

浅谈本科生科研能力培养

李武军

上海交通大学

关键词：本科教育 科研能力

我的研究组里今年毕业的本科生中，三位学生发表了论文，其中两位学生以第一作者身份在中国计算机学会 (CCF) A 类¹会议上发表多篇学术论文。对于本科生，能够在科研上取得这样的成果，是非常难得的。据了解，我所在学校其他老师的研究组以及国内其他大学也有本科生发表过国际顶级论文。因此，对部分本科生实行适度的科研能力培养，是有可能取得一定成效的。

当《中国计算机学会通讯》(CCCF) 专栏编委向我约稿，让我谈谈本科生科研能力培养方面的经验时，我曾经犹豫过，主要原因有两点：第一，只凭借今年 (2013 年) 一届学生的成绩，是否谈经验有点早；第二，按照“惯例”，本科生应当以基础教育为主，过早地涉及科研是否有揠苗助长之嫌？

经过认真考虑后，我觉得在中国高校²教育发展的现阶段，把我的经验写出来与国内同行一

起探讨，还是有一定的意义的。另外，如果本文能引发人们对本科教育的讨论，也会对本科教育起到积极作用。需要注意的是，我的研究方向是机器学习以及相关的应用领域，因此，本文的很多观点可能只适合于计算机应用方向。

本科生参加科研的意义

近年来，中国高校招收教师的门槛不断提高，很多讲师都是由刚走出校门的博士担任。这些老师在刚参加工作的几年里思维非常活跃，有许多好的科研想法和点子。但是，从“资格”来讲，当前很多高校的讲师还不能“指导”研究生。如果本科生能够积极地参加科研，会给这些老师提供非常大的帮助。事实上，两年前我也是因为刚入职，招收的研究生还没有到位，才尝试指导本科生。两年间，我指导的本科生一共发表了 6 篇论文，其中有

4 篇发表在 CCF A 类会议上。很多学生在研究生阶段会选择留在本科阶段所在的研究组继续深造，因此对于具有指导研究生资格的老师，指导本科生参加科研也会给其研究生招生带来很大的好处。另外，目前不少学校的硕士生培养周期只有两年半，学生一般在第一年主修专业课，最后半年找工作，只剩下一年时间用来科研。因此，大部分硕士生科研还没入门时就即将毕业。如果能从本科阶段就开始进行科研训练，科研工作就会深入得多，从而大大增加取得创新性成果的机会。

对于学生来说，参加科研能让他们亲身体会到基础知识的重要性。例如，我的研究组中经常有学生向我反映，在大一、大二学线性代数和概率统计课程的时候，一直怀疑这些知识在计算机学科中是否有用。但经过参加机器学习相关的科研训练后，他们发现这些数学知识能够用来解决实际问题。

¹ A类指国际上少数的顶级刊物和会议。具体信息请参见《中国计算机学会推荐国际学术会议和期刊目录》(<http://www.ccf.org.cn/sites/ccf/paiming.jsp>)。

² 本文的中国高校特指中国大陆高校。

因此,科研训练能提高他们学习基础知识的积极性,让他们从被动接受知识转变为主动学习。对于想在国内继续读研究生的学生来说,本科阶段就参加科研可以大大增加研究生阶段取得创新性成果的机会;对于计划到海外继续深造的学生来说,如果本科阶段有论文发表或者有具体的研究经历,对申请就读海外名校也很有帮助。今年我的研究组中有一位本科生因为在国际人工智能协会会员会议(Association for the Advancement of Artificial Intelligence, AAAI)、国际信息检索大会(ACM Special Interest Group on Information Retrieval, SIGIR)、神经信息处理系统会议(Neural Information Processing Systems, NIPS)上发表论文,获得了美国斯坦福大学的博士生全额奖学金。另一位学生在国际人工智能联合会议(International Joint Conference on Artificial Intelligence, IJCAI)上发表两篇论文,被香港科技大学录取,并获得了每年25万港币的香港政府博士研究生奖学金(HKPFBS)。

当前,中国有些机构的博士生培养存在一些问题:由于博士生生源差,导致博士生培养质量差,进而导致博士生就业差,反过来导致生源更差。而大部分本科生只看到了这个问题的表象,以为读博士没有出路,因此本科毕业就选择逃离科研。但实际情况是,我国的科研机构和企业非常需要具有创新潜质和真正研究能力的科研人员。优秀的博士总是能够找到用武之地,并且在很多场合能发挥关键作用。

我们指导本科生参加科研,可以让他们鉴别什么是好的科研,什么是不好的科研。他们会发现,研究做得好的博士大部分都有不错的前景。这样,我们就能引导部分优秀的本科生选择科研作为自己的职业,从而提高博士生的生源质量,形成良性循环。

什么样的本科生适合参加科研?

原则上,所有本科生都能参加适度的科研工作。本文讨论的对象特指能够投入很大精力参加科研并取得较大进展、而且有可能在本科期间发表高水平论文的学生,他们应具备以下特点。

首先,要对科研感兴趣。俗话说:兴趣是最好的导师。科研工作需要静下心来阅读大量的书籍和文献,需要长期对某个问题深入地分析和思考,进行大量的实验和求证,并在论文写作过程中反复地推敲和修改。在这个过程中,对科研的兴趣和激情是让我们坚持并最终取得突破的最重要因素。因此,我考察学生的一个重要标准就是看他能否静下心来连续十几个小时甚至几天专注地阅读文献、思考问题、观察实验过程。另外,一定要对自己的研究方向感兴趣。对于刚加入研究组的学生,我会先向他们介绍实验室涉及的研究方向,并给他们一定的时间阅读最新的文献资料,然后让他们结合自己的特长选择感兴趣的研究方向。

其次,要勤奋踏实。有一个例

子:有一次临近论文截稿日期,我们的论文还没有得出很好的实验结果,跟我合作的本科生连续几天熬夜做实验,最终从实验结果中发现规律并找到一种改进方法,得出很好结果。这篇论文后来顺利地被CCF A类会议录用。从这个例子可以看出这个学生具有勤奋踏实的特质。相反,有些不踏实也不勤奋的学生,即使导师交给一个简单的实验任务,也迟迟做不出结果,还会找各种理由来搪塞。

最后,要有扎实的基础。我一直认为本科教育应当以基础教育为主。我招收本科生加入研究组的标准是“成绩优秀并且学有余力”。如果学生的成绩不是特别好,我一般都会拒绝,并建议他以课程学习为主。虽然课程成绩和科研能力不一定存在必然的联系,但我发现研究进展不错的往往是那些课程成绩优秀的本科生。在科研过程中,如果有本科生因为参加科研而影响了课程成绩,我也会建议学生停止科研。

本科生参加科研的时间点很重要。我建议从大二结束后的暑假开始,一直延续到本科毕业。一方面是因为这个时候数学、英语以及部分计算机核心基础课都学完了,学生具备了一定的专业素养,另一个原因是大三的本科生还不用花太多的时间来考虑毕业后的打算,能够心无旁骛地学习和钻研。相反,大四的本科生因为需要考虑出国、读研、工作等问题,能够静下心来钻研的可能性较小。有很多老师对本科生的科研培养是从大

四的毕业设计开始的,我认为这并不是一个很好的时机。

导师该如何培养本科生?

中国科学院计算技术研究所的山世光研究员在文献 [1] 中系统地阐述了科研的流程,并完整地介绍了怎样培养研究生从科研入门一步步成长为具有独立科研能力的研究人员。文献 [1] 中关于指导低年级研究生的很多经验也适用于本科生的指导。然而,本科生科研具有其独特性,本文将从目标、方法和管理三方面来重点阐述这些独特的地方。

目标

在文献 [1] 中,科研流程被划分为“问题—思路—算法—实验”四个阶段。如果再加上论文写作,可以将整个科研流程划分为“问题—思路—算法—实验—写作”五个阶段。“问题”指的是通过调研相关领域最前沿的理论和方法,提出具有科学价值的学术问题;“思路”指的是解决问题的具体方案,必要时需要用形式化的手段进行描述,比如机器学习里很多的解决方案可以通过优化一个带约束的目标函数来描述;“算法”指的是将解决方案转变成计算机程序的实现过程,有些算法需要进行数学推导;“实验”指的是通过在数据集上运行算法,并用科学的指标来评测算法的性能,通常需要跟已有算法进行客观、详细的对比;“写

作”指的是将上述流程整理成学术论文。

博士生培养的目标是培养具有独立科研能力的研究人员,需要对上述五个阶段的能力进行全方位培养。但本科生科研能力培养的目标以培养科研兴趣为主,重点定位在“算法”和“实验”阶段的培养,附加一些“思路”和“写作”阶段的培养,弱化“问题”阶段的培养。

方法

要实现上述本科生科研能力培养的目标,尤其是要保证本科生所设计算法的正确性和实验结果的可靠性,需要一套严谨的培养和训练方法。

首先是基础知识的训练。基础知识指的是导师的科研方向所需要的基础理论和方法。例如,很多学校在本科阶段不会开设机器学习相关课程。如果学生选择机器学习作为研究方向,在正式开展科研之前,必须较完整地学过机器学习相关的基础知识。我的经验是充分利用网络公开课,即要求刚加入研究组的学生利用大二暑假至少学习两三门与研究方向相关的网络公开课程。另外,我也会要求学生在大二暑假读完我推荐的两三本与研究方向相关的教材。在大三刚开学时,我会对他们的基础知识学习情况进行了了解和考核。如果能够完成我布置的这些任务,就说明该学生确实对这个研究方向感兴趣,而且具有一定的钻研精神,接下来我就会帮他选择某个研究方向并进行下一步的科研训练。相反,如果某个学生完成

的情况不理想,说明他要么对科研不感兴趣,要么学习不踏实,我一般会建议他退出研究组。

其次是实验技能的训练。实验技能指完成科研实验所需要的编程技巧和实验手段,通常也包括对某些实验平台和工具包的熟练使用。我会根据学生选择的研究方向推荐两三篇有公开代码和数据的重要论文,让学生读懂论文的细节,然后利用论文公开的代码和数据,重复论文中的实验结果,同时熟悉代码和实验的评测流程。经过训练后,学生就能掌握怎样将论文的思路转化成算法,并最终通过程序来验证思路有效性的基本技巧。我还会选择一篇最新的、但是网上没有公开代码的论文,让学生重现论文中的结果。这篇论文里的方法一般是我想改进的,即我的思路来源。这个过程实际上训练学生将学到的基础知识和实验技能灵活运用能力。如果学生能够重现论文中的结果,可以判断该学生已经具备进行实验所需的基本技能。

学生有了扎实的基础知识和熟练的实验技能后,导师就可以找一个科研问题对学生进行科研实战训练。导师给本科生的科研问题必须明确、具体,而且通常情况下导师自己已经有初步的解决思路。这对导师的要求较高,导师必须了解这个领域的前沿进展,并进行过详细的文献调研,给出的问题和解决思路既要有前沿性又要有可行性。我的经验是把本科生的第一个科研题目限定到 3~6 篇参考文献中:其中 1~2 篇来自最近的顶级

会议,包括需要改进的方法,即思路的来源;1~2篇是解决方案和算法设计需要参考的文章;1~2篇实验评测需要参考的文章。在导师的指导下,学生只要把这几篇文章琢磨透,就能较快地上手,完成第一个科研题目的算法和实验阶段。相反,如果导师只给定一个泛泛的题目,让本科生自己去从大量的文献中寻找问题和思路,往往造成研究项目长期没有进展,甚至导致学生丧失对科研的信心。

如果算法和实验取得了好的结果或者有了新的发现,我们需要以论文的形式来发表研究成果,与同行分享。正如文献[1]中所说的,语言并不是写(英文)论文的最大障碍,论文的整体逻辑才是写作的关键。由于知识面有限,本科生写的论文通常不能恰当地表达出自己的工作在整个知识体系中的地位和贡献,这就需要导师花很大的精力来指导和修改。我的经验是,即使最终还是需要导师进行大幅度修改甚至重写,导师也尽可能地宏观上给一些概要性的指导意见,让学生自己多修改几遍,逐步培养和提高学生的写作能力。

大部分学生在本科阶段只能完成一个科研题目或者一篇学术论文,也有少部分优秀的学生可以完成多个科研题目和多篇学术论文。对于后者,我们可以在“思路”阶段进行一定的培养。当学生完成第一个科研题目后,导师可以训练学生去阅读更多的文献,帮助他分析文献中已有方法的优缺点,引导他们提出自己的思路和解决方案。

如果学生能够成功地完成一个科研题目或者一篇学术论文,他就能从中找到成就感,并对科研产生兴趣。我们本科生科研能力培养的目标也就达到了。至于独立科研能力,可以留到博士阶段再培养。

管理

我对研究组里本科生的管理原则是:采用与研究生完全一样的管理制度。每周都有组会,研究生和本科生都必须参加,学生轮流做学术报告,讨论前沿的学术论文或者讲解经典的教材章节。同时,我也会安排单独见面的机会,以了解学生的科研进展,并解答科研中碰到的问题。学生有问题可以随时发邮件给我,我都会尽快回复,如果有必要,我会安排单独见面解答。在整个过程中,本科生跟研究生没有区别。因为每年都有特别优秀的本科生加入我的研究组,有时候本科生反而在组会的讨论中起主导作用。

需要强调的是,有些本科生喜欢同时参加多个老师的研究组,或者频繁更换研究组,这对科研能力培养是不利的。因为本科生大部分精力需要花在课程学习上,参加多个研究组或者频繁更换研究组会造成研究不够深入,很难取得突破性进展。所以,我对学生的要求是从大二暑假开始一直到本科毕业,只能在我的研究组开展科研工作。

结束语

本文的几个要点如下:

1. 本科生科研能力培养以培养科研兴趣为主;
2. 本科生科研能力培养只适合于成绩优秀并且学有余力的学生,不能影响到本科课程的学习;
3. 建议本科生从大二暑假开始参加科研,只加入一个自己最感兴趣的研究组,并且能够坚持到本科毕业;
4. 导师布置的科研问题必须明确具体,既具有前沿性又具有可行性;
5. 科研方向的基础知识很重要,导师可以指导学生有效利用网络公开课来强化基础知识的学习;
6. 导师让学生实现几个前沿的方法并重现论文中的结果可以训练学生的实验技能;
7. 建议导师采用跟研究生完全一样的管理制度。

本文仅是我个人粗浅的认识,希望能够起到抛砖引玉的作用,与国内同行共同探讨本科生科研能力的培养。■



李武军

CCF会员。上海交通大学副教授。主要研究方向为人工智能、机器学习、数据挖掘等。
liwujun@cs.sjtu.edu.cn

参考文献

- [1] 山世光. 浅谈科研流程及其中的师生合作. 中国计算机学会通讯, 2013; 9(3):38~43.