

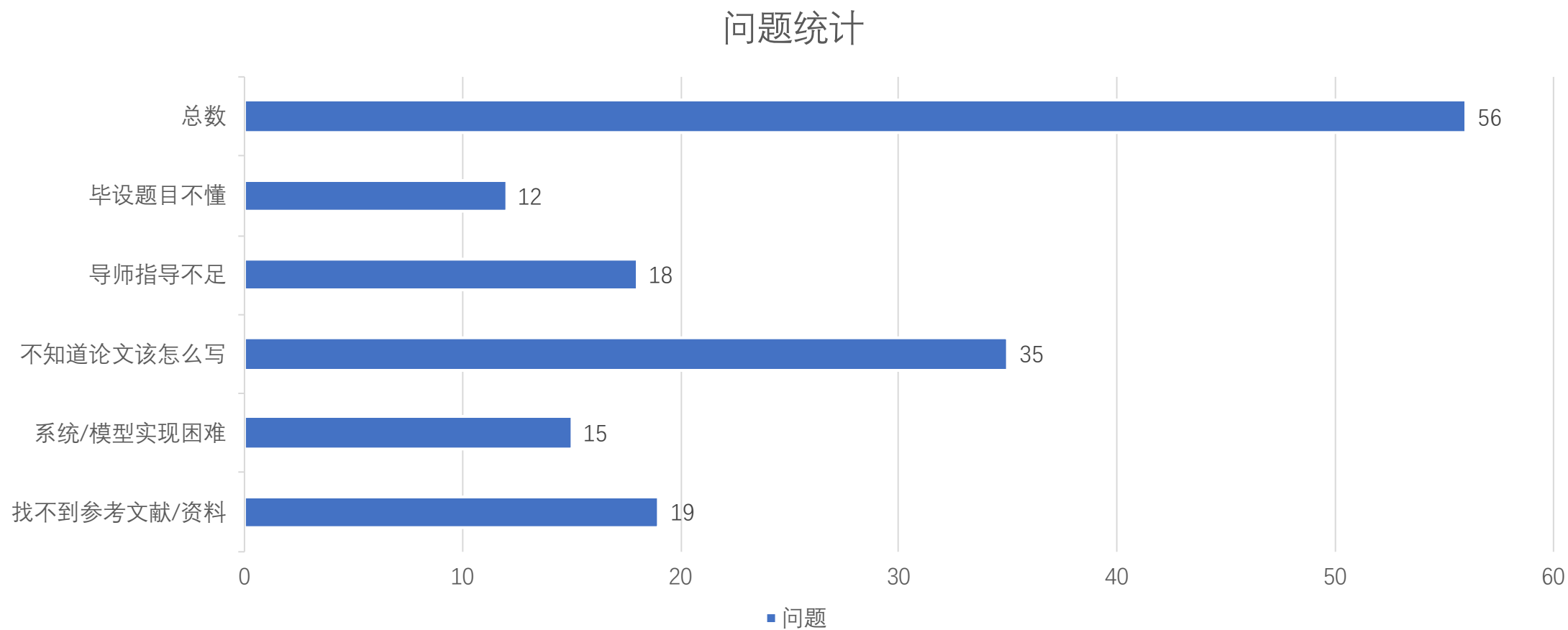
怎样做好计算机毕业设计

2020.2.19

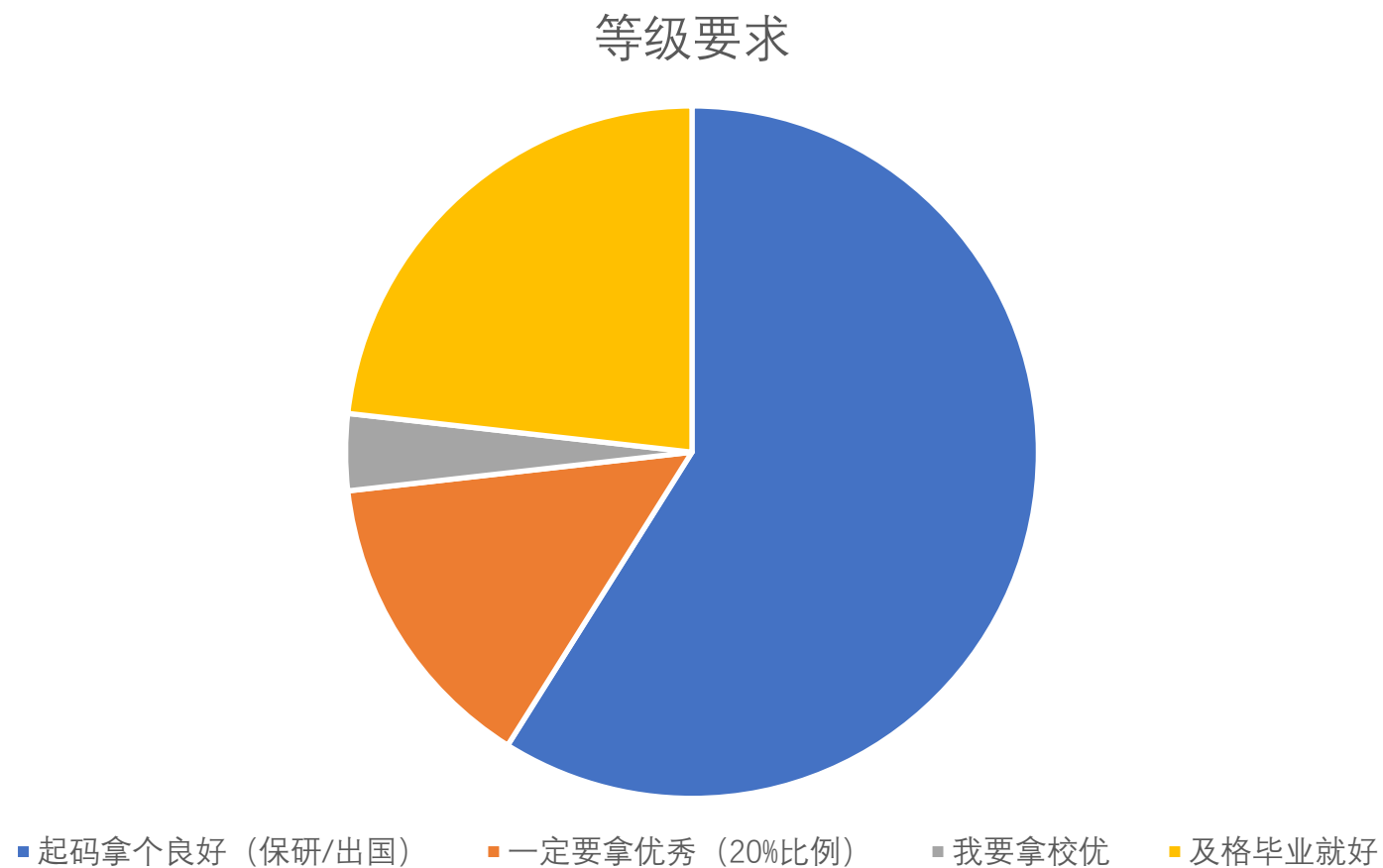
重庆大学计算机学院

吕昱峰

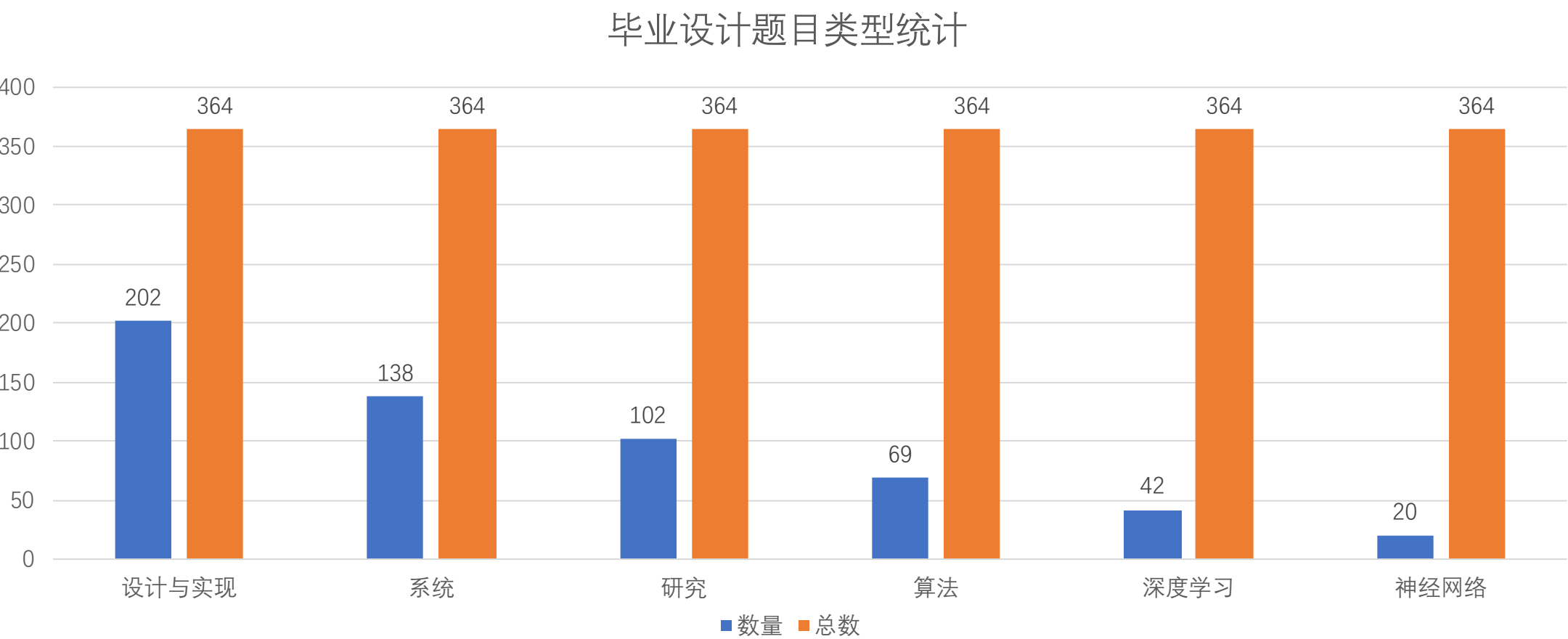
问题统计



毕设完成程度的要求



毕设题目统计



题目不懂怎么办

- 题目大分类：
 - 系统设计与实现
 - 算法/模型研究
- 系统实现类：
 - 技术栈
 - 需求
- 研究类：
 - 搞清楚要研究的问题是什么
 - 研究方向上需要的基础能力有哪些

系统实现类 (Easy)

- 企业IT资产管理平台设计与实现

- 1.通过查阅学习相关开发技术资料以及其他技术资料分析研究，掌握Windows应用开发（C#语言），Windows桌面应用和系统服务的构建，Restful接口技术开发，Web应用开发，熟悉和掌握数据相关应用开发。

需求明确

- 2.系统分为Web展现、Restful接口开发和Windows前端定时检测程序几个部分。Web应用相关开发可选择Java、PHP、C#等。

需要实现的功能明确

技术栈选项众多

最重要的，网上教程一抓一大把

系统实现类 (Medium)

- 基于用户多偏好的实时行车路线推荐系统设计与实现

需要推荐算法，但是有现成的可以使用

- 1、实现基于用户多偏好的实时行车路线推荐系统，系统的基本功能，包括根据用户的输入信息，包括起点、目的地、出发时间、时间预算，及偏好（比如风景质量高、安全系数高）等信息，为用户推荐实时的行车出行路线；为用户建议出行时间，以更好的满足用户偏好。

- 2、设计并实现可视化的用户交互界面，提示用户进行信息输入、将路线推荐结果进行展示。（题目难度适中）

主要还是做交互界面，以实现系统为主，但需要根据需求选择合适的算法

系统实现类 (Hard-Bug)

- 基于FPGA的Q-learning加速器设计与实现

- 1. 学习Q-learning 强化学习算法
- 2. 设计Q-learning算法的硬件设计方案
- 3. 在FPGA上实现Q-learning算法

需要有强化学习支持储备，起码自己复现过Q-learning算法

涉及到矩阵运算的硬件加速知识，本科无相关基础

硬件编程能力必须过关，但是多数人Vivado都不太会用

研究类 (Easy)

- 基于CT图像的肝脏自动分割算法研究与实现

- 1. 熟练掌握MATLAB、Python、C++程序设计等知识；
- 2. 学习图像识别基础知识，掌握图像识别中的基本方法和原理；
- 3. 学习机器学习基本知识，了解深度学习在图像识别中的原理及应用，并学会使用一种深度学习框架，如TensorFlow、PyTorch等；
- 4. 基于机器学习方法，在现有相关文献方法研究基础上，设计并实现一种基于CT图像的肝脏自动分割算法。
- 5. 在腹部CT图像实验数据集上对算法处理效果进行实验验证。

基于DL的图像分割算法已经非常成熟，基本是调库

数据集是现成的

只需要实验验证，没有明确的Baseline，没有SOTA要求

研究类 (Medium)

- 宋词自动生成系统设计与实现

- 1. 宋词语料库分析与词频统计：整理所有的宋词集合，并进行数据转换和处理，统计宋词的单字词、双字词及其频度。
- 2. 宋词自动生成：基于用于输入的词牌，利用语言模型实现宋词生成系统。
- 3. UI界面设计，搭建B/S服务，对宋词生成结果进行展示。

语料库充分，哪怕没有也随处可爬

文本生成任务相对还是较难，生成的内容可能GP不通，但是网上开源Demo不少

最后落脚处还是系统设计，但是文本生成是重点

研究类（Hard 一般是保研导师的题目）

- 机器阅读理解模式下多轮对话系统

- 1. 通过阅读外文文献理解预训练语言模型、多轮对话系统等原理与应用
- 2. 实现基线系统并验证，用于机器阅读理解任务
- 3. 在对话系统环境下进行测试调优、系统改进，提出新的针对性模型设计方案

大量英文文献阅读的要求

一般都会有对比的Baseline或者其他指标要求

需要Performance更好，或者创新性要求高

对待题目

- 搞清楚是什么难度，有任何不确定的问题，抓紧时间联系导师问
- 老师给的任务书上的文献不见得靠谱，如果发现题目和任务书不对应，向老师继续索要
- 不管多难，跟最终结果无关，搞清楚题目抓紧动手

开题报告怎么写

- 一、课题的目的及意义
- 二、重点研究内容及实现途径
- 三、课题预期成果
- 四、进度计划

一、课题的目的及意义

- (一) 课题任务
 - 这里大部分都是从任务书复制过来即可
 - 如果任务书实在没几个字，适当扩充一下
 - 可以跟导师讨论细化任务后填在这部分
- (二) 课题研究意义
 - 一般情况下会有一定的社会意义/研究意义（都是套话，硕士论文里面多的很）
- (三) 国内外研究现状
 - 一定要把任务书给的参考文献读了写上去（xxx于xxx时间，做了xxx，达到了xxx效果/性能/指标得分等等）
 - 除此之外近期的发展可以适当写入（2019.12-2020.02）
 - 做系统开发的同学可以写大厂在这方面的新技术（比如阿里每年出的技术手册）

二、重点研究内容及实现途径

- (一) 重点研究内容
 - 搞清楚最难做的几个点写上去
 - 句子生成连贯性问题（宋词生成题目）
 - 分割边缘准确性问题（CT图像分割题目）
 - 系统实时性如何解决（路线规划推荐系统题目）
 - 废话最好少写，一般一个题目的重点研究内容不超过5个
- (二) 实现途径
 - 根据重点研究内容，对应研究方案/实现路径/技术选型
 - 如果已经有成熟的方案，可以把方案写上

三、课题预期成果

- (一) 预期成果 (展现形式、性能得分等)
 - 简单列举最终达到的效果和形式
 - 如果是研究类题目, 要有对应的指标
 - 研究类题目也最好有个Demo
- (二) 预期成果特色/创新性
 - 能做到什么程度目前不知道, 但是可以适当吹一下
 - 一些副产物也可以写上去, 比如构造了数据集。

四、进度计划

- 如果导师有安排，写导师要求的
- 如果自己有计划，写上
- 如果都没有，大致列一下研究/实现步骤，每个给1-2周即可

怎么搜集资料

- CNKI的文献质量可以用Garbage来形容
- 外文文献和大厂技术报告才是参考文献的重头
- 文献管理必须做！！！！

文献管理软件

- Zotero/Mendeley(外文文献)
- CNKI E-Study (中文文献)
- 为什么要用文献管理软件？
- 统一管理方便查看（说下载pdf存一个单独目录的，你随意）
- Word/Latex参考文献自动插入（包括标号和最后的引文列表）

文献管理软件教程

- Zotero
 - <http://www.jianguoyun.com/s/content/1301.html>
 - 视频教程：<https://www.bilibili.com/video/av45388595>
- Mendeley
 - 视频教程：<https://www.bilibili.com/video/av15751403>
- CNKI E-Study
 - <http://estudy.cnki.net/CNKI%20E-Study%204.0help.pdf>
 - 视频教程：<https://www.cnki.net/elearning/help2.htm>

几个论文检索网站（非CNKI）

- 图书馆可以用的（VPN连内网，图书馆数据库导航找）：
 - [Web of Science](#)（SCI检索）
 - [Elsevier ScienceDirect](#)
 - [Engineering Village（EICompendex）](#)（EI检索）
- 外网可用的：
 - Google Scholar
 - ArXiv

准备好了，何处动手？

- 对于研究类题目：
 - 最好有Baseline复现之后
 - 基本知识可能需要补充，但是不要超过两周，可以边学边做，不然一定颓掉
 - 论文阅读量最好保证5篇以上（该方向经典论文，个位数citation的就别看了）
- 对于开发系统的题目：
 - **需求分析！需求分析！需求分析！**
 - 别着急coding，先把设计做了，软件工程的流程回忆起来
 - 数据库如果只有个user表劝你还是别开工

开源项目/代码该不该用？

- 首先，肯定的，可以用
- 其次，用归用，能不能把人家的github仓库地址写在你的引用上，用了还不引用的这叫学术不端
- Why ?
 - 一般情况下开源项目不会跟你的题目完全一致
 - 重复造轮子确实也很蠢
 - 毕业**设计**重点在于你的设计，不论算法/模型/系统，有针对性的进行改造/改进/适配，都足以达到要求，**唯独抄袭不行**

何时着手写论文？

- 所有人，立刻开始
- 对于做系统的同学：
 - 像设计文档一样，可以先写个概要设计，然后加上引言和结论就是基本的论文了
 - 论文里别贴代码（写实验报告写习惯了的）
- 对于做研究的同学：
 - 先写Proposal
 - 再写实验部分
- @所有人，论文不是顺序写的！！！！

论文重要还是实现重要？

- 都重要，也都不重要
- 论文和实现都是评价的重要组成部分（中期检查）
- 没有实现，论文无异于胡编乱造
- 论文没内容，做得再好也显得单薄（很多因此拿良的）
- 为啥都不重要，因为答辩最重要

毕设对于创新性的要求

- 多数都是要求实现/复现/验证，对创新性没有过多要求
- 保研的题目大多数由导师研究方向/项目输出，基本算是研究生的要求，对创新性要求较高
- 创新性不是毕设的核心要求，能够考察本科四年培养的能力是否达到要求才是
- 有创新一定更加分
- 小的点都是创新，比如UI设计的更人性化/模型在某指标提高了2-3%

如何最大化指导老师的指导

- 不要脸，不要脸，还是不要脸
- 我们学院的学生肯主动联系老师的还是少
- 指导老师们忙得很，别指望主动找你（部分年轻老师会）
- 这是你自己的事情，缺资料、缺设备到哪去找？厚脸皮要
- 微信/QQ电话走起来。
- 你要是每周都主动跟指导老师开一次会，大概保底也得是个良

假如指导老师不靠谱

- （我重还是有这种的）
- 一个指导老师的学生抱团取暖（大多数题目相关性比较强）
- 找类似题目同学
- 找上届学长学姐咨询
- 找TB的，祝你好运

关于抄袭/查重

- 重复率一般要求20%以下，维普查重，一次大概几十RMB
- 一般我们专业很容易控制在10%以下
- 标注好引用，这不算在重复率里
- CSDN/简书那些教程就别去抄了，不然第一次查重你一定想哭

统计问卷的其他问题

- 针对最后的评判标准不清楚
- 毕设时间安排会改变吗/各项时间安排是否依旧如毕设网上一样？何时开始写周记？
- 毕业设计必须要自己独创一种新方法吗？如果就是把现有的技术综合一下，将目前比较前沿的论文复现，可以吗？这样能获得良好或优秀吗？
- 疫情导致无法去实验室/在家没有电脑 网不好
- 怎么开始呢