

大创经验分享

郎骐玮 数据科学与大数据技术专业2020级，吉林大学，软件学院

Agenda

以一个从零开始认识大创的视角

一. 了解大创

二. 决定做大创

三. 运营一个大创

四. 其他

五. Q&A

○、关于我的大创

○、关于我的大创

本人大创经历简介

- 2021年6月（约大一下暑假）进入教育部重点实验室基金项目《基于知识演化的人才网络发现》，参与研究工作
- 2022年5月（大二上中旬）大创立项，《基于学术知识图谱的可定制化学术资源采集方法研究》，国家级立项，第一负责人

一、了解大创

一、了解大创

1. 大创是什么（copy from 王欣和吕帅老师去年的讲座）

- 大创：全名大学生创业创新训练计划
- 课外培养计划之一
- 包括：创新训练项目、创业训练项目、创业实践项目
- 主要面向大二的本科生（大一也可以参与）
- 必须组队申报（3-5人），支持鼓励跨专业跨学科组队
- 每人只能同时参加一个项目，项目完成时间基本为1-2年

一、了解大创

2. 一些时间节点（copy from 吕帅老师去年讲座）

- 4-5月（一般为5月中旬前后）：立项评审，确定立项级别【国家级、省级、校级】
- 9-10月：中期检查/定级评审【提交中期检查报告】
- 第二年4-5月、9-10月：进度检查/结题验收
- 结题项目遴选优秀候选项目（不超过结题数的20%、10%，取整），上报学部PK确定优秀结题

二、决定做大创

二、决定做大创 对你而言

1. 你是否需要大创？
2. 你能否做大创？
3. 大创能带给你什么？

二、决定做大创

你是否需要大创？你能否做大创？

- 能力：课内学业压力、自我学习能力、英语能力、建模能力、编程能力、资料搜寻能力、领导能力、团队协作能力.....
- 精力：学业、科研、学生工作等时间的分配和平衡
- 是否真正对科研或者工程问题感兴趣（针对可能考公、选调、考研的同学）
- 是否对某一领域或方向真正感兴趣或有实际的了解
- 是否有保研或出国的相应需求

二、决定做大创

大创能带给你什么？

- 科研经历

- ✱科研能力的培养：快速进入了解一个领域、快速忘掉原有知识并能把原有思维迁移到新知识的学习上（增量学习）、快速上手模型代码和调试训练

- ✱功利角度：成果产出

- 工程经历

- ✱大型系统的开发过程、痛苦的调试增长bug经验（

- ✱软著、专利

三、运营一个大创

三、运营一个大创

运营时间线

1. 准备
2. 立项
3. 执行
4. 结项

三、运营一个大创

1. 准备 —— 关于“找”

- 找方向
- 找队友
- 找老师

三、运营一个大创

1. 准备 —— 关于“找”

- 找方向

- ♣自己感兴趣的方向

- ♣比较热门 or 还能做且比较好做的方向

- 红海方向尽量不要碰

- 太古老的方向尽量不要碰

- Baseline太高的方向尽量去不要做纯研究刷点的研究

三、运营一个大创

1. 准备 —— 关于“找”

- 找队友

- ♣️ 时间点：找老师前 / 后（推荐前）

- ♣️ 想要自己成为的角色：负责人 / 参与者（本院 or 给别的学院打工） / 挂件（不可取）

- ♣️ 作为负责人要考虑的：自己的能力和队友能力（补充 / 互助）、擅长方向的衡量

三、运营一个大创

1. 准备 —— 关于“找”

- 找老师

- ♣途径：自己上的课 / 学长或同学推荐 / 官网找

- ♣方向：尽量和你想做的方向一致或相近，其余和找方向一致

三、运营一个大创

2. 立项

- 项目怎么来？
- 项目的类型？
- 项目规划怎么做？

三、运营一个大创

2. 立项 —— 项目怎么来

- 项目来源

- ◆自己想：自由度较大，但老师接收难度也较大

- 目前的热点方向

- 自己想做的东西（如开发等）

- ◆老师的课题：自由度较小，但更容易找

- 子课题

三、运营一个大创

2. 立项 —— 项目的类型

难度从简单到难：（可能和其他同学或者老师的说法或分类不同）

- 开发 easy：简单的 APP 开发等
- 工程系统 medium / hard：面向实际问题的复杂系统开发或数据挖掘工作，一般是从 0 到 1 的工作
- 科学问题 hard：对于某种方法或某一问题的深入探究和改进

实际项目并不一定是三者之一，很可能是三者之中某几的结合。

三、运营一个大创

2. 立项 —— 项目规划怎么做

① 明确研究边界、应用场景、项目大小、可完成度等

② 调研现状：领域综述、近三到五年内顶刊顶会论文

◆开发：学技术、找现有的类似应用

◆工程：过程中可能使用的方法，找最新顶刊顶会论文及其开源实现（不一定完全和目的一致）

◆科研：明确“研究链条”

③ 确定技术路线或研究和改进的中心

▸ 明确“创新点”：一定是可实现的、结合现状得出的 idea，绝不是凭空想出的

④ 明确预期成果和时间节点：务必切实合理、是一组本科生能够完成的水平

三、运营一个大创

3. 执行

按前面提到的分类讲述：

- 开发
- 工程系统
- 科学问题

三、运营一个大创

3. 执行 —— 开发

- ① 明确应用需求、可能使用到的技术
- ② 学习开发应用的方法、技术
- ③ 分工并按时间节点进行开发、测试

三、运营一个大创

3. 执行 —— 工程系统

- ① 明确系统的各个模块是什么
- ② 明确各个模块需要应用的方法（在准备阶段应该已经基本明确）
- ③ 一定基于现有的方向或方法去研究：找相关方法的论文，确认实现或复现的难易或者有无开源实现（准备阶段也应已经基本明确），确认已有方法和项目涉及问题的边界的区别
- ④ 获取数据（训练模型需要的数据或系统面向数据的样例），如果有需要可能要建立数据库
- ⑤ 开始实现 v1.0 版本：系统往往需要界面展示给人看才能反应其有效性和准确性，命令行效果过弱（一般对于前端需求不会太高和花哨）
- ⑥ 改进方法或更换方法、模型的迭代和调优

三、运营一个大创

3. 执行 —— 科学问题

① 大量阅读所研究问题或方向的顶刊顶会论文

- ◆找方法：作为 baseline、作为提出新 idea 的基本

- ◆找数据集：研究方向常用的测试模型 performance 的数据集

- ◆复现现有方法 / 找到其开源代码

② 发现现有方法问题，提出 idea

- ◆从原论文中发现问题，并进行小修小补

- ◆从综述或其他同行论文中获得灵感

③ 大量实验验证 idea 正确性

- ◆切忌空对空的想法和讨论 —— 只有实验结果才能证明想法的效果好坏

三、运营一个大创

4. 结项 —— 成果

尽量现实一点.....

- 软著：周期在一个月到三个月不等，相对较水

 * 申请步骤：软团团教程，搜索“软著”（基本流程，细节可能有所不同）

- 论文：周期在3-7天、三个月、半年、一年不等

 * EI / SCI

 - EI 水刊水会：3-7天给结果

 - 常规会议及期刊：会议 deadline 明确，周期相对短；期刊随时可投，周期相对长（看实际情况）

 * CCF A/B/C

- 专利：周期基本为一年

四、其他

四、其他

个人向杂谈

1. 相关资源工具推荐
2. 是否需要延期的考虑
3. 立项等级（可能）和什么有关
4. 导师对项目的影响
5. 做大创的心路历程和个人感想
6. 大创怎么影响保研、出国

四、其他

1. 相关资源工具推荐

- 论文网站

- ◆DBLP, Google Scholar, AMiner, Research Gate

- ◆Web of Science, 知网, ACM数据库..... (图书馆)

- ◆Arxiv

- ◆搜索引擎

四、其他

1. 相关资源工具推荐

- 入门及深入学习资源
 - ◆ 吴恩达：机器学习（Stanford CS229）、机器学习深度学习等Coursera系列课程
 - ◆ 李宏毅：机器学习等系列课程
 - ◆ 李沐：动手学深度学习（D2L）、论文精度系列
 - ◆ 周志华：《机器学习》（西瓜书）
 - ◆ 国外各大高校的网上公开课（Stanford、CMU、MIT等）

四、其他

1. 相关资源工具推荐

- 工具

- ◆ ChatGPT

- ◆ 看论文的工具：知云文献翻译、IcySCI、Zotero、Goodnotes等

- ◆ 搜论文、找论文关系：SciHub、Connected Papers、Scite等

- ◆ 找开源代码：GitHub、Papers with Code、机器之心SOTA

- ◆ 低价靠谱算力：AutoDL

四、其他

2. 是否需要延期的考虑

- 项目实现难度
- 项目进度
- 成果进度

四、其他

3. 立项等级（可能）和什么有关 —— 作为负责人

- 前期工作度
- 成果难度、数量
- 立项答辩表现

✱你对项目是否真正了解

✱你的前期工作是否真正扎实

➡实验：复现结果

➡数据：爬取数据的展示 / 可视化（非现成数据集，一般是现实数据清洗后的结果）

➡系统开发：丐版 v0.1 的展示

✱你对领域现状是否真正清晰

四、其他

4. 导师对项目的影响

- 尽可能做什么方向就找什么老师（但实际上很难）
- 子项目性质项目：老师带一个大项目组，里面有好多子项目，老师的方向未必和每一个子项目一致，但共同指导
- 老师的作用是“掌舵”，科研怎么做、技术怎么学还是大部分要靠自己摸索（科研思维的培养和提升）

四、其他

5. 做大创的心路历程和个人感想

- 过程：心态要好、脑子要活、思维要清
- idea = 大量的论文积累 + 对方法技术扎实的理解 + 个人对领域的洞见
一瞬间的灵感 + 集思广益 + 队友合作的讨论
运气和机缘 实验结果有提升

四、其他

6. 大创怎么影响了保研、出国

- 尽可能有成果：accepted 或有 review，不要在投
- 尽可能不要投水刊水会，但有时为了夏令营、出国申请或结项能有成果，也应当做取舍和选择
- 有成果（论文、软著、专利）尽量早投出去，瞄准 deadline 倒推时间，不要有 idea 被抢或都堆到最后没有会议投或者无法结项

五、Q&A