

《计算科学导论》课程总结报告

|  |  |
| --- | --- |
| 姓 名 | 高铱蔓 |
| 学 号 | 2407010304 |
| 专业班级 | 计科2403 |
| 学 院 | 计算机科学与技术学院 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 课程认识  30% | 问题思考  30% | 格式规范  20% | IT工具  20% | 总分 | 评阅教师 |
|  |  |  |  |  |  |

2024年11月24日

# 1 引言

站在信息时代的潮头，从互联网的普及到智能手机的应用，从大数据分析到人工智能的兴起，计算机的影响无处不在，深刻地改变了人类社会的面貌。在过去的几个月里，《计算科学导论》课程带领我们踏上了这段非凡的学术征途，不仅解锁了计算世界的神秘大门，更为我们揭示了一个充满机遇与挑战的未来。

其实，我在上大学之前，对计算机专业的课程知之甚少，但是，计算科学导论带我了解了专业的基础知识和主线，这给了我一些安定感，同时了解也带来了跟多的兴趣点，可以让我们自行探索计算机这个广阔的天地。但是，课程中，最有意思的其实是老师对当下计算机热点的讲解，这让我对目前的计算机行业充满了兴趣。

# 2 对计算科学导论这门课程的认识、体会

计算科学导论其实是一门入门性的课程，它为学生们总体上介绍了计算机中要学习的内容，等到大学毕业之后，你就会发现它可以作为一个大纲。它的目的是要我们了解计算学科的基本知识和学科方法，引领我们的兴趣，带我们了解计算领域的热点。

在课程的第一章中，我们学习了科学哲学的思想方法，因为我高中副科选了政治，所以对这一部分颇有感触，其中有关哲学和逻辑的部分甚至和高中所学十分相似。一般人来看哲学距离我们很远，认为那是从事人文社科的人所关注的，但是一般的科学思想方法对我们解决问题也具有指导作用，比如遇到一个在编程时，遇到一个过于复杂的问题怎么办？我们可以把握事物的主要矛盾，先忽略次要部分，解决主要部分。关于抽象这一词，因为在课上提到多次，我对它的印象也比较深刻，抽象实际上就是从个性中提取出共性，或是建立一种理想化模型，这也是我们处理问题的一种方式。所以计算科学导论不仅仅是为我们讲解了知识，更为我们提供了解决问题的思想方法。

## 2.1 计算科学与计算机科学有何区别？

首先，科学是关于自然、社会和思维的发展与变化规律的知识体系。科学用来认识世界，技术和工程用来改造世界。 计算科学系统研究描述和变换信息的算法过程，包括其理论、分析、设计、效率分析、实 现和应用。计算学科来源于对数理逻辑、计算模型、算法理论、自动计算机器的研究，形成 于1930年代后期。 而技术以发明为核心，解决“有”“没有”的问题，强调可行性 。工程则是集成科学和技术解决实际问题。

再者，计算科学实际上涵盖了计算机工程和计算机科学，由于最初各高等学校产生计算机科学系的学科生长点所在的院系不同，到了70和80年代，众多大学的计算机科学系逐步分化为计算机科学系和计算机工程系等两大阵营，以后又出现了一些变形，如计算机科学与工程系、数学与计算机科学系、电子与计算机工程系、管理与信息科学系等。由于对计算机科学领域开展工作的侧重点不同，对学科的认识不同，产生了对学科发展道路与人才培养方面的不同认识和诸多争议。为了解决学术界和教育界对计算机科学的认识和分歧，为学科人才培养探索一个科学的发展模式，从 1985 年起，美国 ACM和IEEE/CS 两个学会联手组成攻关小组，试图解决这个问题。联合攻关小组在走访了全美几乎所有知名计算机科学家的基础上，经过近四年的工作，于 1989年提交并发表了一份《计算作为一门学科》的研究报告。在这个报告中，他们认为计算机科学与计算机工程没有区别，本质上是相同的，并有关的科技议使用计算科学一词来涵盖计算这一领域内的所有工作【1】。

## 2.2 AI会不会产生意识？

清华大学北京信息科学与技术国家研究中心助理研究员郭雨晨说：“我们说人有自主意识，是因为人知道自己在干什么。机器则不一样，你对它输入内容，它只是依照程序设定进行反馈。”中国社会科学院科学技术哲学研究室主任段伟文认为，一般意义上，人的自我意识是指对自我具备觉知，但如何认识和理解人类意识更多还是一个哲学问题而不是科学问题，这也是很难明确定义人工智能是否具备意识的原因。同时，脑科学也尚未研究清楚意识的起源【2】。

虽然AI（人工智能）目前展现出的能力可以执行复杂的任务，如图像识别、语音合成、自动驾驶等，但就其本质而言，AI并不具备真正意义上的“意识”或“认知”。AI系统，尤其是深度学习模型，通过大量数据训练来识别模式、作出决策或生成内容。这些模型基于神经网络结构，能模仿大脑某些方面的功能，但与真正的生物大脑相比，缺乏情感、自我意识、意志和主观体验。意识的第一层次是潜意识，这是大多数人类智慧所处的位置，而潜意识则使我们能够做出决定的能力，如发现一张脸。然而，意识中更重要的组成部分是能够保持广泛的思想。大脑的其他部分可以访问这些想法，并且这种能力可以同时访问思想这使得长期计划成为可能，并为我们提供了意识的定性感受。因此，意识被表现为一种涉及高阶思维过程和对自己思想过程的认识的元认知。我们保持广泛思想的能力拓宽了我们的世界时间窗口，并在我们需要在一段时间内保持感官信息时发挥作用。而且，意识通常涉及到自我感知、情绪、意图性和目的性，这些都是当前AI无法实现的。即使AI可以在特定任务上超越人类，比如下棋或玩电子游戏，但它并不能理解这些活动的意义，也没有自身的目标或欲望驱动。而且，AI缺乏情感，AI没有感受快乐、悲伤或其他情感的能力。AI也无法理解上下文，虽然AI在理解语言方面取得了巨大进步，但它仍难以理解语言的微妙含义和文化背景。最后，AI可以生成看似原创的作品，如音乐、诗歌，但这不是源于内在动机，而是基于算法的输出。

随着AI研究的进展，有人提出强人工智能（AGI）的概念，即具有普遍智能的人工智能体，理论上能够理解、学习几乎所有人脑能做的事情。但从广义上讲，人工智能是先进的统计学和应用数学，它利用计算能力的新进展和可用数据的爆炸式增长，为计算机提供新的推理、识别和选择能力。比如机器学习，机器学习（Machine learning）是 AI 最有希望的子集，它是一个旨在教导计算机从大量示例（或“大数据”）中学习并执行任务而无需明确编程的任务的手段。最基本的是，机器学习使用算法来解析数据，从中归纳学习，然后对世界上的某些事情做出决定或预测。不是使用特定指令编码来完成特定任务，而是使用大量数据和算法“训练”机器，使其能够学习如何执行任务或预测结果。又比如深度学习，深度学习(Deep learning)是机器学习中最成功的方法，它基于类似大脑的“神经网络”。在深度学习网格中，电脑拥有“神经元”，它们具有离散的层和与其他“神经元”的连接 ，就像我们自己大脑中的神经元一样。每层神经元都会选择一个特定的特征来学习，例如猫的颜色、外形、行为，而这种层次结构可以让人工智能深入学习如何在图片和视频中辨认一只猫。然而，其仍然不能达到人脑的水平，而且达成这一目标面临着巨大的技术挑战和哲学争议，目前仍是科幻小说和未来学的领域。

所以，我认为AI不会产生意识。

* **2.3 那些专业建议2025届高考考生选择**

毫无疑问，如今是信息时代，AI将会深刻地影响我们的生活。以大模型和生成式人工智能（AI）为主导的技术创新持续取得突破，已成为推动科技和产业变革的关键力量，对经济社会发展和人类文明进步产生着广泛而深远的影响【3】。对于即将步入大学殿堂的高考生来说，选择专业是一项重要的决定，计算机科学无疑是最具吸引力的选择之一。

1. 就业市场广阔：信息技术行业正处于高速发展阶段，对计算机人才的需求量极大。无论是在大型科技公司，还是初创企业，甚至是非科技行业，都有计算机专家的一席之地。毕业后，你可以从事软件工程师、数据分析师、产品经理等多种角色，就业前景十分光明。

2. 薪资待遇优渥：计算机科学毕业生往往享有较高的起薪水平，随着工作经验的积累和个人技能的提升，薪酬涨幅也非常可观。在全球范围内，IT行业的薪资水平都处于领先地位。

3. 创新能力培养：计算机科学教育强调逻辑思维、问题解决和创新精神，这些技能在个人职业生涯中极为宝贵。它教会你如何设计、构建和优化系统，促进科技创新。

4. 适应性强：掌握计算机科学的理论和技术，意味着你有能力在不断变化的技术环境中迅速适应。这不仅局限于某个具体岗位，还可以跨界至其他行业，比如医疗健康、金融服务、教育娱乐等，成为行业的革新者。

5.全球化视野：计算机科学没有国界限制，精通一门外语加上计算机技能，让你能在国际舞台上大展拳脚。跨国公司、远程工作模式更加普及，你的才能将被全世界认可。

6. 满足个人兴趣：如果你热爱技术和解决难题，计算机科学将带给你极大的乐趣和满足感。编程、游戏开发、网络安全等，每一项技能都能找到对应的兴趣点。

# 3 进一步的思考

近年来,随着创新创业的兴起,决策树ID3算法在创新创业分析模型中的应用越来越广泛。决策树ID3算法是一种有效的数据挖掘技术,能够从大量数据中挖掘出有用的信息,为创新创业提供科学依据。

* **3.1 ID3算法的缺点**

首先，传统的ID3算法具有明显的收敛特征,限制了其在创新创业分析模型中的应用。通过实验，当记录数量为5200条时,优化算法的错误率为70 %,而经典算法的错误概率为95 %。当记录数为7200时,优化后的计算方法的误差率为130 %,而经典的ID3计算方法的误差率为170 %。随着记录数量的增加,优化算法和经典算法的错误率都会增加。随着记录数量的增加,优化后的ID3算法的错误率增加幅度小于经典算法【4】。

而且，ID3有偏好属性，ID3倾向于选择具有多个分支点的特征，因为这样更容易达到信息增益最大化。但这并不意味着这样的划分就是最优的，有时候少数几个分支就足以区分所有类别的情况下，过多的分支会导致过拟合。

其次，ID3没有缺失值和连续值处理，原始的ID3算法没有内置机制来处理带有缺失值的数据。虽然可以通过插补或者其他预处理手段弥补这一点，但这增加了额外的复杂性和误差可能性。因为ID3不直接支持连续型特征，需要预先将其离散化才能使用，这也可能导致信息丢失或引入偏差。然后，ID3采用贪婪策略构建决策树，每次只基于当前的最佳分割点进行分裂，而不是考虑整个决策树的全局最优结构。这可能导致局部最优而非全局最优的结果。而且，当训练集中各类别样本数量极不均衡时，ID3可能会偏向多数类别，忽略少数类别的信息。如果没有适当的剪枝措施，ID3容易生成过于复杂的树，过分贴合训练数据，导致泛化能力差。MD原则就是一种剪枝方式。MDL 原则是一种强大的归纳推断方法，是统计建模、模式识别和机器学习的基础。 该原则认为，基于已知的少量数据，最好的解释是能够最大程度压缩数据的解释。 在 MDL 中，如果分类的相应子集实例以及子树中每条路径的描述长度加上子树的描述长度大于当前节点中整个子集实例的描述长度，则子树的子树将被视为不可靠并被修剪【5】。

* **3.2 决策树是基于哪些算法实现的，又是怎样实现的**

最经典的决策树算法是ID3，基于ID3又优化出C4.5和CART等算法。

要解释决策树是如何运行的，需要引入熵这个概念。熵是描述事物混乱程度的一个度量单位。当某个特征的熵值越小，代表它的混乱程度越低，对结果的影响越关键。因此按照熵值从小到大的顺序组合就能得到决策树。这里我们以CART算法为例，在CART算法中，度量指标是基尼系数。基尼系数的计算公式是一减去所有概率的平方。当某个特征的基尼系数越靠近0或者1，则表示特征越关键。我们来演算一遍，假设这里有8次历史战斗数据的特征和结果，数据中的特征分别有玩家水平、单位数量和单位属性，而结果分别是游戏的胜负。第一轮需要从三个特征中做选择，把数据的特征和结果带入公式中，得出玩家水平高的基尼系数为0，玩家水平低的基尼系数为0.375，然后加权平均得到玩家水平的基尼系数为0.1875。同样的方式，我们得到单位数量和单位属性的基尼系数是0.4375。显然，玩家水平更关键。因此作为父节点。第二轮需要从两个特征中做选择，我们保留玩家水平低的样本数据继续计算，得到单位数量和单位属性的基尼系数是0.25，选择其中一个作为父节点。至此，我们成功得到了决策树。

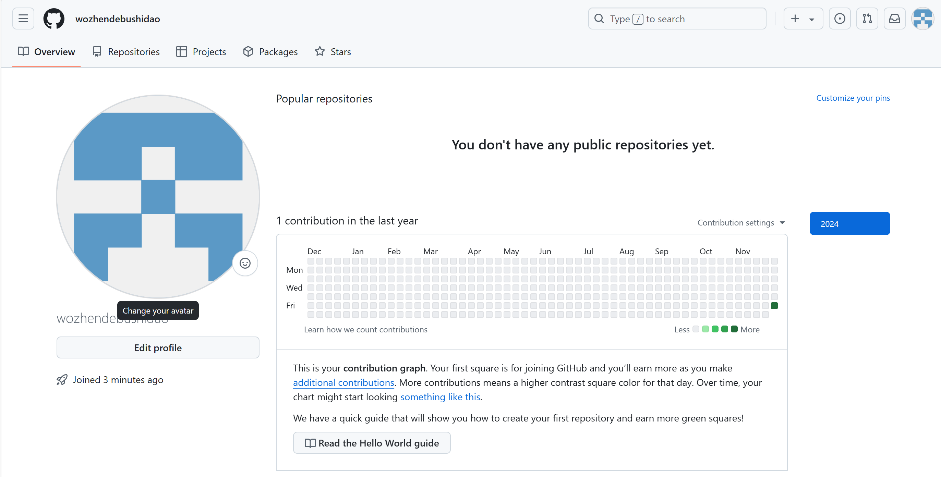
# 4 总结

在《计算科学导论》课程的学习旅程中，我们不仅汲取了丰富的知识，更体验了一次深刻的学术洗礼。自开课以来，我们循序渐进了计算科学的宏伟画卷——从最基础的计算理论开始，逐步深入了解算法设计、数据结构、计算机架构，直至高级领域的机器学习和人工智能。除了硬核知识，我们还获得了诸多软实力的提升。团队协作、批判性思维、持续学习的习惯，都在这个过程中得到了锻炼。更重要的是，我们学会了如何面对不确定性和失败，保持好奇心与探索欲。面向未来，我将带着这份扎实的基础和满腔热情，勇敢迈向下一个里程碑。无论是进一步深化专业知识，或是拓展跨学科视野，计算科学都将是我们强有力的武器。在信息爆炸的时代，新技术层出不穷，但万变不离其宗，扎实的理论根基和持续学习的态度将永远是我们的导航灯塔。让我们携手共进，探索科技的无限可能，塑造一个更加美好的数字未来！

---

# 5 附录

## Github

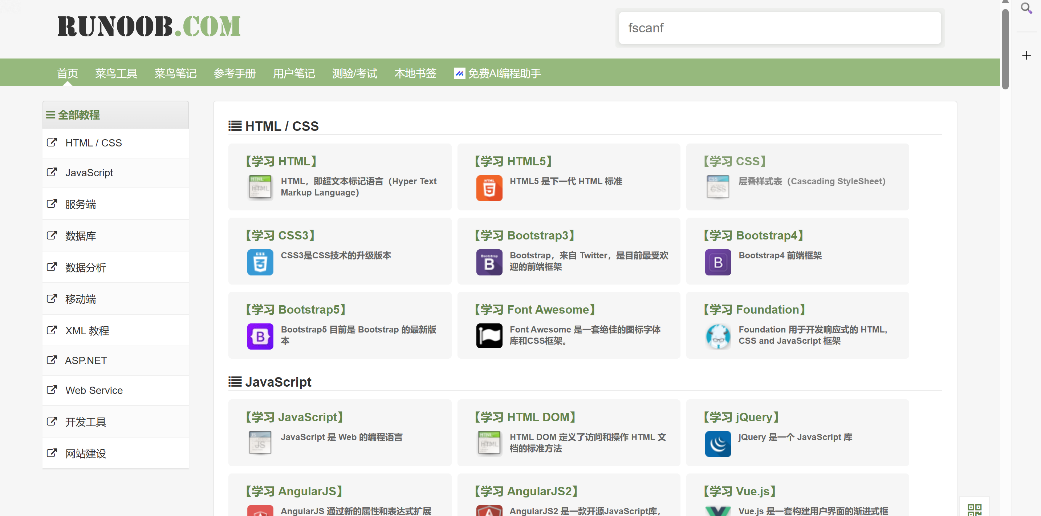
[wozhendebushidao (github.com)](https://github.com/wozhendebushidao)

## 力扣

[Nervous Khorana25T - 力扣（LeetCode）](https://leetcode.cn/u/nervous-khorana25t/)

## 菜鸟教程

[菜鸟教程 - 学的不仅是技术，更是梦想！ (runoob.com)](https://www.runoob.com/)



## CSDN

[2401\_86110488-CSDN博客](https://blog.csdn.net/2401_86110488?spm=1000.2115.3001.5343)



## 稀土掘金

[用户00666155543 的个人主页 - 动态 - 掘金 (juejin.cn)](https://juejin.cn/user/1616716656814864)



## 语雀

[wozhendebuzhidaoa · 语雀 (yuque.com)](https://www.yuque.com/wozhendebuzhidaoa-krn5d)



## 哔哩哔哩



## 观察者



# 参考文献

1. 【2】赵致琢，《计算科学导论》
2. 于雪，魏雨虹，《人工智能可能有自主意识了吗》
3. 王震，《计算机行业加速融合 业界：人工智能电脑产业“代际”不断升级》
4. Linchen Mei, Application and Optimization of Decision Tree ID3 Algorithm in Innovation and Entrepreneurship Analysis Model
5. Rana Ajay, A review of popular decision tree algorithms in data mining