

## 126 | SIGIR 2018论文精读：如何利用对抗学习来增强排序模型的普适性？

2018-07-23 洪亮劫

AI技术内参

[进入课程 >](#)



讲述：初明明

时长 06:58 大小 3.20M



今天，我们继续来精读 SIGIR 2018（国际信息检索研究与发展大会）的论文，今天分享的是本次大会的最佳短论文，标题是《使用对抗学习实现神经排序模型的跨领域正则化》（[Cross Domain Regularization for Neural Ranking Models using Adversarial Learning](#)）。

非常有必要先简单介绍一下这篇文章的作者群，可以说这是一个“明星阵容”。

第一作者丹尼尔·科恩（Daniel Cohen）来自马萨诸塞大学阿默斯特分校（University of Massachusetts Amherst），是计算机科学系的博士生。2017 年曾经在微软研究院剑桥分部实习。这篇获奖论文就是实习项目的总结。在深度学习模型在信息检索的应用这个方向上，科恩已经发表了多篇论文。

第二作者巴斯卡·米特拉 (Bhaskar Mitra) 是微软研究院剑桥分部的主任级科学家。近些年来米特拉在信息检索领域很活跃，并且极力推动深度学习在这个领域的发展，他在这个领域发表了很多篇论文，在过去几年的很多学术会议上，也主持了多个关于深度学习和信息检索相结合的讲座。

第三作者卡特娜·霍夫曼 (Katja Hofmann) 也是来自微软研究院剑桥分部的科学家。霍夫曼在信息检索领域的主要贡献是研究在线排序学习。

论文的最后一位作者，布鲁斯·克罗夫特 (W. Bruce Croft) 是信息检索领域的学术权威，也是科恩的博士导师。他是美国 ACM 院士，还是信息检索领域最高学术荣誉奖杰拉德·索尔顿 (Gerard Salton) 奖的获得者。

## 论文的主要贡献

这篇论文主要涉及了这么两个重要概念的结合。第一个概念是“**跨领域**” (Cross Domain) 信息检索，第二个概念就是“**对抗学习**” (Adversarial Learning) 。

跨领域信息检索主要是指我们需要对一个以上的领域进行搜索。这里所说的领域主要是指不太相同，或者非常不同的文档集合。例如，如果我们要针对体育新闻、金融新闻等进行搜索，这里的“体育”和“金融”就是不同的领域。

跨领域信息检索的核心挑战是我们如何针对不同的领域都能进行有效搜索。比如，如果我们的排序算法本身或者其特性依赖于某个特定领域的信息，例如关于体育方面的搜索，需要依赖体育运动员的名字，那这种信息肯定会很少甚至完全不出现在另外一个领域。因此，**依靠某些领域特有的信息很难做到真正的跨领域信息检索。**

这篇文章的贡献是作者们认为，想要对跨领域的信息进行较好地检索，就**需要训练这样的排序模型：不容易被某一个特定的领域所主导，同时也尽量不偏好某一个特定领域的具体信息。**

如何实现这个目的呢？作者们使用了一种叫做“**对抗学习**”的技术。这也是这篇论文能够获奖的重要原因。

我在这里简单介绍一下对抗学习的原理。对抗学习最初的思想来自于利用深度产生模型解决计算机视觉中的一系列问题。最基本的对抗学习的模式主要是用于产生数据，而且是比较复杂的数据，例如图像。

**对抗学习有两个模块，一个模块叫产生器，一个模块叫判别器。**产生器的作用就是产生数据。判别器的作用是判断产生的数据是不是“真实数据”。产生器的最终目的是产生能够以假乱真的数据来扰乱判别器的判断能力。判别器的最终目的是不断提高判断能力从而能够分辨出数据的真假。

当然，最初的时候，产生器产生数据来源于随机噪声，因此判别器可以很容易地判断数据的真假。但是慢慢的，产生器产生的数据就会越来越接近真实数据，而判别器也很快在这个过程中得到优化，从而能够判别数据的真假。当然，这是一个动态的过程，最终，判别器和产生器的状态会稳定下来。

对抗学习这种思想最初被提出来的时候，主要是应用在计算机视觉领域中，用来产生以假乱真的图片。之后，这种技术被广泛应用到人工智能的各个领域。

这篇论文最大的一个贡献，就是**利用了对抗学习的理念来增强学习到的排序模型的普适性**，从而尽量避免学习到仅仅对一个领域有用的信息。

## 论文的核心方法

具体来说，这篇文章提出了这样一种方法。首先，我们有两套模型，一套是用于学习查询关键词和文档之间的相关关系的；一套是对抗学习的模型。然后，这两套模型的首要任务是更加精准地针对相关的文档和不相关的文档进行建模。这是整个框架里最主要的目标函数。

文章提出框架中新的模块是**利用对抗学习来分别产生相关的和不相关的文档**。具体来说，某一种类型的文档就像我们刚才提到的图片一样，我们希望能够利用产生器来进行产生这类数据。当然，我们依然需要判别器来引导产生器的工作。

在这篇文章中，相关的概念主要是看一个文档是不是某一个领域的。也就是说，我们希望对抗学习能够帮助识别某一个文档是不是来自于一个领域。当对抗学习模型被训练好的时候，对于查询关键词和文档的相关模型，我们会利用一种叫做“**梯度反转**”的技术，强行偏离模型希望去拟合某一个领域的倾向。

从网络结构上看文章提出的模型，查询关键词和文档都需要经过**卷积层、提取层等变换**，然后进行俗称的“**哈达马积**”（Hadamard product），其实就是**对应项乘积**。这样，文档和查询关键词所提取出来的隐含特征就结合在一起。这个结果之后又经过一系列**稠密层的变换**，最终预测一个相关度的标签。对于对抗学习模型来说，对抗中的判别器从刚才所说的架构中提取一些中间层作为输入，然后判断这个文档是不是出现在某个领域中。

## 实验结果

在一个雅虎的搜索数据集以及另外两个数据集上，作者们对论文所提出的模型进行了实验。实验主要是看如果我们在某一个领域上训练出的模型，会不会在另外一个领域上表现优异。

一个不令人意外的结果是，如果我们在全部领域上进行训练数据，自然在所有的领域上效果都不错。当然，文章展示了，如果利用文章提出的方法，针对某一个领域，比如运动类文档，在训练的时候完全移除所有的文档，在测试集上依然有不错的表现。实验的结果比不进行对抗训练的效果要好 5% 以上。

## 小结

今天我为你讲了今年 SIGIR 上的最佳短论文。

一起来回顾下要点：第一，这篇论文主要涉及了两个概念，分别是跨领域信息检索和对抗学习，我们详细介绍了这篇文章的主要贡献，就是利用对抗学习的理念来增强所学排序模型的普适性；第二，我们简要地介绍了文章提出的方法核心内容，训练两套模型，利用对抗学习来分别产生相关的和不相关的文档；第三，我们简单介绍了论文的实验结果，进行对抗训练会有更好的效果。

最后，给你留一个思考题，除了使用对抗训练，你还能想到什么方法，能够比较好地学习到不属于某一个特定领域信息的排序模型？

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 125 | SIGIR 2018论文精读：偏差和“流行度”之间的关系

下一篇 127 | SIGIR 2018论文精读：如何对搜索页面上的点击行为进行序列建模？

## 精选留言 (1)

写留言



吴曦

2018-07-29



是否可以使用word2vec，对不同领域的词句向量计算距离，从而抽取新领域文本的中间特征？