

35|云从科技算法岗武功秘籍

1 云从科技面经汇总资料

第一节
云从面经
汇总资料
(整理: 江大白)
www.jiangdabai.com

- ① 面经汇总参考资料
- ② 面经涉及招聘岗位
- ③ 面试流程时间安排
- ④ 云从面经整理心得

1.1 面经汇总参考资料

① 参考资料:

- (1) 牛客网: 云从科技面经-12 篇, [网页链接](#)
- (2) 知乎面经: [点击进入查看](#)
- (3) 面试圈: [点击进入查看](#)

② 面经框架及参考答案:

- (1) 面经框架及参考答案: [点击进入查看](#)
- (2) 大厂目录及整理心得: [点击进入查看](#)

1.2 面经涉及招聘岗位

(1) 全职岗位类

【机器学习算法工程师】、【计算机视觉算法工程师】、【NLP 算法工程师】

1.3 面试流程时间安排

| 云丛科技面试流程-整理：江大白 | | | |
|-----------------|-------|----------------------------|---------|
| | 面试类型 | 面试流程 | 备注（侧重点） |
| 第一面 | 技术面 | 自我介绍+项目/实习经验 +技术问答+算法编程 | / |
| 第二面 | 技术综合面 | 自我介绍+项目/实习经验 +技术问答+算法编程 | / |
| 第三面 | HR面 | 基础人力问题 | / |

PS：以上流程为大白总结归纳所得，以供参考。

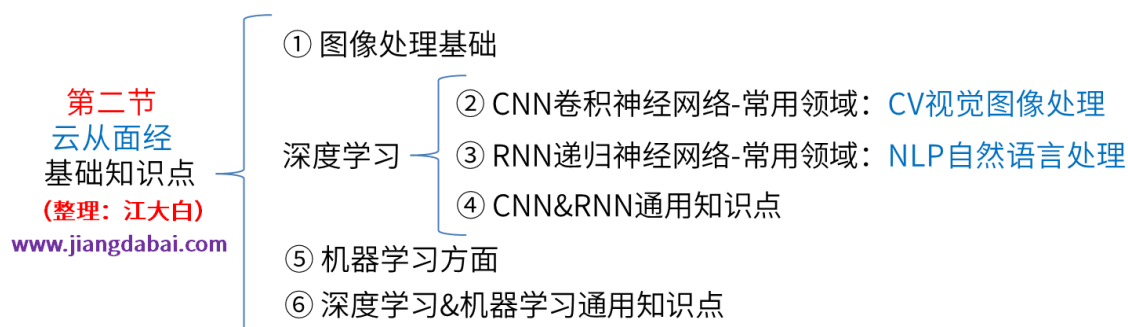
1.4 云丛科技面试心得汇总

★ 会对项目或者经验中的某些算法深挖，比如从头到尾讲一遍，会问很多的细节。最好有一个完全理解的东西。

★ 投得第一家也是面得第一家公司，心里还是挺紧张的。但是面试官很和善，从简单的问题开始问起，很快就调整好了状态（聊技术真的很容易就全身心投入进去了）。可能因为我不是 cs 科班出身，算法问得倒不难，主要就是介绍常见的数据结构，散列表优势是什么，如何实现查找复杂度 $O(1)$ ，写个排序的代码。大多数时间在问项目，问了很多具体细节。

★ HR 很棒，聊天过程非常 nice，当然，技术面试的体验感也是极好的。感觉公司比较年轻，但是已经有不少成果了，成员也大都年轻有朝气。

2 云从科技面经涉及基础知识点



2.1 图像处理基础

2.1.1 讲解相关原理

- 均值滤波的原理是什么?

2.2 深度学习: CNN 卷积神经网络方面

2.2.1 讲解相关原理

2.2.1.1 卷积方面

- deformable convolution 实现原理?
- DW 卷积和普通卷积, 参数量计算量是怎么计算的, 实际运行速度有什么差异?
- 卷积操作, 反卷积操作的意义?

2.2.1.2 网络结构方面

- 简单写一下 resnet block 和 densenet block 的结构, 怎么使用?

2.2.1.3 其他方面

- BN 反向传播更新哪些参数?
- 将 BN 操作合并到之前的卷积操作, 用 numpy 的矩阵操作?

2.3 深度学习：RNN 递归神经网络方面

2.3.1 讲解相关原理

LSTM+Attention 中 attention 的实现方式？

2.4 深度学习：CNN&RNN 通用的问题

2.4.1 基础知识点

- 什么是注意力机制？

2.4.2 模型评价

- 混淆矩阵，如何计算 recall 和 precision？

2.5 传统机器学习方面

2.5.1 讲解相关原理

2.5.1.1 数据准备

无

2.5.1.2 特征工程

① 特征降维

无

② 特征选择

无

2.5.1.3 有监督学习-分类和回归方面

① 逻辑回归 LR

- LR 的原理？

② SVM（支持向量机）

- SVM 的原理？

③ 朴素贝叶斯（Naive Bayes）

- 用贝叶斯公式推导文本共现概率+凸优化中各种方法手推（hessen 阵等）+bert 模型的常规问题？

2.5.1.4 无监督学习-聚类方面

- k-means 的原理讲一下？

2.5.2 手推算法及代码

无

2.6 深度学习&机器学习面经通用知识点

2.6.1 损失函数方面

- 用了哪些损失函数？交叉熵等

2.6.2 激活函数方面

- Sigmoid 和 Softmax 的区别？

2.6.3 正则化方面

- 有哪些正则化方法？batch_normal 的实现方式

2.6.4 过拟合&欠拟合方面

- 如何防止过拟合？哪些措施？

3 云从科技面经涉及项目知识点

第三节
云从面经
项目知识点
(整理: 江大白)
www.jiangdabai.com

- ① 深度学习: CNN卷积神经网络方面
- ② 深度学习: RNN递归神经网络方面
- ③ 强化学习方面
- ④ 机器学习方面

3.1 深度学习: CNN 卷积神经网络方面

3.1.1 目标检测方面

3.1.1.1 讲解原理

- FPN 结构, 解决了什么问题?
- Yolo 的权值是怎么存放的? 训练时标签怎么分配的? loss 怎么计算的?
- 了解哪些特别的小目标检测方法?
- SSD 里 OHEM, 正负样本为什么是 1:3?
- 目标检测里面的评价指标怎么计算? AP、MAP 等
- Yolo 的框和 gt 如何对应?
- 说说 yolo 和 caffe 里面的 padding 有什么区别?
- 说一下目标检测里面的评价指标都怎么算的, ap, map?

3.1.1.2 损失函数

- Focal loss 解决什么问题? 怎么用的? 如何写, 每个参数有什么作用? 拓展到多分类是什么形式?

3.1.1.3 手写代码

- 手写 IOU 计算, 求两个框的 IOU?

- 求 N 个框和 M 个框的 IOU，用 numpy 的矩阵操作

3.2 深度学习：RNN 递归神经网络方面

3.2.1 自然语言处理 NLP

3.2.1.1 讲解原理

① Bert

- 介绍 bert 的原理和预训练模型？

② Transformer

- 介绍 Transorfmer 结构手推计算讲解细节？

3.3 强化学习

无

3.4 机器学习方面

3.4.1 推荐系统

无

4 数据结构与算法分析相关知识点

第四节
云从面经
数据结构与算法分析
(整理：江大白)
www.jiangdabai.com

- ① 数据结构与算法分析：线性表、属、散列表、图等
- ② 算法思想实战及智力题
- ③ 其他方面：数论、计算几何、矩阵运算等
- ④ Leetcode&剑指offer原题

4.1 数据结构与算法分析

4.1.1 线性表

4.1.1.1 数组

- 数组最大子序列？
- k 个长度为 n 的有序数组合并
- 两个排序数组的合并
- 求两个排序数组的交集？
- 两个排序数组的合并？

4.1.1.2 链表

- 递归和非递归的方式反转单向链表？

4.1.2 排序

- 从 10 亿数据里面找出 topK，说说最坏最好和平均情况下的时间复杂度？
- 排序的代码怎么写？快排怎么写？

4.2 算法思想实战及智力题

4.2.1 算法思想实战

- 一维数组，索引差为宽度，两边值较小的为高，求最大矩形面积？

4.2.2 智力题

- 10 个不同颜色的球，大中小三个箱子，那么大箱子中球的数量是奇数的所有可能？
- 9 个球，一个轻的，称重找出来，-用天平两次就可以找出来？

4.3 其他方面

- 常见的数据结构，散列表优势是什么？
- 如何实现查找复杂度 $O(1)$ ，用双指针？

4.4 Leetcode&剑指 offer 原题

无

5 编程高频问题：Python&C/C++方面

第五节
云从面经
编程高频问题
(整理: 江大白)
www.jiangdabai.com

① Python方面：网络框架、基础知识、手写代码相关

② C/C++ 方面：基础知识、手写代码相关

5.1 python 方面

5.1.1 网络框架方面

5.1.1.1 Pytorch 相关

- 怎么加载数据，自己实现不用 pytorch 的 dataloader 应该是什么思路？

5.1.1.2 Tensorflow 相关

- Tensorflow 中 input 如何提升训练速度？
- Tensorflow 中 lstm 的实现方式有哪两种？

5.1.2 基础知识

5.1.2.1 区别比较

- list 和 tuple 的区别？

5.1.2.2 讲解原理

- python 的语言特性，比如面向对象，rlock、git 锁，命名空间之类

5.2 C/C++方面

5.2.1 基础知识

5.2.1.1 区别比较

- 静态成员变量函数和普通的成员变量函数有什么区别？

5.2.1.2 讲解原理

- C++中 const 的作用？

6 操作系统高频问题：数据库&线程&常用命令等

第六节
云从面经
操作系统高频问题
(整理: 江大白)
www.jiangdabai.com

- ① 数据库方面：基础知识、手写代码相关
- ② 操作系统方面：TCP、线程&进程、常用命令相关

6.1 数据库方面

无

6.2 操作系统方面

6.2.1 常用命令

- Linux 的指令，查看进程，查看线程，定位，统计文本内容之类？

6.2.2 其他问题

- git 有哪些常用操作，gitlab 一般有哪些权限 (owner,developer,guest 之类的) ?

7 技术&产品&开放性问题

无