



《NLP算法工程师》

|  |  |
| --- | --- |
| **专 业** | **计算机科学与技术** |
| **姓 名** | **田阔** |
| **日 期** | **2022/8/29** |

**摘 要**

目前自然语言处理领域虽然处于高速发展的阶段，但是由于疫情影响，NLP相关岗位需求量减少，同时NLP领域相关毕业生不断增多，这些都使NLP就业难度增大，而且总的来说NLP就业也更偏向研究领域，更专注于领域前沿知识。

**目 录**

[第1章 行业概述与应届生就业情况 - 1 -](#_Toc114087354)

[1.1行业概述 - 1 -](#_Toc114087355)

[1.2应届生就业状况 - 1 -](#_Toc114087356)

[第2章 岗位职能介绍 - 2 -](#_Toc114087357)

[第3章 行业需要的主要能力与技术栈 - 2 -](#_Toc114087358)

[3.1行业需要的主要能力 - 2 -](#_Toc114087359)

[3.2 技术栈 - 3 -](#_Toc114087360)

[第4章 专业知识在行业中的实践路径 - 4 -](#_Toc114087361)

[第5章 就业准备建议 - 5 -](#_Toc114087362)

[5.1平时准备 - 5 -](#_Toc114087363)

[5.2 专业相关 - 5 -](#_Toc114087364)

[结 论 - 6 -](#_Toc114087365)

[**参考文献** - 7 -](#_Toc114087366)

第1章 行业概述与应届生就业情况

## 1.1行业概述

自然语言处理(NLP，Natural Language Processing)在百度百科[1]上的定义为：研究[人](https://baike.baidu.com/item/%E4%BA%BA/13020851)与[计算机](https://baike.baidu.com/item/%E8%AE%A1%E7%AE%97%E6%9C%BA/140338)[交互](https://baike.baidu.com/item/%E4%BA%A4%E4%BA%92/6964417)的[语言](https://baike.baidu.com/item/%E8%AF%AD%E8%A8%80/72744)问题的一门[学科](https://baike.baidu.com/item/%E5%AD%A6%E7%A7%91/2634099)。

处理自然语言的关键是要让计算机“理解”自然语言，所以自然语言处理又叫做自然语言理解(NLU，NaturalLanguage Understanding)，也称为[计算语言学](https://baike.baidu.com/item/%E8%AE%A1%E7%AE%97%E8%AF%AD%E8%A8%80%E5%AD%A6/1036659" \t "https://baike.baidu.com/item/NLP%EF%BC%88%E8%87%AA%E7%84%B6%E8%AF%AD%E8%A8%80%E5%A4%84%E7%90%86%EF%BC%89%E6%8A%80%E6%9C%AF/_blank)(Computational Linguistics)。一方面它是[语言信息处理](https://baike.baidu.com/item/%E8%AF%AD%E8%A8%80%E4%BF%A1%E6%81%AF%E5%A4%84%E7%90%86" \t "https://baike.baidu.com/item/NLP%EF%BC%88%E8%87%AA%E7%84%B6%E8%AF%AD%E8%A8%80%E5%A4%84%E7%90%86%EF%BC%89%E6%8A%80%E6%9C%AF/_blank)的一个分支，另一方面它是人工智能(AI, Artificial Intelligence)的核心课题之一。

从名字我们就可以看出NLP算法工程师主要从事自然语言处理相关的工作，包含：数据处理、对话系统、文本分类、机器翻译、实体识别、大规模文本数据挖掘等工作方向。

## 1.2应届生就业状况

今年受互联网寒冬影响人工智能高速增长的势头有所放缓，2022年增长速度较2021年下降24%。2021年增速75%。



图 1 2022年增速 图 2 2021年增速

据非官方平台统计，nlp岗位数：

2022-07：职位47个，对比去年同期下降了78%，  
 2022-06：职位258个，对比去年同期增长了65%，  
 2022-05：职位57个，对比去年同期下降了67%，  
 2022-04：职位60个，对比去年同期下降了77%

在需求量下降的情况下，人工智能相关的毕业生却越来越多，就业形势已经严峻起来。



图 3 NLP岗位要求

我们看到企业NLP岗位的要求大多需要多年工作经验，说明在岗位减少，求职者增多的双重影响下，企业对岗位的要求已经逐渐变高，这对应届生也是非常不利的。

第2章 岗位职能介绍

通过各个公司的求职信息，NLP岗位职能可以大致归为以下两类：

1. 应用：构建或研发情感分析，知识图谱，问答系统，推荐系统，中文分词，命名实体识别，句法分析等相关模型或者应用系统
2. 研究：大部分公司都要求求职者有紧跟前沿算法和技术，推动自然语言处理在实际应用领域的性能优化和落地的能力，即在工作中具备一定的科研能力。

这两类职能一般都要求同时具有，即能够解决实际问题还能够不断学习前沿技术。

第3章 行业需要的主要能力与技术栈

## 3.1行业需要的主要能力

职能要求主要分为学历，语言，算法，工作经验方面要求：

1. 学历：大厂普遍要求自然语言处理、模式识别、机器学习等相关专业硕士及以上。整体要求基本在本科学历以上



图 4 学历要求

1. 编程语言：有较强的算法基础和编码能力，熟练掌握C/C++，Python，Java等至少一门语言。
2. 专业相关：悉常用的自然语言算法，如word2vec、LDA、HMM、CRF等；熟悉深度学习在自然语言处理中的应用；
3. 工作经验：基本要求有一定的工作经验



图 5 NLP经验要求

1. 研究相关：如果有较强的研究能力，比如在领域顶级会议如ACL、EMNLP、NAACL等发表过高质量论文，在就业方面会有比较大的优势。

## 3.2 技术栈

有底层原理，基础技术，领域任务，落地场景等四个层次。

·底层原理：有机器学习，神经网络，逻辑推理，经验规则四个方面的内容。

·基础技术：语义embedding,语义相似性计算，分词，关系抽取，词性标注，句法分析，命名实体识别，语言模型等。要掌握基本的自然语言算法如word2vec、LDA（隐含狄利克雷分布）、HMM（隐含马尔柯夫模型）、CRF（条件随机场）等。一定的深度学习基础：LSTM、BiLSTM-CRF、TextCNN、注意力机制Attention、Transformer、Bert等

·领域任务：要熟悉自然语言处理的各个细分任务：文本分类，信息抽取，信息检索，阅读理解，机器翻译，语音识别，问答系统，推荐系统，多轮交互，自动摘要，知识图谱，情感计算等。

·落地场景：搜索，广告，信息流，智能客服，直播等。[3]

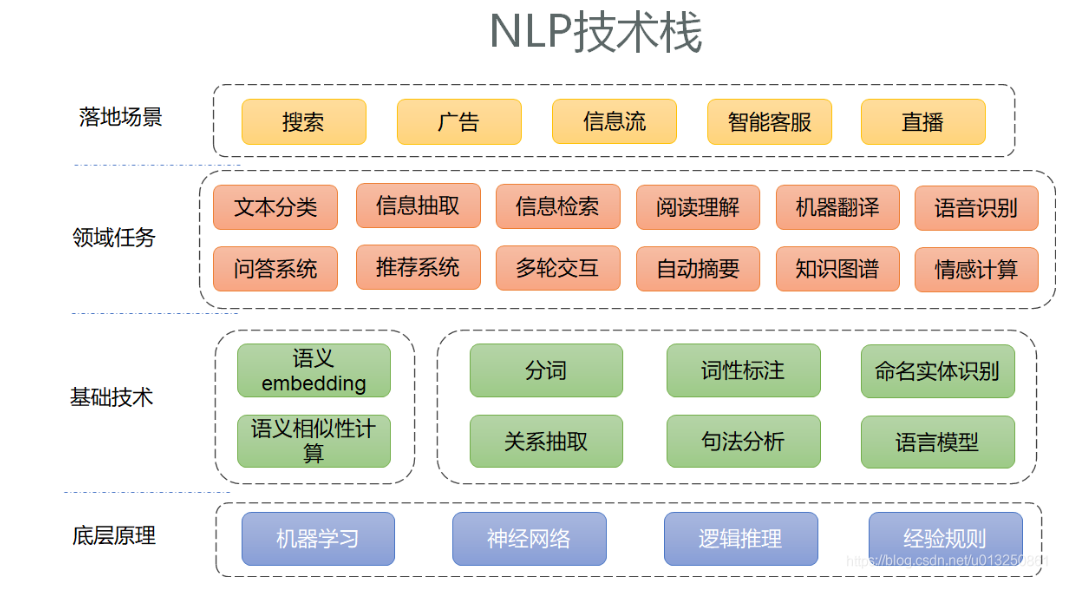


图 6 NLP技术栈

第4章 专业知识在行业中的实践路径

人工智能发展到今天已经深深的融入各行各业，作为NLP作为人工智能的主要方向，已经有许多成熟的落地场景。

·新闻：通过文本分类，对新闻分类显示，通过新闻摘要技术，可以准确的为针对待发布的新闻进行摘要，从而提高传统人工提炼摘要的效率。

·搜索引擎：通过NLP技术，不断提高引擎理解自然语言的能力，从而提供更加优质的服务。

·金融领域：通过大规模数据分析，利用NLP技术，从海量咨询中挖掘结构化信息，达到金融舆情分析的目的。

·对话系统：各行各业的智能客服已经基本普及，能够实现某一领域下订单，简单问答等功能，大幅度降低人力需求，其中使用了语音识别，自然语言理解，自然语言生成，数据库等技术。

第5章 就业准备建议

当今NLP技术在工业领域的运用已经基本成熟，并不断向前发展。并且由于深度学习的兴起，人工智能技术对算力的要求越来越高。企业相对于高校科研院所有着较大的优势，我们可以看到当今深度学习的发展基本上离不开各个主要企业的研究。

所以NLP在工业领域的需求和其他传统行业也不尽相同，头部企业对于研究性人才需求很大，招聘岗位在整体上也更偏向于研究型。

## 5.1平时准备

a）代码题，这是算法面试大厂必考，可以刷一刷Leetcode，牛客平台的算法题，如果还有闲暇时间，可以尝试下每周日上午10:30到12点的Leetcode周日比赛。（在面试中hard难度一般比较少见，我个人觉得是没有必要额外准备的）。

b) 算法理论知识，起码把自己在简历上写的东西都搞明白吧，包括基本的原理，对比其他方法的不同，和实际使用遇到的问题。有的面试确实会问很多简历上没有的算法，但是要同时掌握全部的算法理论还是一个比较有难度的事情的，所以我建议先掌握简历上写的内容就好。

c) 有条件的话尽可能去参加实习，没有实习经历的话也要去网上找下开源模型亲身跑下代码，最好根据自己需要做下改动。一方面是相关的实习经历很加分，另一方面是有去亲身实习后你才能进一步了解这个行业，再决定是否从事这个行业。打打比赛，水水论文也都行。

d) 看点前沿论文，不需要太多，但是对自己感兴趣的领域，你要知道现在主流的方法有什么，有什么优势或者弊端吗？深度学习是一个日新月异的领域，需要时刻保持着对前沿技术的追踪。[2]

## 5.2 专业相关

1. 学历：相关专业的研究生及以上，因为人工智能领域的火热，所以求职也卷的越来越厉害，除了一些本科期间有突出工作（比如发了几篇顶会）的人，大厂基本要求硕士以上学历。
2. 成果：要有相关的项目经历支撑，最好能有论文加持。

结 论

从今年就业形势来看，NLP领域在近几年应该都是一个比较热门的方向，想要拿到一个好的offer，需要在学历，成果各个方面都有一定的资本，总的来说大厂的NLP岗位更侧重于研发方向，需要有一定的科研能力，想要进入这个行业的同学，需要认真思考自己是否真的对这个方向有兴趣。

**参考文献**

[1] [NLP（自然语言处理）技术\_百度百科 (baidu.com)](https://baike.baidu.com/item/NLP%EF%BC%88%E8%87%AA%E7%84%B6%E8%AF%AD%E8%A8%80%E5%A4%84%E7%90%86%EF%BC%89%E6%8A%80%E6%9C%AF/8102216?fr=aladdin)

[2] [NLP算法面试的一些个人建议 - 知乎 (zhihu.com)](https://zhuanlan.zhihu.com/p/472921394)

[3] [人工智能-自然语言处理(NLP)：NLP技术栈\_（csdn.net）](http://人工智能-自然语言处理(NLP)：NLP技术栈_（www.csdn.net）)