

31|欢聚集团算法岗武功秘籍

1 欢聚集团面经汇总资料

第一节
欢聚集团面经
汇总资料
(整理: 江大白)
www.jiangdabai.com

- 1.1 面经汇总参考资料
- 1.2 面经涉及招聘岗位
- 1.3 面试流程时间安排
- 1.4 欢聚集团面经整理心得

1.1 面经汇总参考资料

① 参考资料:

- (1) 牛客网: 欢聚集团面经-16 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

【推荐算法工程师】、【机器学习算法工程师】、【NLP 算法工程师】、【视频编解码算法工程师】、【语音深度学习工程师】

1.3 面试流程时间安排

欢聚集团面试流程-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	很看重项目及知识点细节 算法编程能力也看重
第二面	技术面	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。

其他注意点：

- 最后有的人可能会有总监面，主要是自我介绍，聊聊他们团队正在做的事情

1.4 欢聚集团面试心得汇总

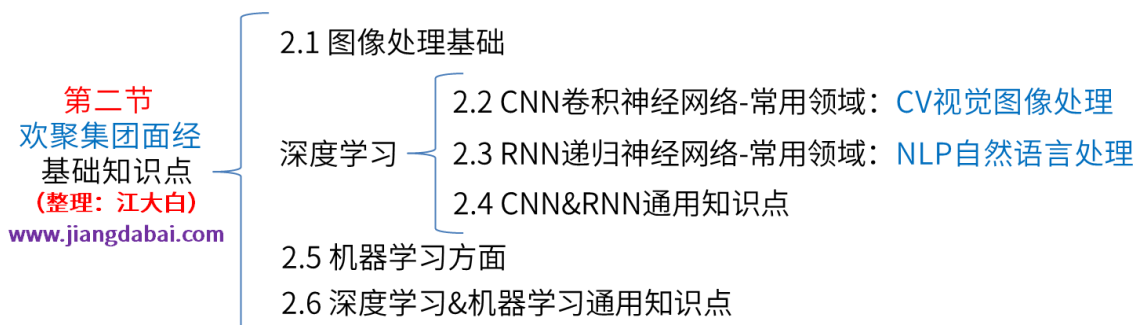
★ 不是问的很深，以 nlp 为主，介绍项目以及里面的细节

★ 实习！算法岗如果不是科班出身，实验室没有项目，一定要去大厂实习。 coding!

相较于前两年，编程能力在面试算法工程师过程中，所占比重越来越大。

★ 面试时，问算法原理，会让边介绍，边进行公式推导(问熟悉的机器学习算法有哪些？回答一个就要推一个公式)

2 欢聚集团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

- 图像左右翻转在 matlab 里面怎么实现? 口头说一下。

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

- dropout 的原理讲一下?

2.2.1.2 其他方面

- CNN 用在文本里和用在图像里有什么区别,用在文本里时卷积核的宽度代表什么?
- Layer Normalization 的作用?

2.3 深度学习: RNN 递归神经网络方面

无

2.4 深度学习：CNN&RNN 通用的问题

2.4.1 基础知识点

- 样本不平衡怎么解决的？
- CNN 和 LSTM 都可以用于分类，两者用于分类有什么区别？
- 说一下 Attention，Attention 有哪些变种，为什么 Attention 模型的 F 指标还比不上作为 baseline 的 textCNN？
- 偏差和方差的原理？
- Attention 的原理是什么？
- Attention 和 RNN 的区别？

2.4.2 模型评价

- AUC 怎么计算，AUC 的含义
- ROC 曲线怎么画？复杂度分析

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

- LDA 解释下？它里面狄利克雷分布怎么应用的？
- 讲一下 tag 提取怎么做的，讲一下 LDA，讲一下隐狄利克雷分布，里面有个辛普森采样了解吗？

② 特征选择

无

2.5.1.3 有监督学习-分类和回归方面

① 分类回归树（集成学习）

- boost 和 bagging 的区别？
- 问随机森林如果判断过拟合，过拟合之后如何调参？
- 随机森林的原理？
- 随机森林为什么可以减小方差？

② 逻辑回归 LR

- LR 和 DNN 的优缺点？

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

- 分类中为什么交叉熵损失函数比均方误差损失函数更常用？

2.6.2 网络优化梯度下降方面

- 梯度优化算法讲一下？

2.6.3 过拟合&欠拟合方面

- 解决过拟合的方法？

3 欢聚集团面经涉及项目知识点

第三节
欢聚集团面经
项目知识点
(整理: 江大白)
www.jiangdabai.com

- 3.1 深度学习: CNN卷积神经网络方面
- 3.2 深度学习: RNN递归神经网络方面
- 3.3 强化学习方面
- 3.4 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

- Faster R-CNN 是如何解决正负样本不平衡的问题?

答: 限制正负样本比例为 1:1, 如果正样本不足, 就用负样本补充, 这种方法后面研究工作用的不多。通常针对类别不平衡问题可以从调整样本数或修改 loss weight 两方面去解决, 常用的方法有 OHEM、OHNM、class balanced loss 和 Focal loss

可参照:

https://blog.csdn.net/weixin_35653315/article/details/78327408#commentsedit

3.1.2 视频编解码

- 一直围绕项目经历 (关于视频编码) 展开询问, 以及项目希望能够应用到具体哪些场景中?

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

① Transformer

- 画 Transformer, 讲一下?

② CRF

- 讲讲 CRF?CRF 和 HMM 的区别，从有向无向图的角度呢？从其他角度呢？CRF 和深度学习的结合还知道哪些？

③ Word2vec

- Word2vec 原理？
- 除了 word2vec 还有哪些词向量，以为比较神经网络向量，答个 glove，然后让比较，我说不会，面完想想我可以说说 tf-idf，词频啊等？

④ 其他

- Cbow 和 skip-gram 哪个训练更快，哪个效果更好，为什么？
- 词向量的评价指标？
- 说下项目中句子相似度匹配的模型？
- 基于检索的问答都有哪些？
- 词向量的评价指标？
- 讲事件分类。数据有多少，样本不平衡怎么解决的，CNN 用在文本里和用在图像里有什么区别，用在文本里时卷积核的宽度代表什么，你怎么选的，为什么要这么选？CNN 和 LSTM 都可以用于分类，两者用于分类有什么区别？说一下 Attention，Attention 有哪些变种，为什么 Attention 模型的 F 指标还比不上作为 baseline 的 textCNN？最后为什么选择 Attention 模型？词向量用什么训练的，数据量有多少，怎么评价词向量的质量的？词向量的维度是多少，为什么要选这个维度？文本分类中的多义词问题可以怎么解决？

3.3 强化学习

无

3.4 机器学习方面

3.4.1 推荐系统

- 关于推荐算法的各种场景如何处理？各种 brain storm，面试官部分赞同，一边聊天一边推进思考过程。
- 协同过滤的原理讲一下？
- 讲讲 CRF。CRF 和 HMM 的区别，从有向无向图的角度呢？从其他角度呢？。CRF 和深度学习的结合还知道哪些？

4 数据结构与算法分析相关知识点

第四节
欢聚集团面经
数据结构与算法分析
(整理: 江大白)
www.jiangdabai.com

- 4.1 数据结构与算法分析：线性表、属、散列表、图等
- 4.2 算法思想实战及智力题
- 4.3 其他方面：数论、计算几何、矩阵运算等
- 4.4 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

- 无序数组求中位数，复杂度分析

升级版：不断的往数组里 push 元素，如何求中位数，复杂度分析

再升级：大规模的数组求中位数

4.1.2 树

- 快排和二叉树层次遍历？

4.1.3 排序

- 快排和二叉树层次遍历？

4.2 算法思想实战及智力题

4.2.1 算法思想实战

- 高楼大厦遮挡问题，一个地平线，给很多楼的左右 x 坐标以及楼高，求有遮挡的面积和？

4.3 其他方面

无

4.4 Leetcode&剑指 offer 原题

无

5 编程高频问题：Python&C/C++方面

第五节
欢聚集团面经
编程高频问题
(整理: 江大白)
www.jiangdabai.com

- 5.1 Python方面：网络框架、基础知识、手写代码相关
- 5.2 C/C++ 方面：基础知识、手写代码相关

5.1 python 方面

5.1.1 基础知识

5.1.1.1 区别比较

- python map 函数和列表表达式的区别？

5.1.1.2 讲解原理

- python 中怎么导入第三方文件？
- python 垃圾回收机制，装饰器

5.2 C/C++方面

无

6 操作系统高频问题：数据库&线程&常用命令等

第六节
欢聚集团面经
操作系统高频问题
(整理: 江大白)
www.jiangdabai.com

6.1 数据库方面：基础知识、手写代码相关

6.2 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 线程和进程相关

- 线程和进程有什么区别？

7 技术&产品&开放性问题

无