

104 | 如何快速学习国际顶级学术会议的内容？

2018-06-01 洪亮劼

AI技术内参

[进入课程 >](#)



讲述：初明明

时长 08:09 大小 3.74M



我们在专栏中已经多次分享了计算机科学（特别是机器学习、数据挖掘和人工智能）领域顶级学术会议的会议论文，向你介绍了这些会议中一些最新的学术和研究成果，希望这些内容能够起到抛砖引玉的作用。

最近今年，随着机器学习和人工智能的逐渐火热，越来越多的研究人员投入到这些领域中，这些顶级会议的投稿论文数和收录稿件数都成倍增长，不少会议会有几百篇论文被收录。这就给希望学习这些论文的人，不仅是初学者，甚至是长期在这些领域工作的研究人员和实践者，都带来了不小的挑战。

现实情况下，尝试去看完一个会议的所有论文是不现实的，从某种程度上看也是不必要的。那么，怎样才能从这些会议中快速地、有效率地吸收到更多的内容呢？想要事半功倍地学习顶级学术会议的论文，究竟有没有什么方法可以帮助我们实现这个目的呢？

今天我就结合自己的经验，和你分享如何快速学习顶级会议的内容。

顶级会议的基本内容架构

在学习任何顶级会议的论文内容之前，我们其实最需要了解的就是这些会议本身有什么规律，或者说会议的内容安排上是不是遵循了一定的逻辑架构。这里，我们首先来看一看这些顶级会议在内容安排上的规律。

毫无疑问，所有顶级会议的核心内容是收录的论文。这一点，我已经反复提及了。有一个细节可能是初学者容易忽视的，那就是很多会议都对收录的论文进行了区分，用于区分出最优秀的论文和一般的论文。

这种区分一般有两种形式。第一种形式是论文在会议上的表现形式有区别。最优秀的少量论文是时长 20 ~ 30 分钟的**演讲报告**，而其他大量的论文是时间很短的报告，或者有的会议干脆就没有这些论文报告时间。第二种形式是这些论文在会议论文集中发表的长度有区别，这也就有了所谓**长论文**和**短论文**的区分。通常情况下，长论文是 8 ~ 11 页，而短论文则是 2 ~ 6 页不等。

除了收录的论文以外，顶级会议往往还有很多精彩的内容。你可以关注的第二个有价值的内容形式是“**讲座**” (Tutorial)。“讲座”的形式常常有半天的、全天的，以及超过一天由几个研究者针对某一个特定主题的集中式分享和讨论。一般来说，“讲座”往往是一些最新的热点问题，但同时有了一整套阶段性的成果，这时候，在这个领域有突出贡献的研究者，就利用“讲座”的模式把这些内容给总结出来进行分享。

第三种你可以关注的内容形式就是“**研讨班**”。“研讨班”是针对某一个特定的近期热门的主题，由一些学者或者工程师组织的小型会议。和“讲座”一样，“研讨班”往往也是半天到一天不等。“研讨班”被认为是小型会议，因此有时候也还会收录一些论文。然而，这些论文的数量和质量都有很大的差别。

从组织的角度来讲，“研讨班”的目的主要是吸引更多人对某个主题的关注，以及给目前针对这个主题工作的研究者们提供一个交流的场所和平台。从这个角度来讲，很多“研讨班”分享的内容可能都是“还在进行的工作” (Work In Progress)。所以，“研讨班”所关注的热点往往比“讲座”还要激进。近年来，机器学习和人工智能的热点越来越多，有的会议有多达十几个甚至二十几个“研讨班”在大会期间举行。

最后，还有一个可以关注的则是很多会议和研讨班中都有的“**邀请演讲**”（Invited Talk）或者叫作“**主题演讲**”（Keynote Talk）。如果是亲自参加这些会议或者有视频录像，那么这些演讲往往还是值得一听的。从会议组织的角度来说，这些演讲往往是邀请在本领域有重大贡献的学者或者是相关领域的权威学者。因此，通过这些演讲，我们可以了解到这些领域的重要学者都在针对什么问题进行思考，以及他们眼中这些领域的发展方向是怎样的。

上面介绍了几个顶级会议值得关注的点，需要注意的是，并非所有会议都有“研讨班”或者“讲座”。

如何精选内容

了解了这些顶级会议内容的一个基本架构以后，我们现在来聊一聊如何从这些不同形式的内容中精选出你可以学习的信息。

首先，我们来说一说论文。刚才提到了，论文可能分为长论文、短论文以及报告论文和非报告论文。那么，有一些比较粗浅的方法可以帮助你筛选。例如，针对长论文和短论文，可以首先关注长论文。对于报告论文和非报告论文，可以先关注报告论文。利用这种简单的方法，我们已经过滤掉了至少一半以上的论文数量。这里的一个基本假设是，会议安排这些论文的区别其实就是一个最基本的质量控制。

另外，很多会议把所有的论文分为不同的“组”（Track）。每一个组其实有一个主题，例如搜索、推荐、用户建模等等。这其实也是帮助你通过论文内容的类别进行区分。所以，如果你已经知道自己感兴趣的内容，那就可以直奔主题。再庞大的会议，针对某一个主题的论文数目往往也是非常有限的。

还有一种便捷的方法就是“跟踪”某一些学者。当然，这需要你对这些领域有哪些活跃的学者、有哪些权威的学者要先有一个了解。你心里有这么一个清单，那么，只关注这些你清单上的学者在这个会议上发表的论文就是一个非常简洁的方法。

举一反三，还有一些和追踪学者类似的办法，比如你可以追踪某几个研究组或者公司。当然，这也需要你对这些领域有一定的了解。

对于论文来说，另外一个精选内容就是获奖论文。虽然对于获奖论文在长期来看是不是最具有学术价值一直存在争议，但是作为快速学习和扩宽视野的方法，浏览这些获奖论文以及获奖作者的一些附带研究，可以说是了解一个领域的另外一个捷径。

说完了论文之后，我们来看一看“讲座”。其实，这里也有一个窍门，那就是**如果你是一个初学者，我建议你不要一上来就去看一个会议的论文，而是从“讲座”入手。**

我们前面提到过，“讲座”是学者已经总结好的最近已经形成一定系统性的研究成果。这里，“总结好”和“系统性”这两个特点非常重要。这意味着“讲座”里面的内容往往不是一篇论文的介绍，而是多篇相对于一组主题论文的总结。这一组论文以及相关主题极有可能已经形成了一个稳定的内容块，而“讲座”的目的就是希望更多人认可、学习、普及这个内容块。有很多“讲座”的内容在几年后就成为了这些领域核心经典思想的一部分。因此，**学习“讲座”其实就是一个走捷径的过程。**

另一方面，因为“讲座”具有总结性质，所以肯定会提供很多具体论文的文献资料，这些资料可以作为进一步学习的“指南针”。正因为这个性质，对于“讲座”内容的把握你可以做到先从概念上和大思想上对一个主题进行把握，例如看你自己能不能简单地把这个“讲座”中的核心技术、核心思想复述出来。同时，你也可以来衡量一下到底需不需要对这个“讲座”具体指引的论文进行阅读。绝大多数的时候，我们对于“讲座”的把握在这个层次其实就够了。

最后，我们来说一说“研讨班”的内容。和“讲座”不同的是，“研讨班”的内容因为其性质，往往比较杂。我们刚才也提到了，“研讨班”重在为一线的研发人员提供一个交流的平台。因此，他们的交流是第一目的。而且，“研讨班”的很多方向和主题都比较新，因此对于初学者而言，需要甄别哪些内容值得细看哪些内容可以忽略。总体说来，这里面的内容都不推荐给初学者。当然，如果你对某一个领域已经追踪了一段时间，有了一定的了解，那么“研讨班”无疑是最前沿的讨论平台之一。

小结

今天我为你梳理了如何快速学习顶级会议的内容，希望这篇文章可以帮助你在海量的会议内容中找到自己感兴趣的信息。

一起来回顾下要点：第一，我们讲了顶级会议的内容架构；第二，我们简要地介绍了如何在这些内容中进行快速筛选。

最后，给你留一个思考题，拿到一篇 10 页的长论文，究竟该如何快速阅读呢？

欢迎你给我留言，和我一起讨论。

AI 技术内参

你的360度人工智能信息助理

洪亮劼

Etsy 数据科学主管
前雅虎研究院资深科学家



新版升级：点击「👤 请朋友读」，10位好友免费读，邀请订阅更有**现金**奖励。

© 版权归极客邦科技所有，未经许可不得传播售卖。页面已增加防盗追踪，如有侵权极客邦将依法追究其法律责任。

上一篇 103 | The Web 2018论文精读：如何从文本中提取高元关系？

下一篇 复盘 3 | 自然语言处理及文本处理核心技术模块

精选留言 (3)

写留言



sky

2018-06-02

1

先看摘要和introduction，看是否对该论文真的感兴趣，再看conclusion，看论文的结果讨论是否对自己有启发，如果感兴趣再详细精读，但是有个问题是，怎么样才算把一篇文章彻底读懂了呢，公式的每个细节都需要弄明白吗？

展开



Bo

2018-06-01

1

洪老师实力派，能不能说一下第一次写KDD之类的英文paper经验，该怎么写？框架，思路，组织语句等等，非常需要哈😊



米乐乐果

2018-06-01



非常受益，授人以渔😊

展开▼

拼课微信：171614366!