

对自己深度学习方向的论文有idea，可是工程实践能力跟不上，实验搞不定怎么办？

极市平台 2021-09-03 22:00:00 手机阅读 跟

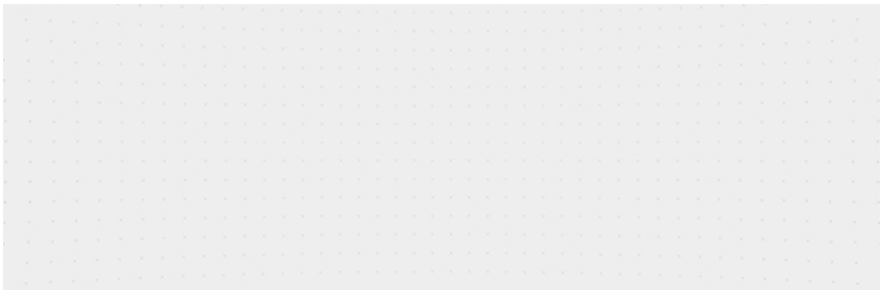
以下文章来源于NLP情报局，作者Giant



NLP情报局

务实浪漫的算法小哥哥。知乎ID：Giant

↑ 点击[蓝字](#) 关注极市平台



作者 | Giant、叶小飞

来源 | 知乎问答

编辑 | 极市平台

极市导读

最困扰科研人（科研狗）或许不是idea的构思，而是繁琐的实验和实践。没有接触过神经网络就要实现idea，自闭的学术之路该怎么打通？本文汇总了一些知乎er在该问题下的精华回答，答主们从自身的经验出发讲述如何快速上道科研路！>>加入极市CV技术交流群，走在计算机视觉的最前沿

原问题背景：本人7月份才接触现在的课题，看了将近40多篇相关英文文献了吧，有了几个idea，开组会，老板们也通过了，接下来就是实现idea，可是自己之前没有接触神经网络方面的，现在要改实验代码，觉得好难啊，是一块硬石头。快泄气。。但是没有退路，必须得做。因为没有正反馈，所以总是很打击自己.....自我怀疑，我都自闭了。。但是，同时也告诉自己，别人也是从不会到会，从无到有。广大的知友们，有没有什么好的建议，让我可以少走点坑，顺利一点...

原问题链接：<https://www.zhihu.com/question/348376942>

回答一

作者：Giant

来源链接：<https://www.zhihu.com/question/348376942/answer/1917476012>

去年五一，我正在洛阳旅行，本已收到了腾讯公司的offer，准备假期过后去实习。这时候导师突然来了电话，让我6月份回实验室做科研。

一瞬间我觉得自己好悲惨，大厂实习泡汤了，研究生最后一段自由时光也成了泡影。

壹伴图



月发文数目： **

月平均阅读： **

文章工具

已发

采集图文 合成多

采集样式 查看

科研的主题是Text2SQL，就是把一句话翻译成对应的SQL语句。这是我毕业的开题方向，当初对NLP这个朝阳领域很感兴趣，有几个不错的idea。可我也就止步于此，没尝试过代码实现，看的论文不到10篇，处境很尴尬。

放弃了实习的念头，我在6月底重回实验室后，经历半年多我终于对Text2SQL领域有了清晰的认识。完成毕业论文的同时，产出了paper和专利，并在10月份获得了耶鲁大学Text2SQL比赛的全球第二名。

Leaderboard - Exact Set Match without Values

For exact matching evaluation, instead of simply conducting string comparison between the predicted and gold SQL queries, we decompose each SQL into several clauses, and conduct set comparison in each SQL clause. Please refer to the paper and the Github page for more details. Notice: Test results after May 02, 2020 are reported on the new release (collected some annotation errors).

Rank	Model	Dev	Test
1 May 02, 2020	RATSQL v3 + BERT (DB content used) Microsoft Research (Wang and Shin et al., ACL '20) code	69.7	65.6
2 Sep. 8, 2020	YCSQL + BERT (DB content used) Anonymous	-	65.3
3 Sep. 8, 2020	ShadowGNN (DB content used) Anonymous	-	64.8
4 May 31, 2020	AuxNet + BART (DB content used) Anonymous	70.0	61.9
4 Dec 13, 2019	RATSQL v2 + BERT (DB content used) Microsoft Research (Wang and Shin et al., ACL '20) code	65.8	61.9
6 May 31, 2020	AuxNet + BART Anonymous	68.0	61.3
7 Feb 18, 2020	RYANSQL v2 + BERT Kakao Enterprise (Choi et al., '20)	70.6	60.6

从开始的懵懂、工程能力不足到最终收获了还不错的结果，我对这半年多经历用了3个方面来概括：

- 1. 阅读领域内近3-5年顶会论文
- 2. 站在巨人的肩膀上，强化工程能力
- 3. 充分利用学校和实验室资源

一、阅读领域内近3-5年顶会论文

阅读已有工作是开展科研的必经之路，既是为了站在一个制高点饱览研究领域这块蛋糕，也是为了提升学术嗅觉与idea能力。同时，可以避免想出了一个很好的idea立即开始复现，结果提交论文时才发现，这个点早就被别人做掉了的极端现象。

高效收集paper的4个途径：

- 1) 在研究领域的公开赛事或榜单上学习Top名次的解决方案（一般都有相应的论文链接）。比如Text2SQL领域有：WikiSQL、TableQA、Spider、CoSQL等比赛，里边有大量值得学习的p

aper。

2) 收集2-3篇综述论文。一篇好的综述概括了该领域的整体进展、已有工作以及潜在的研究方向等等，能够给予我们启发，事半功倍。

3) 在谷歌学术上通过关键词搜索相关论文。论文质量可以根据引用数量、会议等级来衡量。

Google 学术搜索

Text SQL

文章

找到约 245,000 条结果 (用时0.06秒)

时间不限

2021以来

2020以来

2017以来

自定义范围...

按相关性排序

按日期排序

不限语言

中文网页

简体中文网页

☐ 包括专利

☒ 包含引用

创建快讯

Improving text-to-sql evaluation methodology

Q Finegan-Dollak, JK Kummerfeld, L Zhang... - arXiv preprint arXiv ..., 2018 - arxiv.org

To be informative, an evaluation must measure how well systems generalize to realistic, unseen data. We identify limitations of and propose improvements to current evaluations of text-to-SQL systems. First, we compare human-generated and automatically generated ...

☆ 被引用次数: 99 相关文章 所有 8 个版本

Typesql: Knowledge-based type-aware neural text-to-sql generation

L Yu, Z Li, Z Zhang, B Zhang, Q Bader... - arXiv preprint arXiv:1804.09769, 2018 - arxiv.org

Interacting with relational databases through natural language helps users of any background easily query and analyze a vast amount of data. This requires a system that understands users' questions and converts them to SQL queries automatically. In this paper ...

☆ 被引用次数: 121 相关文章 所有 6 个版本

Rat-sql: Relation-aware schema encoding and linking for text-to-sql parsers

B Wang, B Shin, X Liu, Q Polozov... - arXiv preprint arXiv ..., 2019 - arxiv.org

When translating natural language questions into SQL queries to answer questions from a database, contemporary semantic parsing models struggle to generalize to unseen database schemas. The generalization challenge lies in (a) encoding the database relations ...

☆ 被引用次数: 68 相关文章 所有 5 个版本

Hybrid ranking network for text-to-sql

Q Lyu, K Chakrabarti, S Hathi, S Kundu... - arXiv preprint arXiv ..., 2020 - arxiv.org

In this paper, we study how to leverage pre-trained language models in Text-to-SQL. We argue that previous approaches under utilize the base language models by concatenating all columns together with the NL question and feeding them into the base language model in ...

☆ 被引用次数: 16 相关文章 所有 2 个版本

Towards complex text-to-sql in cross-domain database with intermediate representation

J Guo, Z Zhan, Y Gao, Y Xiao, JG Lou, T Liu... - arXiv preprint arXiv ..., 2019 - arxiv.org

We present a neural approach called IRNet for complex and cross-domain Text-to-SQL. IRNet aims to address two challenges: 1) the mismatch between intents expressed in natural language (NL) and the implementation details in SQL; 2) the challenge in predicting ...

☆ 被引用次数: 103 相关文章 所有 4 个版本

点击某一篇文章的“被引用次数”链接，可以跳转进入引用了这篇论文的界面，继续寻找有价值的论文。很多论文其实没有精读的有必要，我们在收集了一批论文后，不妨先读读摘要、实验、结论部分，再确定是否需要通篇精读，以提高效率。

Improving text-to-sql evaluation methodology

☒ 在引用文章中搜索

Spider: A large-scale human-labeled dataset for complex and cross-domain semantic parsing and text-to-sql task

T Yu, R Zhang, K Yang, M Yasunaga, D Wang... - arXiv preprint arXiv ..., 2018 - arxiv.org

We present Spider, a large-scale, complex and cross-domain semantic parsing and text-to-SQL dataset annotated by 11 college students. It consists of 10,181 questions and 5,693 unique complex SQL queries on 200 databases with multiple tables, covering 138 different ...

☆ 被引用次数: 177 相关文章 所有 6 个版本

Towards complex text-to-sql in cross-domain database with intermediate representation

J Guo, Z Zhan, Y Gao, Y Xiao, JG Lou, T Liu... - arXiv preprint arXiv ..., 2019 - arxiv.org

We present a neural approach called IRNet for complex and cross-domain Text-to-SQL. IRNet aims to address two challenges: 1) the mismatch between intents expressed in natural language (NL) and the implementation details in SQL; 2) the challenge in predicting ...

☆ 被引用次数: 103 相关文章 所有 4 个版本

Syntaxsqlnet: Syntax tree networks for complex and cross-domain text-to-sql task

T Yu, M Yasunaga, K Yang, R Zhang, D Wang... - arXiv preprint arXiv ..., 2018 - arxiv.org

Most existing studies in text-to-SQL tasks do not require generating complex SQL queries with multiple clauses or sub-queries, and generalizing to new, unseen databases. In this paper we propose SyntaxSQLNet, a syntax tree network to address the complex and cross ...

☆ 被引用次数: 83 相关文章 所有 6 个版本

当然一些经典的paper是一定要看的，比如做NLP肯定不能错过Transformer、BERT等等。

4) 在Github上搜索资源整合项目。如果这个领域比较热门或者正处于上升期，一般都会有热心用户分享自己整理的资料。例如，关于Text2SQL我已经整理好了一份大礼包，包含了背景、论文、数据集、解决方案、应用案例等内容，帮助感兴趣的小伙伴们节省时间：<https://github.com/yeichens/NL2SQL>

二、站在巨人的肩膀上，强化工程能力

有了学术积累和idea后，就需要快速复现idea。刚开始我有一个误区：科研一定要从0做到1。

后来导师告诉我大可不必这样。聪明的人会先参考别人的解决方案(特别是SOTA)，在学习过程中强化工程和coding水平，形成更优雅的代码风格。这就像站在巨人的肩膀上，我可以基于他们的工作实现自己的想法，取其精华，去其糟粕。

还是以Text2SQL为例。这个任务的数据处理部分特别繁琐，需要同时考虑文本和数据库信息，仅预处理就包含了上千行代码。我参考了Spider上的Top方案后，发现大家在这部分有很多思路是共通的，完全可以借鉴。这样一来，我可以把更多精力放在模型结构设计和后处理上，这两者对最终结果有关键影响。

有的同学表示说，“自己之前没有太多接触神经网络，现在要改实验代码，觉得好难”，这就涉及基本功问题了。我们每天需要额外挤时间来补漏洞，提升自己的coding能力。

如果喜欢看书，我推荐Keras之父的《Python深度学习》和李沐老师的《动手学深度学习》。两本书我都认真看过，非常经典和通俗易懂。

三、充分利用学校和实验室资源

最后一个关键是学会充分利用现有资源，这个资源包括学术资源、人力资源、硬件资源。

学术资源

最直接的学术资源就是实验室的师兄师姐们，还有大Boss——导师。

师兄中肯定有人发过paper，无论是写作还是关于编程和技巧，和他们聊过之后都给我带来了启发。导师是我研究领域中的权威人物，可以帮助我把关idea是否work、是否有足够的竞争力，甚至在没有任何思路的时候点醒我，提供有价值的idea。

人力资源

导师精力有限，往往神龙见首不见尾，不可能顾及所有人，所以带领学弟学妹们的工作一般留给了高年级同学。如果研究的方向他们也感兴趣，大家完全可以一起合作。例如让coding能力强的师弟帮忙做一些实验，coding弱些的同学分析数据和badcase，便于我迭代模型。不仅让他们有所收获，也帮助我分担了很多工作量，一举多得。

如果能带学弟学妹们投中论文，我想他们一定会从心里发出感激的，没准也更愿意在今后的工作中加上我的名字。

硬件资源

俗话说巧妇难为无米之炊，做科研有了idea和coding能力，没有机器也还是白搭。所以无论是在实验室还是公司，要善于利用资源。

我在实验室时，老板很慷慨的采购了服务器，还有Tesla V100显卡让我们做实验。实验室24h可以进出，提供免费饮用水和咖啡。如果做出了成绩（比如会议中稿，比赛拿下SOTA）会有丰厚的物质奖励。

如果实验室没有这些硬件资源，应该主动和老板沟通，请他帮忙协助解决。

总结

路漫漫其修远兮，吾将上下而求索。

做科研是一条艰辛的路，特别是从0到1的过程。但是通向光明之路，从来没有一帆风顺的选择。

为了避免拖延症和摸鱼，我会给自己罗列各个阶段的deadline，然后从学术调研、coding实验、多次迭代方面进行攻坚。如果真的感觉很累，我会停下来离开实验室，去球场和朋友打球；或者跑跑步，呼吸新鲜空气，让自己清醒放松。

最后推荐几个深度学习方向科研工作神器吧，希望能为同学们的科研道路提供帮助：

arxiv：论文收录网站（防止自己的idea被剽窃，完成论文后可以先挂上去，证明原创性）；

PaperwithCode：论文和代码的结合工具，包含了很多NLP、CV方向的SOTA论文和模型实现；

dblp：计算机英文资料收集网站，支持各种字段搜索（会议、作者、时间等）；

Connected Paper：论文引用信息可视化工具，分析出文献的前世今生；

NLPIndex：NLP学术搜索神器，同时链接了paper、code、graph，是上述上个神器的综合体；

DeepI：翻译神器，特点是地道，更接近真人翻译（略胜于谷歌翻译）

diagrams：画图好帮手，制作高大上的流程图、模型图必备，可以在线导出为pdf高清矢量图（放大后图片细节不失真）

- 参考资料
- [1] <https://github.com/yechens/NL2SQL>
- [2] arxiv: <https://arxiv.org/>
- [3] PaperwithCode: <https://paperswithcode.com/>
- [4] dblp: <https://dblp.uni-trier.de/>
- [5] ConnectedPaper: <https://www.connectedpapers.com/>

回答一

作者：叶小飞

来源链接：<https://www.zhihu.com/question/348376942/answer/1923857788>

题主和我反着，我一直比较困扰提出很novel的idea，但我实现idea的速度是很快的，从头搭一个一般复杂的pipeline并调通基本都能控制在一周之内。题主按我这四步走，一定很快就能上道！

第一步：找到baseline

论文一开始对深度学习不熟时，切勿自己蒙着头写代码。既然题主已经看了四十多篇论文，并且有了自己的idea，那么肯定知道自己要做的方向有哪些前人写下的不错的论文，从这些论文中挑选出一到两篇作为你的baseline，作为你的代码开端。

一般一个好的baseline有两个特点：

- 1) 它一定有**开源代码**，而且有较完整的文档（readme）。如果它的github星蛮多的更好，说明很多人用过，质量没问题。
- 2) 它发表的时间是**近两到三年**的。深度学习的发展速度实在过快，时间久远又没人维护的，它用的代码框架什么的可能会过时，比如我之前做colorization, 有几篇不错的都是用的caffe, 比较难拿来直接用。
- 3) 它里面的**公式整体较为清晰**，这样方便你对照着代码来看公式，当你**明白了如何把理论的公式转化为代码时，你基本上道了**。

第二步：深入baseline代码

当你找到了不错的baseline论文后，就要深入阅读、使用它的代码了。在你基础较差时，**提高代码能力最好的方法绝对不是头铁地去拷贝，而是阅读别人写的优质代码**。阅读代码也是有诀窍的，一般分为这么几步：

1. 按照github的文档**配置好你的环境**。几年前配置环境其实是个蛮痛苦的事，那时候我还用caffe,简直就是噩梦，但现在配置环境变得特别傻瓜式，pytorch/tf都特别好配置。
2. 配置好环境后，按它readme的步骤下载相应的数据，运行相应的脚本，**让它的训练先能流畅地跑通**。优质的开源代码基本都能跑通，如果在中间报了错，一般都是library版本不符。但如果是涉及到算法内部的问题，就要走到下一步了。
3. 当你把开源代码跑起来或者遇到算法里的bug之后，**用一个好的ide(例如pycharm)进入debug模式**。在这个模式下的最大好处，就是你可以拿着实际的数据一步一步看代码，哪一步看不懂就对照论文的流程加谷歌一下API的使用方法，很多你干看看不懂的代码结合着tensor shape或者处理过后的数据的shape/value就能看懂。**在这一步一步debug运行代码的过程里**，其实你对**深度学习完整的pipeline怎么写基本就有了一个比较好的认识**。如果时间允许，建议把你找到的几篇baseline都这样跑一遍，并且写个笔记记录下大概流程。

第三步：搭建你自己的pipeline


我一般喜欢根据自己特定的任务从头搭建整个pipeline(包括数据I/O, 预处理，模型搭建，loss定义、训练与验证、测试等等)，但鉴于题主刚接触深度学习代码，可以从第二步中**挑选一个不错的baseline作为你的基础代码，站在它的肩膀上进行修剪**。一般你要做的方向虽然可以用baseline的算法，但数据集往往不同，所以在别人的基础上搭建自己的pipeline最重要的是修改数据读取和预处理部分。在这一步**先不要**把自己的模型加进去，用baseline的模型能跑通你自己的数据即可。

第四步：将你的核心算法融入pipeline

经过了上面三步之后，你已经对深度学习训练预测的流程有了一个比较清晰的认识，这个时候可以开始把你自己的idea加入pipeline. 一般来说你的idea无非是提出了更好的模型，或者更好的训练机制。模型或者loss你在网上一般都能找到整体架构较为类似的代码，比如你要做轻量级图像分割，那大概率可以借鉴Unet，网上一搜一大堆，按照他们的格式仿写即可。这里有个小诀窍，你搭建自己的模型搭好后可以先不放入自己的pipeline里真枪实战的训练，先生成几个随机的tensor扔进去做inference, 看看能不能跑通，输出的shape是否正确。

最后一点点建议，新手想快速入门首选pytorch, tf1.x对新手挺不友好的，而用tf2的研究人员比较少，容易找不到现成的代码。

如果觉得有用，就请分享到朋友圈吧！



极市平台

为计算机视觉开发者提供全流程算法开发训练平台，以及大咖技术分享、社区交流、竞...
765篇原创内容

公众号

△点击卡片关注极市平台，获取最新CV干货
公众号后台回复“CVPR21检测”获取CVPR2021目标检测论文下载~

极市干货

- 深度学习环境搭建：如何配置一台深度学习工作站？
- 实操教程：OpenVINO2021.4+YOLOX目标检测模型测试部署 | 为什么你的显卡利用率总是0%？
- 算法技巧 (trick)：图像分类算法优化技巧 | 21个深度学习调参的实用技巧



CV技术社群邀请函



△长按添加极市小助手

添加极市小助手微信 (ID : cvmart4)

备注：姓名-学校/公司-研究方向-城市（如：小极-北大-目标检测-深圳）

即可申请加入极市 目标检测/图像分割/工业检测/人脸/医学影像/3D/SLAM/自动驾驶/超分辨率/姿态估计/ReID/GAN/图像增强/OCR/视频理解 等技术交流群

每月大咖直播分享、真实项目需求对接、求职内推、算法竞赛、干货资讯汇总、与 10000+ 来自港科大、北大、清华、中科院、CMU、腾讯、百度等名校名企视觉开发者互动交流~

觉得有用麻烦给个在看啦~ 

喜欢此内容的人还喜欢

有趣大会 · ACL2022 (长论文篇)

看个通俗理解吧



卷到纯数学：MyEncyclopedia号主亲历并总结了一份AI工程师的纯数学课程学习之路

MyEncyclopedia



【表面缺陷检测】基于yolov5的PCB表面缺陷检测(附代码)

Alexplore

