

43| 搜狐算法岗武功秘籍

1 搜狐面经汇总资料

第一节
搜狐面经
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 搜狐面经整理心得

1.1 面经汇总参考资料

① 参考资料:

- (1) 牛客网: 搜狐面经-9 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 面经参考答案:

- (1) 面经答案: [点击进入查看](#)

1.2 面经涉及招聘岗位

- (1) 实习岗位类

【推荐算法实习生】

- (2) 全职岗位类

【推荐智能中台算法工程师】

1.3 面试流程时间安排

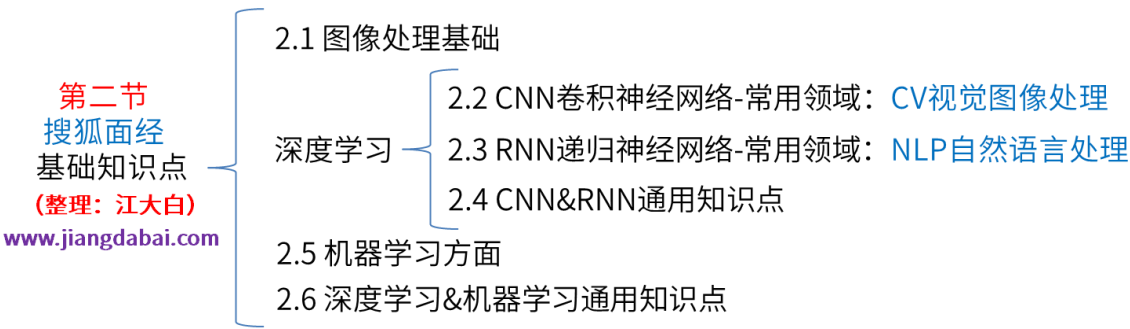
搜狐面试流程-整理：江大白			
	面试类型	面试流程	备注（侧重点）
第一面	技术面	自我介绍+项目/实习经验 +技术问答+算法编程	项目问的很细
第二面	技术面	自我介绍+项目/实习经验 +技术问答+算法编程	/
第三面	总监面	自我介绍+项目经验+公司发展	关注技术在场景中的应用
第四面	HR面	基础人力问题	/

PS：以上流程为大白总结归纳所得，以供参考。

1.4 搜狐面试心得汇总

- ★ 编程相关的知识比较少
- ★ 感觉问计算机视觉相关的问题比较少，机器学习方面的问题比较多。

2 搜狐面经涉及基础知识点



2.1 图像处理基础

无

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 中训练和测试阶段有什么区别？

2.2.1.2 其他方面

- BN 的原理？

2.3 深度学习：RNN 递归神经网络方面

2.3.1 讲解相关原理

- GRU 和 LSTM 的原理讲一下？
- LSTM 相比 RNN 的优点？
- 针对项目，为什么用 Bi-LSTM，数据量这么小，为什么不用传统的 CRF，CRF 的损失函数，怎么做解码的？

2.4 深度学习：CNN&RNN 通用的问题

无

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

- 协同过滤和矩阵分解？

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）

- xgb 和 lgb 的区别？（非常高频，我每次也不光说个 level-wise 和 leaf-wise，还会说 xgb 里面的 block 怎么样的，lgb 怎样从空间和时间上优化，直方图是怎么构建的都会说）
- 问到了 bagging 和 boosting 的区别？
- 针对项目问，为什么要选择 xgb 做特征选择，lr 也可以啊，那说一下你这个场景下，lr 的弊端，xgb 的优势吧？
- xgb 和 lgb 的区别？（非常高频，我每次也不光说个 level-wise 和 leaf-wise，还会说 xgb 里面的 block 怎么样的，lgb 怎样从空间和时间上优化，直方图是怎么构建的都会说）
- xgb 里面不同参数主要是什么作用，问了防止过拟合有哪些参数可以调节？

② 决策树（DT）

- 介绍决策树模型及其扩展？

2.5.1.4 无监督学习-聚类方面

- 为什么用 DBSCAN 呢，k-mean 为什么不用呢，两者有什么区别？

2.5.2 手推算法及代码

- 手推 LR

- 手推 gbd

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

- 手推 logistic 分类公式

2.6.2 激活函数方面

- sigmoid 函数求导？

2.6.3 网络优化梯度下降方面

- 使用了哪些优化算法，有去实际比较过吗？

2.6.4 过拟合&欠拟合方面

- 实际中遇到过拟合了吗？怎么解决的？

3 搜狐面经涉及项目知识点

第三节
搜狐面经
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

3.1.1.1 讲解原理

- RCNN 系列不同算法的不同点？

3.1.1.2 手写代码

- nms 的代码写一下？

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

- 有了解什么最新的深度模型，transformer，bert 等，bert 和 XL-Net 的区别？
- Bert 的 Multi-Head Attention？

② Transformer

- 介绍 transformer, multihead attention？

③ Word2vec

- word2vec 负样本采样？

④ 其他

- 针对项目，用的 fasttext 做的，query 和 title 多少维啊，这个模型说一下？
- 讲一些 query 和 title 的一些文本相似度的方法？

3.3 强化学习

无

3.4 机器学习方面

3.4.1 推荐系统

3.4.1.1 讲解原理

- fm ffm wide&deep deepfm 的原理?
- 高维 id 特征怎么处理?

4 数据结构与算法分析相关知识点

第四节
搜狐面经
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

- 一个乱序数组, 最长的连续数列有多长?

4.1.1.2 字符串

- 字符串的所有排列?

4.1.2 排序

- 有什么稳定的排序么? 简述思路。
- 快排思想和时间复杂度, 最差是什么情况, 时间复杂度是多少?

- 写一下堆排序？

4.2 算法思想实战及智力题

4.2.1 算法思想实战

- 一个扔骰子的问题，求数学期望？
- 哈希表和红黑树的原理？

4.3 其他方面

无

4.4 Leetcode&剑指 offer 原题

- 剑指 offer 原题：字符串的所有排列

5 编程高频问题：Python&C/C++方面

第五节
搜狐面经
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

无

5.2 C/C++方面

5.2.1 基础知识

- Python 中 import 导包是 pip 安装的包还是文件夹下的包（如果有同名包）？ 如果

有同名包，文件命名会出什么问题？

- " __if__ "==" __main__ "的作用是什么？

6 操作系统高频问题：数据库&线程&常用命令等

第六节
搜狐面经
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 常用命令

- Linux 的常用命令 chmod

7 技术&产品&开放性问题

7.1 技术方面

- 用户冷启动怎么做？