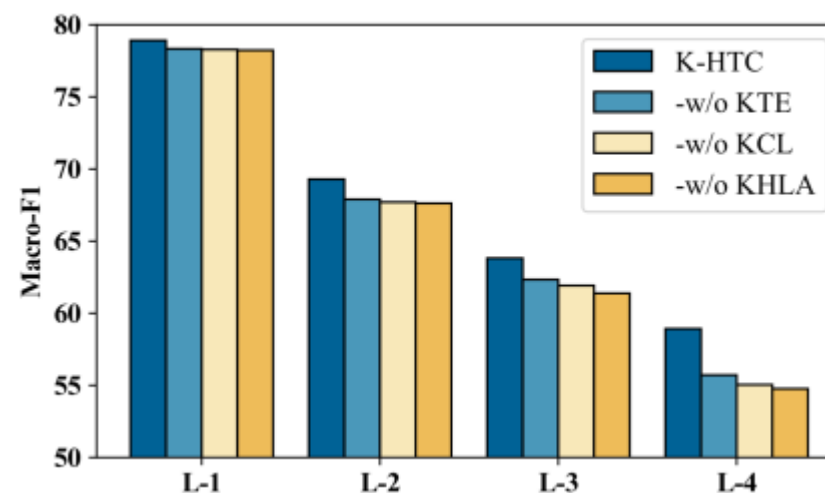
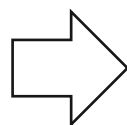
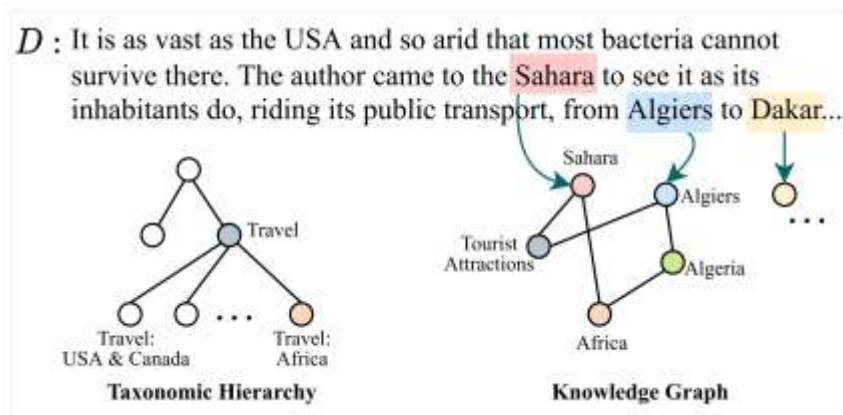
$$L_c^{ij} = -\beta_{ij} \log \frac{e^{-d(z_i, z_j)/\tau}}{\sum_{k \in g(i)} e^{-d(z_i, z_k)/\tau}},$$
$$c_{ij} = |C_i \cap C_j|, \quad \beta_{ij} = \frac{c_{ij}}{\sum_{k \in g(i)} c_{ik}},$$

# 有了 idea 之后，如何实现？



## □ 做好实验设计

- 不要单单关注主实验，并不是最终效果好就万事大吉了
- 主实验之外，学会设计一些能够验证motivation的实验
- 思考一个让人眼前一亮的实验
- 做好实验的log记录、checkpoints，方便之后使用



# Outline



- 怎么开始科研?
- 我没有 idea , 该怎么办?
- 有了 idea 之后, 如何实现?
- **怎么把 *idea* 转换成一篇科研论文?**

# 怎么把 idea 转换成一篇科研论文?



## □ 梳理论文逻辑

- 整理模型设计和实验的逻辑
- 先确定论文的整体框架，现有骨架再去丰富内容
- 第一次写文章时候，可以让导师或者师兄师姐推荐一篇写的比较好的paper进行参考。  
重要的是把逻辑讲清楚，而非意识流的创新！
  - ✓ Background
  - ✓ Related Work
  - ✓ Challenge
  - ✓ Solution
- 等到写过 2 篇文章之后，可以在写作上进行一些新的尝试。



# 怎么把 idea 转换成一篇科研论文?



## □ 论文写作

- 注意写作积累，平时阅读时遇到比较好的表达可以记录下来备用（Enhancing → Empowering）
- 翻译软件
- ChatGPT



## □ 绘图

- 模型图和文字同样重要！
- 选择一些好用的绘图软件，图片的模块/配色也注意复用



# Outline



- 怎么开始科研?
- 我没有 idea , 该怎么办?
- 有了 idea 之后, 如何实现?
- 怎么把 idea 转换成一篇科研论文?
- **写在最后**

# 写在最后



- 做科研需要努力，但更重要的是一个好的心态
  - 为什么我感觉我很努力，但是就是没有进展？
    - ✓ 科研不是上课考试，很多回报是延迟到来的
  - 和我同级的xxx、xxx都已经有了文章了？
    - ✓ Peer Pressure是前进的动力，但是要更多的关注自己的进步
  - 为什么我的文章总是被拒绝？
    - ✓ 拒稿、修订也是一个自我提高的过程
    - ✓ 要辩证的看待审稿人给出的意见，不要盲目迷信

# 写在最后



## □ 做科研需要努力，但更重要的是一个好的心态

还有人在第 6 次投稿努力后，终于收获了第一篇顶会接收论文。



TryProfessional4078 · 6 hr. ago

7655 accepted (poster). Finally got my first accepted paper at the 6th attempt. I'm so happy.

↑ 8 ↓ Reply Share ...

# 写在最后



## □ 科研也并不是生活的全部



热爱生命 汪国真  
我不去想是否能够成功  
既然选择了远方  
便只顾风雨兼程  
我不去想能否赢得爱情  
既然钟情于玫瑰  
就勇敢地吐露真诚  
我不去想身后会不会袭来寒风冷雨  
既然目标是地平线  
留给世界的只能是背影  
我不去想未来是平坦还是泥泞  
只要热爱生命  
一切  
都在意料之中



# **Thank You for Listening !**

**Ye Liu 2023/09/26**