



人工智能

算法岗

江湖武林秘籍(下)

献给各路豪杰



算法岗位面试必备

现有一本

江湖之中风云莫测

武林盟



# 人工智能算法岗江湖目录

## 武林秘籍（上册）

|                       |         |
|-----------------------|---------|
| <b>武林秘籍练功必读</b> ..... | 第 1 页   |
| 1. 字节跳动算法岗武功秘籍.....   | 第 6 页   |
| 2. 阿里巴巴算法岗武功秘籍.....   | 第 68 页  |
| 3. 腾讯算法岗武功秘籍.....     | 第 104 页 |
| 4. 百度算法岗武功秘籍.....     | 第 137 页 |
| 5. 华为算法岗武功秘籍.....     | 第 180 页 |
| 6. 美团算法岗武功秘籍.....     | 第 205 页 |
| 7. 京东算法岗武功秘籍.....     | 第 229 页 |
| 8. 网易算法岗武功秘籍.....     | 第 256 页 |

## 武林秘籍（中册）

|                        |         |
|------------------------|---------|
| 9. 拼多多算法岗武功秘籍.....     | 第 276 页 |
| 10. Vivo 算法岗武功秘籍.....  | 第 298 页 |
| 11. 招银网络算法岗武功秘籍.....   | 第 311 页 |
| 12. 360 公司算法岗武功秘籍..... | 第 326 页 |
| 13. 海康威视算法岗武功秘籍.....   | 第 339 页 |
| 14. 快手算法岗武功秘籍.....     | 第 354 页 |
| 15. 小米算法岗武功秘籍.....     | 第 375 页 |
| 16. 深信服算法岗武功秘籍.....    | 第 392 页 |
| 17. 作业帮算法岗武功秘籍.....    | 第 405 页 |
| 18. 滴滴算法岗武功秘籍.....     | 第 420 页 |
| 19. 蚂蚁金服算法岗武功秘籍.....   | 第 439 页 |
| 20. 顺丰科技算法岗武功秘籍.....   | 第 453 页 |
| 21. 依图科技算法岗武功秘籍.....   | 第 464 页 |
| 22. 旷视科技算法岗武功秘籍.....   | 第 476 页 |

|                       |         |
|-----------------------|---------|
| 23. 微软算法岗武功秘籍.....    | 第 494 页 |
| 24. Oppo 算法岗武功秘籍..... | 第 507 页 |
| 25. Bigo 算法岗武功秘籍..... | 第 518 页 |
| 26. 猿辅导算法岗武功秘籍.....   | 第 534 页 |

## 武林秘籍（下册）

|                        |         |
|------------------------|---------|
| 27. 中兴算法岗武功秘籍.....     | 第 546 页 |
| 28. 商汤科技算法岗武功秘籍.....   | 第 557 页 |
| 29. CVTE 算法岗武功秘籍.....  | 第 573 页 |
| 30. 大华算法岗武功秘籍.....     | 第 583 页 |
| 31. 欢聚集团算法岗武功秘籍.....   | 第 600 页 |
| 32. 平安科技算法岗武功秘籍.....   | 第 610 页 |
| 33. 大疆算法岗武功秘籍.....     | 第 622 页 |
| 34. 蘑菇街算法岗武功秘籍.....    | 第 631 页 |
| 35. 云丛科技算法岗武功秘籍.....   | 第 642 页 |
| 36. 追一科技算法岗武功秘籍.....   | 第 653 页 |
| 37. 爱奇艺算法岗武功秘籍.....    | 第 663 页 |
| 38. 搜狗算法岗武功秘籍.....     | 第 674 页 |
| 39. 地平线机器人算法岗武功秘籍..... | 第 683 页 |
| 40. 58 集团算法岗武功秘籍.....  | 第 693 页 |
| 41. Keep 算法岗武功秘籍.....  | 第 704 页 |
| 42. 寒武纪算法岗武功秘籍.....    | 第 712 页 |
| 43. 搜狐算法岗武功秘籍.....     | 第 724 页 |
| 44. 有赞算法岗武功秘籍.....     | 第 733 页 |
| 45. 知乎算法岗武功秘籍.....     | 第 742 页 |

# 27|中兴算法岗武功秘籍

## 1 中兴面经汇总资料

第一节  
中兴面经  
汇总资料  
(整理: 江大白)  
[www.jiangdabai.com](http://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 中兴面经整理心得

### 1.1 面经汇总参考资料

#### ① 参考资料:

- (1) 牛客网: 中兴面经-33 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&答案&目录&心得:

- (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

【ZTE 南京算法工程师】、【图像算法工程师】、【蓝剑计划编码算法工程师】、【ZTE 优招算法工程师】、【AI 算法工程师】



## 1.3 面试流程时间安排

| 中兴面试流程-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<br>+技术问答+算法编程 | /           |
| 第二面           | 综合面  | 自我介绍+项目/实习经验<br>+技术问答      | 结合实际应用问解决思路 |
| 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 中兴面试心得汇总

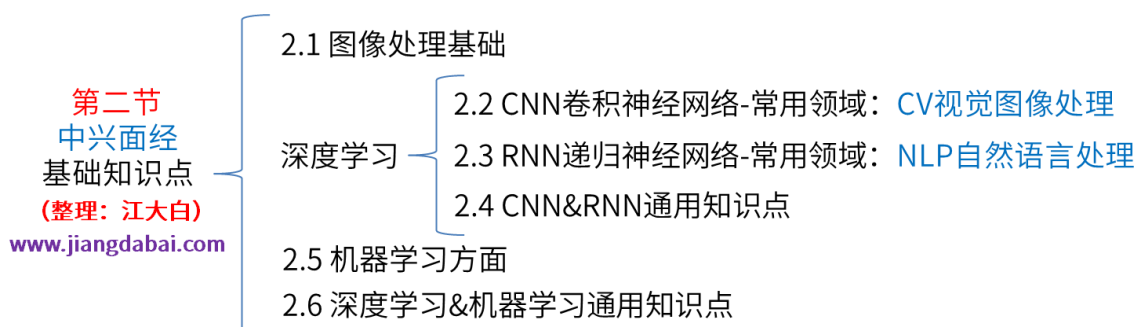
- ★ 感觉招的人比较多，问的问题比较开放，完全底层的问的很少，感觉不是很深入
- ★ 感觉都是发散的面，深挖项目细节比较少，算法题问的也比较少
- ★ 因为公司比较大，所以去校招的时候，有的时候会问。

转岗问题：愿不愿意通信算法转开发？（算法岗位少）

地点问题：为什么工作地点选择深圳，愿不愿意选南京？

- ★ 写进简历的项目一定要好好捋清楚思路，因为有的面试官会被问很深。

## 2 中兴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

- 传统的特征提取方法？
- canny 原理讲一下？
- 主要使用的 opencv 吗，提取了那些特征？
- 颜色特征是在哪个颜色空间下提取的？
- 用过边缘提取吗？(canny,laplace)
- 霍夫变换了解吗？霍夫直线检测的原理知道吗？
- 霍夫变换的一个点对应原图的什么？

## 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

- 卷积核有哪些？
- 介绍一下 Dropout？

#### 2.2.1.2 网络结构方面

- Resnet 短路连接作用？
- Vgg 网络什么特点，用到了哪几种卷积核？

## 2.3 深度学习：RNN 递归神经网络方面

### 2.3.1 讲解相关原理

- 介绍 RNN 和 LSTM？

### 2.3.2 手绘网络原理

- 画 RNN 图写一下公式，根据图介绍一下？

## 2.4 深度学习：CNN&RNN 通用的问题

### 2.4.1 基础知识点

- 样本不平衡如何不平衡怎么办？

## 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

- PCA 的原理介绍一下？

##### ② 特征选择

无

#### 2.5.1.3 有监督学习-分类和回归方面

- SVM 支持向量垂直还是平行、SVM 的拉格朗日做什么的？

#### 2.5.1.4 无监督学习-聚类方面

- Kmeans 聚类的原理？什么数据不能聚？
- kmeans 聚类初始值的选择？

## 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 激活函数方面

- 激活函数有什么作用？

### 2.6.2 压缩&剪枝&量化&加速

- 模型量化会吗？

### 2.6.3 过拟合&欠拟合方面

- 什么是过拟合，如何解决？

## 3 中兴面经涉及项目知识点

第三节  
中兴面经  
项目知识点  
(整理: 江大白)  
[www.jiangdabai.com](http://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

- Yolo3 的 Anchor 怎么计算？
- Faster RCNN 怎么改的？



- RPN 能不能多分类？

#### 3.1.1.2 损失函数

- Yolo3 的损失函数是哪些？

#### 3.1.2 图像分类

- 分类了解吗，知道哪些分类算法，介绍一下？

### 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

##### ① Bert

- 对 Bert 的理解，优势和劣势

##### ② Word2vec

- word2vec 随机初始化和 onehot 的区别？

### 3.3 强化学习

无

### 3.4 机器学习方面

无

## 4 数据结构与算法分析相关知识点

第四节  
中兴面经  
数据结构与算法分析  
(整理: 江大白)  
[www.jiangdabai.com](http://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 链表

- 双向链表如何插入结点

#### 4.1.2 树

##### 4.1.2.1 二叉树

- 二叉树怎么遍历

##### 4.1.2.2 堆

- 有哪些数据结构？图、栈等
- 栈实现，写代码？
- 堆栈区别？堆栈中分别存放什么？栈修饰函数有什么作用？

#### 4.1.3 排序

- 归并排序知道吗？与快排序、堆排序有什么显著的优势？
- 为什么快排会用的最多？

- 排序算法的稳定性了解吗？解释一下是什么稳定
- 堆排序的原理

## 4.2 算法思想实战及智力题

无

## 4.3 其他方面

- 求一个序列左边大于该数，右边小于该数的数有多少个，说下思路？

## 4.4 Leetcode&剑指 offer 原题

无

# 5 编程高频问题：Python&C/C++方面

第五节  
中兴面经  
编程高频问题  
(整理: 江大白)  
[www.jiangdabai.com](http://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 网络框架方面

- 深度学习框架会哪几种？

### 5.1.2 基础知识

#### 5.1.2.1 线程相关

- 多进程实现？多线程实现？

### 5.1.2.2 区别比较

- 浅拷贝和深拷贝区别？

### 5.1.2.3 讲解原理

- Python 数据类型有哪些？6 大数据类型。
- Python 如何连接数据库？
- Python 列表求长度，列表如何增加元素？

## 5.2 C/C++方面

### 5.2.1 基础知识

#### 5.2.1.1 内存相关

- 数组存在堆里还是栈里？
- 讲讲垃圾回收机制？

#### 5.2.1.2 区别比较

- C++与 python 区别？
- 重载与覆盖区别？

#### 5.2.1.3 讲解原理

- 面向对象详细介绍一下，继承多态封装？
- C++多态怎么实现？
- 虚函数关键词？
- 子类重载父类使用什么关键字？

## 6 操作系统高频问题：数据库&线程&常用命令等

第六节  
中兴面经  
操作系统高频问题  
(整理: 江大白)  
[www.jiangdabai.com](http://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 知道哪些命令？
- LINUX 系统下怎么设置文件权限，怎么提升用户权限。答：chmod u+x,chmod 777,root 权限
- 问 chmod 777 中每个 7 是什么意思？答：不知道，我之后会看 LINUX 系统方面的知识

#### 6.2.2 其他问题

- 对内存熟悉吗？碰到过内存泄漏问题吗？怎么解决的？

## 7 技术&产品&开放性问题

### 7.1 技术方面

- pandas 数据处理。分类函数是什么？遇到大的文本数据内存读不下怎么办？

答：分类函数 groupby；大的数据可以考虑降低数值精度，比如将 float64(默认数值类型)转换为 float32 甚至 float16。

- 代码风格如何？代码碰到警告解决吗？怎么解决？
- 场景题：如果数据量不是很大的话，怎么做文本匹配，对数据怎么处理？



## 28|商汤科技算法岗武功秘籍

### 1 商汤科技面经汇总资料

第一节  
商汤面经  
汇总资料  
(整理: 江大白)  
[www.jiangdabai.com](http://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 商汤面经整理心得

#### 1.1 面经汇总参考资料

##### ① 参考资料:

- (1) 牛客网: 商汤科技面经-24 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&答案&目录&心得:

- (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) 实习岗位类

【见习研究院员】、【计算机视觉暑期实习岗】

##### (2) 全职岗位类

【视觉算法研究员】、【计算机算法工程师】

## 1.3 面试流程时间安排

| 商汤科技面试流程-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<br>+技术问答+算法编程 | 项目+算法+基础知识 |
| 第二面             | 技术面       | 自我介绍+项目/实习经验<br>+技术问答+算法编程 | /          |
| 第三面             | 技术Leader面 | 自我介绍+项目经验+公司发展             | 解决问题的思路    |
| 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 商汤科技面试心得汇总

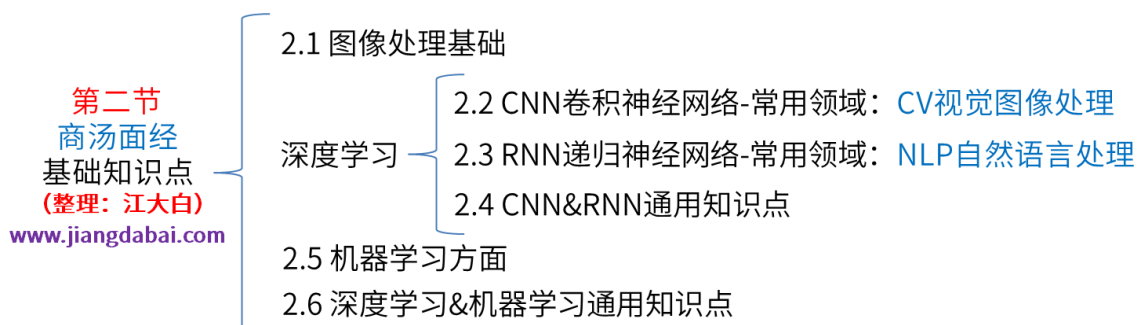
★ 感觉公司技术大佬很多，而且面试问的非常细，特别考虑你的思考和基础能力。主要考察整体把握和积累。特别是你接触的方面，比如 mobilenet 为什么快？有多少层，有多少参数，如果不是深入解除的话，很难重点记忆。

★ 最好根据项目的一个点，吸引面试官注意，给出一个轮廓，再不断细化，同时可以分点答题，思路清晰。

★ 比较硬核，不太好啃。文章和项目会剥茧抽丝，每一个细节都问的非常深入。

★ 在商汤基本上每个人都要熟悉 Linux 和 C++，当然 python 每个人都会。

## 2 商汤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

- 简单介绍下 SIFT，你怎么保证双目相机同时曝光？
- 光流图如何计算的？光流图应变如何计算的？
- 有没有什么数学方法能够去除特征矩阵中的噪声？
- bfs 实现最大连通域？
- 对一个灰度图进行均值滤波操作，写一个函数，可以使用 numpy，要求输入输出的图片和输入的图片的 shape 保持一致。

### 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

- 普通卷积、DW PW 卷积计算量推导？
- 卷积的底层实现/加速技巧

##### 2.2.1.2 池化方面

- Max Pooling 反向传播怎么算？ROI Align 和 ROI pooling 有什么区别？

### 2.2.1.3 网络结构方面

- 网络的基础知识，如一些基础网络以及变体？
- 残差网络结构原理？
- detnet 的原理？
- SENET 原理，接在 ResNet 还有 inception 的什么位置？
- BN、IN、LN、GN 原理，BN 为什么有效？
- mobile net shuffle net 具体结构，如何降低计算量，给了我一个 DW 卷积具体实例让我算降低了多少计算量？
- Inception v1 为什么能提升性能？
- MobileNet v1、V2 介绍？
- MobileNet V2 中的 Residual 结构最先是哪个网络提出来的
- Resnet inception 结构对比？
- 各种网络结构中，resnet 为什么好？
- mobilenet 为什么快？
- 怎么证明 mobilenet 的快是因为用了深度可分离卷积？mobilenet 有多少层？mobilenet 有多少参数？
- 了解哪些轻量级的检测器？
- 类别不平衡用 Inception 好还是 Resnet 好？
- 细粒度分类用 inception 好还是 Resnet 好？

### 2.2.1.4 其他方面

- 如何解决类内的检测？
- Batch Normalization 的作用，BN 的处理是否在通道层面？
- 关于 BN 层。可学习参数，BN 层的作用，在训练阶段和预测阶段的有什么不同，了解 GN 吗？
- 梯度消失的原因是什么？

- 现有两个特征向量，怎么分析他们的相似度？
- 那么在反向传播时同一个 mini-batch 在共享卷积层的末端是否需要除以 batch size？为什么？
- 多任务 CNN 训练过程中不同任务的数据输入顺序是怎样的？一个 mini-batch 中的输入是怎样组成的？

### 2.2.2 数学计算

- 给出 feature map 和卷积核尺寸，求参数量

### 2.2.3 手写算法代码

- numpy 写个 batch norm 层

### 2.2.4 激活函数类

- Softmax+celoss 工程上如何防止上下溢出？

## 2.3 深度学习：RNN 递归神经网络方面

### 2.3.1 讲解相关原理

- RNN 为什么 long-term dependency 做不好？

## 2.4 深度学习：CNN&RNN 通用的问题

### 2.4.1 基础知识点

- 注意力机制怎么实现？

## 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

- 讲讲 PCA？矩阵中的特征值和特征向量是什么意思？
- PCA 从特征值分解角度如何解释（建议要看下 pca 推导）
- PCA 分解是怎么做的？

##### 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）

- 随机森林介绍一下？

##### ② SVM（支持向量机）

- SVM 是什么意思，手推 SVM？

### 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

- 常用损失函数介绍，包括公式写出来，项目中用的损失函数，为什么？
- 介绍一下 triple loss？
- 为什么 smooth L1 Loss 更有效？
- 不同任务的损失是如何设置的？

### 2.6.2 激活函数方面

- Sigmoid 和 Softmax 的区别？

### 2.6.3 网络优化梯度下降方面

- 梯度优化过程原理介绍？
- SGD、Adam 之类优化的原理？
- SGD 使用 mini batch 优化和使用所有优化样本优化哪个更好，为什么？
- 分布式：如何实现分布式 sgd？

### 2.6.4 压缩&剪枝&量化&加速

- 如何修改轻量级的模型？
- 介绍 Mimic 知识蒸馏是怎么做的？
- 现在有什么降低模型复杂度的方法，我说了两点，使用一些降低计算量的结构比如 mobile net shuffle net，然后模型剪枝？
- 模型剪枝的方法，具体细节？
- Tensorrt 原理，如果不做量化，还可以加速吗？

## 2.6.5 其他方面

- 正负样本不平衡的解决？

## 3 商汤科技面经涉及项目知识点

第三节  
商汤面经  
项目知识点  
(整理: 江大白)  
[www.jiangdabai.com](http://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

- 为什么说 resnet101 不适合目标检测？
- 小目标在 FPN 的什么位置检测？
- IOUNet 的原理，能有多少提升？
- NMS 和 soft-nms、softer-nms 的原理？
- 讲 FPN 经验，proposal 怎么映射回多 level featuremap 上？
- 如何解决 multiscale 问题？
- FCOS 这种模型的思想？
- anchor based 和 anchor free，anchor 哪些缺点？
- 为什么要用 weight BCE+dice loss 的损失，focal loss 了解吗？
- 视觉显著性检测和深度学习中常见的 attention 有什么关系？
- 检测、分割、分类熟悉哪些？分类网络有哪些？



- Faster-RCNN 整个从输入到输出的框架流程,用的 roi pooling 还是 roi align?
- RPN 的原理?
- 说一下 SSD 的原理, 和 faster-rcnn 的异同点?
- faster-rcnn 的回归目标是如何定义的? 用的什么 loss?
- 常用的目标检测算法, one-stage 和 two-stage 的区别?
- CornerNet 介绍, CornerPooling 是怎么做的, 怎么解决 cornernet 检测物体合并为一个框的问题?
- 目标检测在工程中应用有没有遇到一些问题? 检测类别冲突怎么办?
- 为什么 ssd 比 faster rcnn 慢, 介绍 r-fcn, 介绍 ohem?
- 目标检测的回归 loss 是什么, 为什么这么用?
- 单阶段比双阶段难训练?
- 讲一下 Faster RCNN 的 RPN, 为什么前景背景分类用 sigmoid 而不是 softmax, 讲一下 FPN, region proposal 时怎么知道 roi 是属于哪一个 feature map 的。
- 简单介绍 centernet 和 cornet?
- 使用 ssd-mobilenetv1-coco 的预训练模型, finetune 后的 map 是 0.956@IOU0.5, map 的误报率是多少?
- 针对小目标, 目标检测的解决措施?
- anchor-free 为什么重新火起来?
- nms 很耗时吗? 时间复杂度? 一般预测时会有多少个候选框?
- NMS 是怎么做的简述?
- yolo 的 V1, V2, V3 有什么不同怎么改进的?

### 3.1.1.2 损失函数

- Focal loss 具体怎么操作?
- Focal loss 是如何进行难分样本挖掘的? (公式我有点忘了, 就说了大概, 主要思想是说出

来了，因为是取幂，对于简分样本损失会比较小，取幂后会更小，反传时更新梯度会很小很小，基本可以忽略，但对于难分，取幂后还是很大)

### 3.1.1.3 手写代码

- nms 的代码手写，并分析时间复杂度，比如  $O(N^2)$
- 目标检测中的 IOU 计算（给定两个框的左上角和右下角坐标，结构体自己定义），如果用 GPU 实现的流程（分 block 和 thread 等）

### 3.1.2 目标追踪

- SiamRPN 原理讲一下？

### 3.1.3 图像分割

- 聊语义分割的发展点，分割的发展可能性，无监督学习在分割的应用？
- 如何解决分割的样本不均衡的问题？
- 如何解决车道线分割有车阻拦视线导致分割结果断裂的问题？
- 讨论语义分割未来可能的发展点？

### 3.1.4 关键点检测

- 在人脸关键点和检测中的 mimic 是怎么做的？为什么不在 logits 输出上做？用 l2 loss 吗？
- 人脸关键点使用 pose 做 multitask 为什么 landmark 会有提升？
- 如何利用 OpenCV 实现的人脸对齐？

## 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

- 懂马尔科夫吗？
- 深度学习中提取词向量算法列举、原理、作用以及大致一个发展历程？

### 3.3 强化学习

无

### 3.4 机器学习方面

无

## 4 数据结构与算法分析相关知识

第四节  
商汤面经  
数据结构与算法分析  
(整理: 江大白)  
[www.jiangdabai.com](http://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

- 求数组中出现次数超过一半的数字？
- 旋转数组，找最小值，能否用递归做？
- 数组的最大子区间和？如果数组是二维的怎么办？就是一个矩阵，求里面最大的一个子矩阵的和，如果数组首尾相连怎么办？
- 用一个 3\*4 的二维指针数组，完成数组顺时针翻转？
- 有序数组合并？
- 最大连续子串的区间？
- 无序数组中，只有 1 个数重复了奇数次，其他数都重复了偶数次，如何找到？
- 多个数组，都是有序的，想求 topk？

- 给你一组乱序的数组  $a$ ，要找出一组  $l, r$ ，使得  $a[r] - a[l]$  最大？
- 寻找数组中唯一/唯二只出现一次的数？

#### 4.1.1.2 链表

- 单链表的翻转？
- 链表倒数第  $k$  个节点？
- 写一个链表的快速排序，自定义结点。

#### 4.1.1.3 字符串

- 两个栈实现队列，并计算时间和空间复杂度？

### 4.1.2 树

#### 4.1.2.1 二叉树

- 定义一颗多叉树，并从  $\{0\ 1; 0\ 2; 0\ 3; 0\ 4; 3\ 4; 3\ 5; 3\ 6\}$  build 这棵树？
- 判断二叉树是否包含另一二叉树？
- 之子形打印二叉树？

#### 4.1.2.2 堆

- 从一百万个数里面找前 1000 大的（堆排，建一个大小为 1000 的小根堆），除了堆排呢？  
(快排，二分+partition)

## 4.2 算法思想实战及智力题

### 4.2.1 算法思想实战

- 两个 queue 实现 stack

### 4.2.2 智力题

- $M \times N$  的方格中有多少个正方形、多少个矩形、有多少种不同面积矩形？

- 在没有能量损失的理想台球桌上任意击球，满足什么条件下球必然进洞？
- 给定  $n \times n$  的正方形，每个点都有权值，有些点是障碍物，问从左上角走到右下角的最小花费？

## 4.3 其他方面

### 4.3.1 数论

- 斐波那契数列及优化？

### 4.3.2 计算几何

- 已知二维平面上有很多点，求出最多有多少点在同一条直线上？

### 4.3.3 概率分析

- 已有两个能生成 0 到 1 之间的数，并且这些数是均匀分布的随机生成器，给定一个任意的三角形，如何能在三角形内等概率随机的生成一个点？(然后对于多边形呢)

## 4.4 Leetcode&剑指 offer 原题

- Leetcode200 或 695：相似，不同是把 1 当作了海，0 当作岛屿（没有本质区别），然后统计岛屿面积，将岛屿面积小于给定阈值  $t$  的全置为 1。

## 5 编程高频问题：Python&C/C++方面

第五节  
商汤面经  
编程高频问题  
(整理: 江大白)  
[www.jiangdabai.com](http://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 网络框架方面

#### 5.1.1.1 Pytorch 相关

- 使用 Pytorch 对 cifar10 数据集分类的整个代码流程，构建模型的过程是怎么样的？
- pytorch 多 gpu 训练机制的原理，优化器以及网络参数保存机制？

#### 5.1.1.2 Tensorflow 相关

- 对于 tensorflow 和 pytorch 的了解，越多越好？
- tensorflow 训练执行过程，forward 的代码过程？

### 5.1.2 基础知识

#### 5.1.2.1 区别比较

- list 和 tuple 的区别？
- Python is 和 == 的区别？

#### 5.1.2.2 讲解原理

- Import 搜索过程？
- python 常用的库报一遍？
- python 里如何实现类似 c++ 里引用？（在函数里改变基础类型，这里回答可以用 list 传入，进行修改）
- python 装饰器？
- python 类似引用？

## 5.2 C/C++方面

### 5.2.1 基础知识

#### 5.2.1.1 区别比较

- 浅拷贝和深拷贝？

#### 5.2.1.2 讲解原理

- C++多态、引用和指针区别、类的概念？
- C 里多态虚函数？
- vector 的相关问题？
- 虚函数怎么实现的，怎么用的？

#### 5.2.1.3 讲解应用

- cpp 中的 delete 和 malloc 的区别？
- cpp 中 vector 的底层实现原理，如何进行内存分配？
- 一个 cpp 程序的内存分配是什么样的？

## 6 操作系统高频问题：数据库&线程&常用命令等

第六节  
商汤面经  
操作系统高频问题  
(整理: 江大白)  
[www.jiangdabai.com](http://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

- 假如文件中有很多文件，每个文件中又有很多文件，如何删除全部文件？

### 6.2.2 常用命令

- github 的常用操作：上传、合并、分支之类的？
- linux 的常用操作：查看文件大小、删除文件、查看文件行数？

## 7 技术&产品&开放性问题

### 7.1 技术方面

- 比赛中如何解决样本不平衡，以及用了什么训练技巧，为什么能起作用，其原理是什么？  
数据增广方法？
- 你认为目前 video 和知识蒸馏这两个方向的挑战和可以改进的地方在哪？
- 你觉得网络模型和硬件平台是什么关系？
- 怎么计算视差的，说一下 SGBM 原理，你怎么根据立体像对同名点计算大地坐标？



# 29|CVTE 算法岗武功秘籍

## 1 CVTE 面经汇总资料

第一节  
CVTE面经  
汇总资料  
(整理: 江大白)  
[www.jiangdabai.com](http://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 CVTE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: CVTE 面经-20 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&答案&目录&心得:

- (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

【深度学习算法工程师】、【图形计算算法工程师】、【数据挖掘算法工程师】、【NLP 算法工程师】

## 1.3 面试流程时间安排

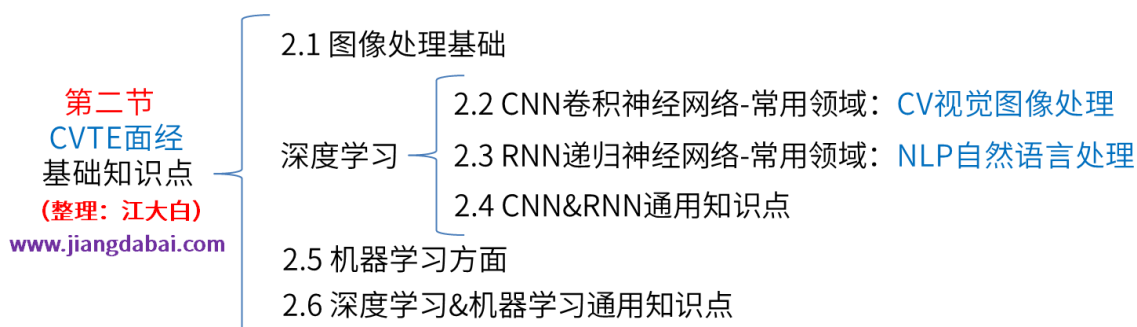
| CVTE面试流程-整理：江大白 |      |                            |               |
|-----------------|------|----------------------------|---------------|
|                 | 面试类型 | 面试流程                       | 备注（侧重点）       |
| 第一面             | 技术面  | 自我介绍+项目/实习经验<br>+技术问答+算法编程 | /             |
| 第二面             | 技术面  | 自我介绍+项目/实习经验<br>+技术问答      | 处理问题的综合能力和压力面 |
| 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 CVTE 面试心得汇总

- ★ 喜欢从传统的视觉讲起，所以需要将传统的图像算法都进行掌握
- ★ 技术面问的都比较基础，基本稳绕简历发问，全程没有让写编程题，没有问基础数据结构。HR 面问题也相对温和，没有问一些极端的问题，感觉不错。
- ★ 一面的小哥对各种算法的底层数学问的超详细，基本上把机器学习和 CV 的知识问了个遍，好多问题都没答上来。问了很多个问题，考察了很多方面，比如你对前沿研究方向的把握、论文阅读能力等等。

## 2 CVTE 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

- 传统的图像特征有哪些？
- Sift 特征为什么能实现尺度不变性？
- Hough 直线检测的原理？
- Hough 变换，harris 角点检测算子的原理？
- SURF 与 SIFT 的对比，为什么 SURF 计算速度比 SIFT 快？
- 给出一堆大小不一的矩形框，快速求矩形框的灰度值之和？

## 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

- CNN 的原理，卷积为何有效？
- 1\*1 卷积的作用？（面试官补充：主要是起到降维或者是获取到更多特征的作用）

#### 2.2.1.2 池化方面

- 解释一下 pooling？

#### 2.2.1.3 其他方面

- 问 BN 的作用，有哪些 BN 算法？
- 梯度消失与梯度爆炸的产生原因，如何避免？

## 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、GRU、RNN 的原理以及差别？
- LSTM 和 RNN 有什么区别，解决了什么问题？

## 2.4 深度学习：CNN&RNN 通用的问题

### 2.4.1 基础知识点

- 偏差、方差区别？模型比较复杂的时候，偏差和方差的变化：偏差变小，方差变大
- 解释一下方差？
- 数据不平衡的时候的处理方法，比如说正样本多？
- 数据增强的方法？
- 数据增强用了什么噪声，与实际噪声的区别？
- 为什么训练时 Loss 会出现 NAN？

## 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

- PCA 的实现过程，推导 PCA？
- SVD 的原理？

- 特征值和特征向量奇异值
- SVD 分解局部路径规划算法

## ② 特征选择

- 特征工程如何做的，有哪些特征，为何使用这些特征？

### 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）

- 传统的机器学习了解哪些？
- RF 和 Adaboost 的异同？
- 阐述 GBDT、xgboost、lightGBM 的异同，xgb 的优势，lgb 的优势，二者的并行如何实现？

#### ② K 近邻（KNN）

- KNN 和 K-means 的原理说一下？

#### ③ SVM（支持向量机）

- 阐述 SVM 原理，为何使用对偶性以及其对偶性原理，SVM 如何解决多分类，SVM 与 LR 对比？
- SVM 和 LR 的共同点和不同点？

### 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

- loss function 说几种？

### 2.6.2 网络优化梯度下降方面

- 梯度下降和牛顿法的区别？

### 2.6.3 正则化方面

- L1 和 L2 正则化的区别？
- L1 正则为何有稀疏性？

### 2.6.4 过拟合&欠拟合方面

- 解释一下什么叫过拟合，如何防止过拟合？过拟合的解决方案？
- 模型过拟合了怎么办?(数据增强、修改模型、L1 和 L2 正则化等)
- 假如网络不收敛，你会采取哪些措施？

## 3 CVTE 面经涉及项目知识点

第三节  
CVTE面经  
项目知识点  
(整理: 江大白)  
[www.jiangdabai.com](http://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

- 解释一下 ROI Pooling?

#### 3.1.2 超分辨

- 做过哪些传统的超分或者去噪的算法？
- 多尺度如何分别在三个图像去噪，图像超分辨率任务起作用？
- 图像去噪是去高频，超分是加高频，如何作用？

### 3.1.3 目标重识别

- 对行人重识别的损失函数了解多少？（说三元组损失，让解释一下）

## 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 HMM 隐马尔科夫模型

- HMM 与 CRF 的原理，差别，是全局最优还是局部最优以及为什么？

##### ② Word2vec

- 阐述 word2vec 的原理？
- 问 word2vec 的两种优化方法，说下分层 softmax 是怎么做的？
- word2vec 的优点和缺点，是如何解决 oov 的问题的？

##### ③ NER（命名实体识别）

- 做 NER 和 RE 时采用了什么模型？效果目前怎样？
- RE 时的数据标注如何进行的？

##### ④ 其他

- 对于一个取值较多的类别变量在不能进行 onehot 的情况下如何使用？
- 知识图谱的构建过程是怎样的？
- 图谱推理目前采用了什么方法？
- 表示学习如何使用在 NLP 任务中，举例说明？
- 做文本表征特征的方法有什么？

### 3.3 强化学习

无

### 3.4 机器学习方面

无

## 4 数据结构与算法分析相关知识点

第四节  
CVTE面经  
数据结构与算法分析  
(整理: 江大白)  
[www.jiangdabai.com](http://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 树

- 为何使用哈弗曼树，如何构建哈夫曼，负采样的原理？

### 4.2 算法思想实战及智力题

#### 4.2.1 算法思想实战

- 有一个城市名称列表，如何判断语句中是否出现了列表中的城市，一开始说了一种最简单的算法，但是时间复杂度很高，在面试官提示下提出用 kmp 算法来降低复杂度

### 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  
CVTE面经  
编程高频问题  
(整理: 江大白)  
[www.jiangdabai.com](http://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 网络框架方面

- 怎么处理大批量数据的读入？（基于 pytorch 或 tensorflow 回答）

#### 5.1.2 基础知识

- Python 数据类型中，不可变的有哪些，可变的有哪些？（不可变的是元组，可变的是字典、列表）

### 5.2 C/C++方面

- list 内存分配虚函数与纯虚函数的区别实现一个 share 类 lambda 函数

## 6 操作系统高频问题：数据库&线程&常用命令等

第六节  
CVTE面经  
操作系统高频问题  
(整理: 江大白)  
[www.jiangdabai.com](http://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

- 动态窗 Linux 查找文件的命令有哪些 Linux 两台主机进行通信采用什么命令 Linux 进错目录?
- 用 cd 命令如何返回 vector?

## 7 技术&产品&开放性问题

### 7.1 技术方面

- 单模型最高与模型融合的结果差距多少? 如何进行的融合?
- 无人机竞赛中所使用的路径规划算法是什么? 是如何对天气进行预测的? 比赛中对上分的手段是什么?
- 有一个城市名称列表, 如何判断语句中是否出现了列表中的城市, 一开始说了一种最简单的算法, 但是时间复杂度很高, 在面试官提示下提出用 kmp 算法来降低复杂度, 面试官接着问有没有更好的方法, 我已经想不出来了, 面试官就给我介绍前缀树的做法。

### 7.2 开放性问题

- "记录微波炉里烘培食物的过程, 自动选出最有趣的一段视频" 如果让我去实现这个项目, 我会怎么做?

# 30| 大华算法岗武功秘籍

## 1 大华面经汇总资料

第一节  
大华面经  
汇总资料  
(整理: 江大白)  
[www.jiangdabai.com](http://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 大华面经整理心得

### 1.1 面经汇总参考资料

#### ① 参考资料:

- (1) 牛客网: 大华面经-25 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&答案&目录&心得:

- (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

【图像算法工程师】、【算法应用工程师】、【感知算法工程师】、【算法移植优化工程师】、【机器人算法工程师】

## 1.3 面试流程时间安排

| 大华面试流程-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<br>+技术问答+算法编程 | /       |
| 第二面           | 技术面  | 自我介绍+项目/实习经验<br>+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：**

- 有时候第一面是电话面

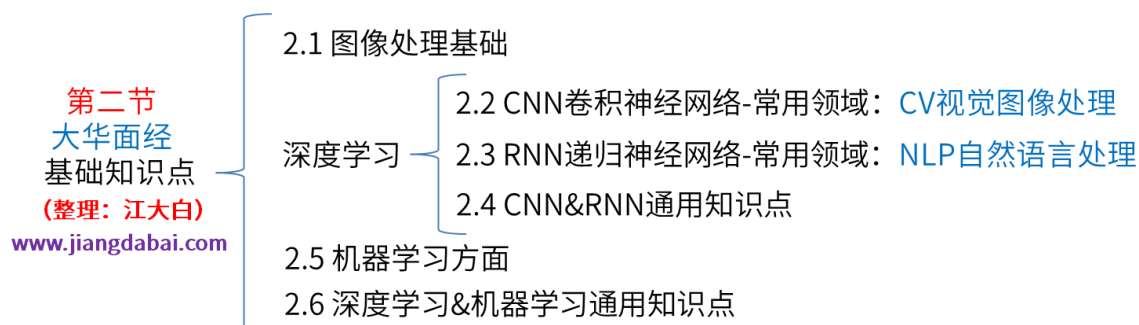
## 1.4 大华面试心得汇总

★ 会问些项目，还有一些基础的细节，问的特别细的那种。比如图像的某些细节，感觉传统的图像处理也会用到。

★ 算法逻辑问题很少问。传统的和 CV 的都会问。

★ 大华的面试官都挺好的，问题也很有深度，当时很多都没答上来，特别是二面很多数学问题，数学渣渣瑟瑟发抖，没想到最终还是给了 HR 面。

## 2 大华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

- 常用颜色空间, 色彩空间以及色彩直方图介绍一下?
- 传统的图像特征提取算法了解吗? SIFT 特征提取方法也讲解一下?
- 传统的图像特征, 鞍点是啥?
- 传统边缘检测算子有哪些? Hough 直线检测原理讲一下?
- KCF 介绍、缺点、MOSSE、CSK 、KCF 的区别?
- 介绍下 HOG, HOG 梯度直方图如何建立?
- 光流法讲解一下?
- 双线性插值算法原理讲一下?
- Gabor 滤波器和 LBP 的原理知道吗?
- 高斯滤波的原理是什么?
- 说出图像变化的矩阵公式?
- 仿射变换有几个自由度? 给 5 个点坐标, 如何求解仿射变换的参数?
- 图像质量评价指标
- 色彩空间

- 去噪方法
- BM3D 的原理和优点
- 传统图像处理了解吗？检测直线和圆可以用哪些算法？

### 2.1.2 手写算法代码

- 手写灰度直方图代码？
- 手撕图像读取，crop，缩放代码？

## 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

- 卷积层原理、反向传播的计算过程？
- $1 \times 1$  卷积的作用？
- 深度可分离卷积的作用？
- 卷积优化的操作有哪些？
- dropout 是伯努利还是二项分布？
- 介绍一下深度可分卷积？
- 深度可分卷积和传统卷积计算量和参数量的对比？

#### 2.2.1.2 池化方面

- Pooling 层原理、反向传播的计算过程？
- 最大池化的反向传播是怎么做到的？

#### 2.2.1.3 网络结构方面

- ResNet 用过吗？Resnet 的各个版本介绍一下？

- Inception 的各个版本及其改进?
- Mobilenet 的三个版本的区别?
- MobileNetv1、v2 在 imagenet 上复现?
- MobileNetv1 和 v2 的区别, relu6 为什么低精度好?
- 轻量级网络还有哪些?
- 知道哪些深度学习的网络结构 (ResNet、DenseNet、VGG、Inception、MobileNet、Shufflenet)
- Inception 的四个版本
- MobileNet 的三个版本
- 画一下 shufflenetv2 的结构, 解释作用

#### 2.2.1.4 其他方面

- BN 层原理、作用、参数介绍? BN 层测试时怎么办?
- BN 中的参数有几个, 具体是多少 (NCHW) ?
- 介绍一下 Batchnorm?  $\gamma$  和  $\beta$  有什么用?
- 梯度消失和梯度爆炸的原因、解决方法?
- 训练和测试阶段不一样的操作有哪些? (比如 dropout)
- finetune 的原因, 原理, 步骤?
- CBAM 在实际分类中和 SENet 表现差不多, 你觉得可能的原因?

#### 2.2.2 激活函数类

- Softmax 损失公式、具体计算过程推导?

## 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 扩充细节？

- LSTM 与 RNN 的区别，为什么不用 GRU？

(1) GRU 和 LSTM 的性能在很多任务上不分伯仲。

(2) GRU 参数更少因此更容易收敛，但是数据集很大的情况下，LSTM 表达性能更好。

(3) 从结构上来说，GRU 只有两个门（update 和 reset），LSTM 有三个门（forget, input, output），GRU 直接将 hidden state 传给下一个单元，而 LSTM 则用 memory cell 把 hidden state 包装起来）

## 2.4 深度学习：CNN&RNN 通用的问题

### 2.4.1 基础知识点

- 神经网络初始化方法有哪些些？

- 说一下 kaiming 初始化要解决什么问题？

- 了解 SeNet 吗，channel attention 怎么做的？

### 2.4.2 模型评价

- 如何评价多个模型的性能的好坏，ROC 曲线代表了什么？

- auc 的定义？

- P、R 和 AP 有什么关系？



## 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

- CCA 和 PCA 的区别?
- 介绍一下 PCA?
- PCA 和 LDA 的原理?

##### 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）

- 说一些有哪些分类器? (说了 SVM、LR), 用两分钟介绍一下 SVM
- 传统机器学习, Xgboost 和 GBDT 的区别?

##### ② SVM（支持向量机）

- SVM 和 Adaboost 的流程讲一下?
- SVM 的原理说一下?

#### 2.5.1.4 无监督学习-聚类方面

- 讲一下 kmeans 算法? 说一下初始聚类中心的选择有没有什么优化的方法?
- 聚类算法有哪些?

## 2.5.2 手推算法及代码

- 手撕 svd，分解的三个矩阵分别代表了什么？迭代法推到，最小二乘估计。

## 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

- 常用的损失函数？
- Loss 的设计规则？
- Smooth L1 为什么小于 0.5 的时候不直接用 L1？

### 2.6.2 激活函数方面

- 知道哪些激活函数？
- ReLU 为什么可以防止梯度消失？
- Sigmoid 不好在哪？

### 2.6.3 网络优化梯度下降方面

- 常用的网络优化方法有哪些？
- Adam 的缺点是什么？
- 手撕牛顿法和梯度下降法，数学推导及区别？
- deep learning 中梯度优化的几种经典优化方法的数学推导
- 了解过优化算法吗？说说二阶梯度优化方法？（经常问）

### 2.6.4 压缩&剪枝&量化&加速

- 了解过压缩和量化吗？
- NCNN 量化出问题怎么解决？
- 剪枝有哪些方法？

- 介绍一下模型压缩？有哪些方法？量化之后精度掉了多少，有什么改进的方法？

## 2.6.5 过拟合&欠拟合方面

- 过拟合解决的方法？

## 2.6.6 其他方面

- em 算法 e 步和 m 步是什么？
- 马氏距离是什么？介绍一下？
- 快速傅里叶变换和离散傅里叶变换的区别？
- 512 个特征点是怎么比较的？用什么方法？（余弦相似度）

# 3 大华面经涉及项目知识点

第三节  
大华面经  
项目知识点  
(整理: 江大白)  
[www.jiangdabai.com](http://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

- YOLOv1 YOLOv2 YOLOv3 YOLOv4 的发展？
- Yolo anchor 的选取讲一下？（kmeans）
- NMS 的变种说一下？
- RPN 怎么计算 box 的实际坐标？
- SSD 算法的难负样本挖掘，产生的 anchor boxes 对应原图还是特征图？

- Faster RCNN 在经过 RPN 后，产生的 anchor boxes 数量，产生的感兴趣区域对应原图还是特征图？
- Faster RCNN 的损失讲一下。softmax 损失函数的表达式 (softmax + 交叉熵损失)，Smooth L1 损失表达式，为什么好？
- Anchor Free 的目标检测了解吗？说一下
- RetinaFace 了解吗？说一下。它的损失函数是什么？
- 如何加强小目标检测？答：多尺度、fpn
- 说一下检测中是怎么处理正负样本不平衡的？

### 3.1.1.2 损失函数

- Focal loss 讲一下？alpha 的作用是什么？

### 3.1.1.3 手写代码

- 非极大值抑制代码写一下？

## 3.1.2 目标追踪

- 基于神经网络的跟踪算法了解吗？感觉想问 centertrack 之类的？

## 3.1.3 图像分割

- 了解 Mask R-CNN 吗？说一下损失函数？smooth L1 loss 的函数表达式？

## 3.1.4 关键点检测

- 人脸关键点从一幅图到另一副图空间转换算法手撕，4 个未知数 5 个方程怎么求解手撕？

## 3.1.5 目标重识别

- MAP 讲一下，目标检测的 MAP 和 ReID 中的 MAP 的区别，ReID 中 MAP 怎么计算？

## 3.2 深度学习：RNN 递归神经网络方面

无

## 3.3 强化学习

无

## 3.4 机器学习方面

无

# 4 数据结构与算法分析相关知识点

第四节  
大华面经  
数据结构与算法分析  
(整理: 江大白)  
[www.jiangdabai.com](http://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

- 找出无限大数组中第 K 个

## 4.2 算法思想实战及智力题

无

## 4.3 其他方面

### 4.3.1 数论

- 雅可比矩阵了解过吗？
- KL 散度的定义，度量分布相似度的缺点是什么？
- 线性回归的求解方式？
- 方差协方差的定义？
- 为什么顺着梯度的负方向更新参数下降最快？

## 4.4 Leetcode&剑指 offer 原题

无

## 5 编程高频问题：Python&C/C++方面

第五节  
大华面经  
编程高频问题  
(整理: 江大白)  
[www.jiangdabai.com](http://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 网络框架方面

##### 5.1.1.1 Pytorch 相关

- 深度学习把很多操作封装成独立的层，那为什么 pytorch 中把 softmax 和 loss 层封装到一起，原因是什么？
- pytorch 转 openvino 为什么精度降低了，问题在哪？

### 5.1.1.2 Tensorflow 相关

- 看过 pytorch, tensorflow 源码吗? 智能指针讲一下

### 5.1.1.3 Caffe 相关

- Caffe 和 tensorflow 的区别?

### 5.1.1.4 其他

- 平时用的框架, 这些深度学习框架之间的差别?
- Pytorch 和 tensorflow 的区别, 了解这些深度学习框架的底层区别吗? 用深度框架自己写过层吗?
- openvino 为什么快? 原理?

## 5.1.2 基础知识

### 5.1.2.1 线程相关

- python 多线程?

### 5.1.2.2 内存相关

- python 深浅拷贝?

### 5.1.2.3 区别比较

- 说一下 python list 和 tuple 的区别? python 的 list 是可迭代的嘛?
- python 中的 “==” 和 is 有什么区别?
- 集合、列表、元组的区别?
- 命名变量, 首字母变成下划线有什么用?
- python list 和 tuple

### 5.1.2.4 讲解原理

- python 列表和元组、列表去重?

- pass 语句的作用？
- with 关键字有什么作用？
- python 对列表去重？
- python 全局解释器锁？
- python 怎么定义一个类？类的成员变量和对象的成员变量有什么区别？
- 修饰符有什么作用？
- Static 的用法？

#### 5.1.2.5 讲解应用

- 如何遍历字典的所有键和值？
- python 进程慢的原因？
- list 怎么去重

## 5.2 C/C++方面

### 5.2.1 基础知识

#### 5.2.1.1 内存相关

- 介绍一下 static 静态变量，怎么存放？
- 主要考察了 static 变量的性质，在函数内定义，不随着调用结束而销毁，C++内存模型。

#### 5.2.1.2 区别比较

- C++指针和引用？
- C++中指针和引用的区别是什么？指针的 sizeof 是多少？
- float 在 32 和 64 上的差距
- sizeof 和 strlen 差别
- C++怎么样？new 和 malloc 的区别？



### 5.2.1.3 讲解原理

- 类和对象有什么关系？
- 单例模式和工厂模式知道吗？
- 判断 struct 的字节数？
- static 关键字作用？
- static 加到变量和函数上的作用？
- 局部变量的作用域？
- 底层 const 的具体含义？
- 在 C 语言中 static 的作用？
- 浮点数怎么比较相等（精度）

### 5.2.1.4 讲解应用

- cuda 如何确定 block 及 thread 等
- float 单精度 6.5 在机器里怎么存的？

## 5.2.2 手写代码相关

- 读程序题两道，一个是考无符号整型，一个是考 switch case 语句+break
- 实现均值滤波

# 6 操作系统高频问题：数据库&线程&常用命令等

第六节  
大华面经  
操作系统高频问题  
(整理: 江大白)  
[www.jiangdabai.com](http://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

### 6.1.1 基础问题

- Mysql 中的左连接和右连接有什么区别？

## 6.2 操作系统方面

### 6.2.1 线程和进程相关

#### 6.2.2.1 讲解原理

- 深拷贝和浅拷贝的区别？

#### 6.2.2.2 讲解应用

- 进程除了使用锁之外如何同步？

### 6.2.3 常用命令

- Linux 怎么知道一个文件夹下的文件个数？
- Linux 中 top 命令有什么作用？系统之间文件传输使用什么命令？
- 修改文件权限：chmod +x file
- 软硬链接的区别
- 查看内存的命令
- 查看一个文件夹中大于 3 天的所有文件（子文件下的也要）
- Linux 怎么查看隐藏文件？怎么统计一个文件夹下文件数量？

### 6.2.4 其他问题

- ipc 是什么缩写？

## 7 技术&产品&开放性问题

### 7.1 技术方面

- 如果模型再 rgb 上训练，给你一张 bgr 图片，在不改变输入图片通道顺序的情况下，如何使用这个模型？（不能重训，可以改模型）
- 如何降低误报率（举例：车辆违规，如果一直误报交警不满意）？
- 如何求灰度图的中值？

# 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](http://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&答案&目录&心得:

- (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<br>+技术问答+算法编程 | 很看重项目及知识点细节<br>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<br>+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

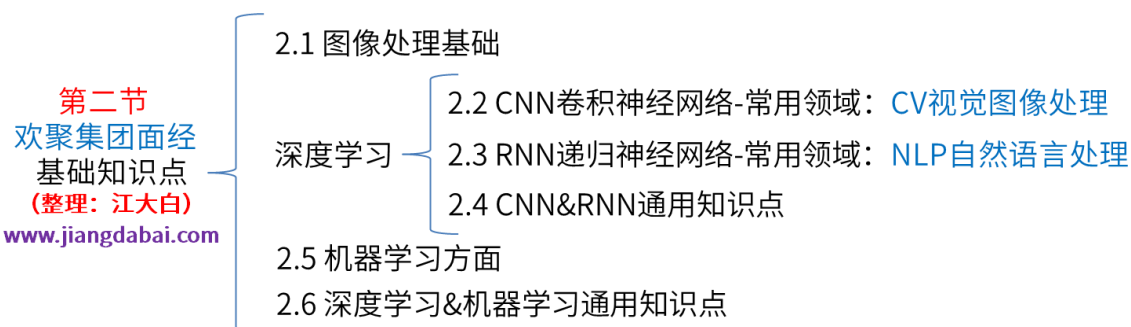
## 1.4 欢聚集团面试心得汇总

★ 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？它里面帝立刻雷分布怎么应用的？
- 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](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](http://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](http://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](http://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

# 32|平安科技算法岗武功秘籍

## 1 平安科技面经汇总资料

第一节  
平安科技面经  
汇总资料  
(整理: 江大白)  
[www.jiangdabai.com](http://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 平安科技面经整理心得

### 1.1 面经汇总参考资料

#### ① 参考资料:

- (1) 牛客网: 平安科技面经-22 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&答案&目录&心得:

- (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) 实习岗位类

【深度学习算法实习生】

#### (2) 全职岗位类

【平安科技医疗组算法工程师】、【平安科技智能引擎部算法工程师】、【智慧农业算法工程师】、  
【平安产险算法工程师】、【平安人寿 AI 研发团队算法工程师】、【平安图像算法工程师】

## 1.3 面试流程时间安排

| 平安科技面试流程整理：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<br>+技术问答+算法编程 | 问题比较广，<br>项目+基础知识点 |
| 第二面            | 技术综合面 | 自我介绍+项目/实习经验<br>+技术问答+算法编程 | 会从项目延伸到基础知识        |
| 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 平安科技面试心得汇总

- ★ 不一定每一轮都有算法逻辑题
- ★ 感觉还是要多关注基础，面试时在编程能力和数学理论上起码要有一项说得过去。
- ★ 感觉会问基础的知识点，但主要是这些，不会进行很多的发散，延伸，所以最好所有知识点抖知道
- ★ 整体感觉：问得问题中规中矩、一般是人工智能的基础+项目深挖+少数算法题目

## 2 平安科技面经涉及基础知识点

第二节  
平安科技面经  
基础知识点  
(整理：江大白)  
[www.jiangdabai.com](http://www.jiangdabai.com)

2.1 图像处理基础

深度学习

2.2 CNN卷积神经网络-常用领域：CV视觉图像处理

2.3 RNN递归神经网络-常用领域：NLP自然语言处理

2.4 CNN&RNN通用知识点

2.5 机器学习方面

2.6 深度学习&机器学习通用知识点

## 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

- 为什么要用  $1 \times 1$  卷积？讲一下作用？
- Inception V3 的  $1 \times 1$  卷积和 ResNet50 的  $1 \times 1$  卷积有什么不同？
- $1 \times 1$  卷积在 Resnet 中的几种作用？
- Dropout 原理

#### 2.2.1.2 池化方面

- Pooling 的作用？
- 池化层如何反向传播？
- 上采样有哪些方式？

#### 2.2.1.3 网络结构方面

- 介绍一下 Inception V3 和 ResNet50？
- Inception V3 基于之前的 Inception 系列最大的改进是啥？
- ResNet 的特点有哪些？（答跳跃连接和瓶颈层）
- Resnet 为啥能解决梯度消失，怎么做的，能推导吗？
- VGG 网络讲讲怎么做的？
- 讲一下 AlexNet 和 VGG 的区别？



#### 2.2.1.4 其他方面

- 梯度爆炸，梯度消失，梯度弥散是什么，为什么会出现这种情况以及处理办法？
- 图像的相似性测度有哪些？

#### 2.2.2 数学计算

- 卷积核的参数计算

#### 2.2.3 激活函数类

- Softmax 的原理了解吗？

### 2.3 深度学习：RNN 递归神经网络方面

#### 2.3.1 讲解相关原理

- RNN，LSTM 推导

### 2.4 深度学习：CNN&RNN 通用的问题

#### 2.4.1 基础知识点

- 讲一下 attention 机制，有哪几种 attention 机制？

### 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

- 降维方法，PCA 降维的具体步骤？

#### 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）

- 讲一下集成方法，Bagging, Boosting, Stacking?

##### A.基于 bagging: 随机森林

- 讲一下随机森林(RF)的原理（集成学习里面的 Bagging）、随机森林怎么调参？
- 树的深度一般是多少？怎么来决定树的深度。

##### B.基于 boosting: Adaboost、GDBT、XGBoost

- GBDT, XGBOOST 的原理讲一下？

#### ② 逻辑回归 LR

- LR 推导？
- 介绍一下 LR 及其损失函数？

#### ③ SVM（支持向量机）

- SVM 所有核函数的了解应用，损失函数？
- SVM 为什么用 hinge loss, 解释 hinge loss, 讲一讲 SVM 的损失函数？为什么要用 Hinge Loss?
- SVM 如何解决非线性分类？知道哪些核函数，说一下？
- SVM 核函数性质

#### ④ 朴素贝叶斯（Naive Bayes）

- 朴素贝叶斯基本原理和预测过程？

#### 2.5.1.4 无监督学习-聚类方面

- 讲一下 K-means 的原理？

#### 2.5.2 手推算法及代码

- 写公式并说一下 KNN、kmeans？
- 写一下 LR 和 SVM 的损失函数
- SVM 推导

### 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

- 交叉熵的公式写一下？
- 损失函数为什么用 `binary_cross_entropy` 不用 MSE？
- 知道哪些机器学习常用的损失函数？结合具体的机器学习算法来讲。(回答了分类：交叉熵损失，合页损失函数，回归：L1 和 L2 损失，还说了一个 Focal Loss)

#### 2.6.2 网络优化梯度下降方面

- 讲一下梯度下降，原理？

#### 2.6.3 正则化方面

- 正则化和未正则化有什么区别？未正则化对  $w$  和  $b$  的影响？
- L1, L2 正则化的原理？

#### 2.6.4 压缩&剪枝&量化&加速

- 了解过模型压缩方面的知识吗？

## 2.6.5 过拟合&欠拟合方面

- 过拟合欠拟合为啥？ 怎么解决过拟合。

## 2.6.6 其他方面

- 类别不平衡怎么解决的？（从数据层面和算法层面来回答）

# 3 平安科技面经涉及项目知识点

第三节  
平安科技面经  
项目知识点  
(整理: 江大白)  
[www.jiangdabai.com](http://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

- 介绍一下瓶颈层？为什么要用瓶颈层？

#### 3.1.1.2 损失函数

- 用过 Focal Loss 吗？讲一下 原理。
- L1 和 L2 损失在目标检测中的应用知道吗？

#### 3.1.1.3 手写代码

- 写非极大值抑制 NMS，用 python 或者 c++都行。

### 3.1.2 图像分割

- 知道 U-Net 吗？

### 3.1.3 图像分类

- 在分类任务中会用到哪些 loss function?
- 为什么分类用交叉熵而不用 MSE?

答：用 MSE 计算 loss 的问题在于，通过 Softmax 输出的曲线是波动的，有很多局部的极值点，即非凸优化问题 (non-convex)。

而 cross entropy 计算 loss，则依旧是一个凸优化问题，用梯度下降求解时，凸优化问题有很好的收敛特性。

- 如果分类的数据图像每一类只有几张，你会用什么方法？（少样本训练，用知识迁移/词袋模型/数据增强）

## 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 Bert

- bert， bert 的各种变形，bert 结合知识图谱，bert 结合多模态， bert 压缩， bert 优缺点？
- 为什么项目中采用了 albert，不是用 bert， albert 怎么实现降低参数量，除了 Albert 还了解其他根据 bert 改进的模型吗
- bert 的效果为什么那么好？（mask 机制）
- transformer 跟 bert 在 position embedding 上有什么区别？

#### ② Word2vec

- Word2vec 实施过程，遇到的困难之类的？
- word2vec 和 glove 的原理讲一下？

#### ③ 其他

- 了解 textCNN 吗，为什么可以在 NLP 领域取得成功？

## 3.3 强化学习

无

## 3.4 机器学习方面

### 3.4.1 推荐系统

- CRF 和 HMM 的原理讲一下？
- CRF 跟 HMM 的区别，CRF 的损失函数是什么？

## 4 数据结构与算法分析相关知识点

第四节  
平安科技面经  
数据结构与算法分析  
(整理: 江大白)  
[www.jiangdabai.com](http://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

- 给定一个数组，找出数量大于数组长度的  $1/2$  的数字，不能用哈希表
- 无序数组怎么在最短时间内求出和为目标值的两个数？

### 4.1.2 排序

- 快排、快速排序的思想时间复杂度？

## 4.2 算法思想实战及智力题

### 4.2.1 算法思想实战

● 字典里面的 key 是长度为 2~5 的字符串，另外一个字符串从中间去掉连续的字符，最多去掉 6 个，剩下的字符串连接在字典中和 key 匹配，得到匹配的字符串的 value？

● 求连续子数组最大和，

输入：[1, 4, -5, 9, 8, 3, -6]

输出：9, 8, 3

● 众数，不能用哈希表和排序

● 最长连续子序列

### 4.2.2 智力题

● 你有三个朋友在深圳，你想知道深圳的天气，你给每个朋友都打了电话。

● 三个人都说现在正在下雨，三个人说谎的概率都是  $1/3$ ，问深圳下雨的概率是多大？

## 4.3 其他方面

### 4.3.1 矩阵运算

● 给一个矩阵，问从左上角到右下角有多少种走法？

### 4.3.2 其他

● 简历里有写到数学基础，所以面试官让我说一下怎么自己实现两个矩阵相乘的程序；然后问了我对于刚获得的数据，如何判断其分布，这时候我用了正态分布举例，所以面试官接着问，为何实际生活中大部分数据服从正态分布或者拟正态分布；在知道数据的分布之后，如何进行后续处理？

## 4.4 Leetcode&剑指 offer 原题

- Leetcode：螺旋矩阵

## 5 编程高频问题：Python&C/C++方面

第五节  
平安科技面经  
编程高频问题  
(整理：江大白)  
[www.jiangdabai.com](http://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

- Python 用过哪些库或者框架？
- 序列去重怎么做？
- 讲一下 apply 函数的作用？

#### 5.1.2 手写代码相关

- numpy 问题：怎么取出一个矩阵中某一行的最大值(说出具体的函数)

### 5.2 C/C++方面

#### 5.2.1 基础知识

- 简历上写了 C，问了智能指针、Class 和 struct 的区别？



## 6 操作系统高频问题：数据库&线程&常用命令等

第六节  
平安科技面经  
操作系统高频问题  
(整理: 江大白)  
[www.jiangdabai.com](http://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 TCP 协议相关

- 三次握手的原理

#### 6.2.2 线程和进程相关

- 多线程的原理讲一下？

## 7 技术&产品&开放性问题

### 7.1 技术方面

- 知道奥卡姆剃刀准则吗？说说看。
- 场景题：假设有 10000 个样本，只有 1000 个样本有详细的框框出了目标的位置，另外 9000 张没有框，只是有类别标签而已，问怎么训练这个网络？

# 33|大疆算法岗武功秘籍

## 1 大疆面经汇总资料

第一节  
大疆面经  
汇总资料  
(整理: 江大白)  
[www.jiangdabai.com](http://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 大疆面经整理心得

### 1.1 面经汇总参考资料

#### ① 参考资料:

- (1) 牛客网: 大疆面经-20 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&答案&目录&心得:

- (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

【多源融合算法工程师】、【机器学习算法工程师】、【机器人算法工程师】、【传感器算法工程师】、  
【感知算法工程师】、【视频编码解码算法工程师】

## 1.3 面试流程时间安排

| 大疆面试流程-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<br>+技术问答+算法编程 | 项目+基础知识            |
| 第二面           | 技术面  | 自我介绍+项目/实习经验<br>+技术问答+算法编程 | 会从工程化的角度，<br>聊项目经验 |
| 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。

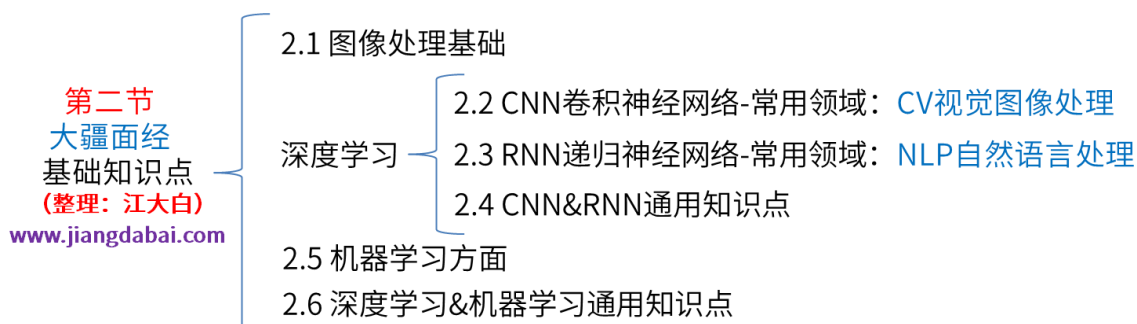
**其他注意点：**

- 第一面之前，有的人可能会有笔试

## 1.4 大疆面试心得汇总

★ 在笔试题目中，单选题其实就是靠你有无相关的知识背景，并不难，大疆感觉是在做移动机器人，像我因为概率论好久没有看了、凸优化那块也不是很懂，所以有一些是懵的。

## 2 大疆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

- 腐蚀膨胀的原理讲一下？
- 摩尔纹讲一下？
- 相机畸变的原因？
- 图像存储格式知道吗？
- 图像去噪的原理？

## 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

- 笔试多选题：增大感受野的方法:dilated? pooling? 网络&实际感受野不同? 增加卷积层的channel 能扩大感受野?
- 笔试简答题：保持网络层数不变，写出增加感受野的方法

#### 2.2.2 数学计算

- 笔试简答题：给了原图  $L \times L$ ，输入通道数，步长  $s$ ，padding，卷积大小  $k \times k$ ，输出通道数，求特征图大小？求乘法次数？

## 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 模型评价

- ROC 曲线 & AUC 值的原理讲一下？

## 2.5 传统机器学习方面

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

- Loss function 为什么选用 CC 和 LCC, 和 MSE 的区别在哪儿？CC 是否能够实现明度鲁棒？

### 2.6.2 激活函数方面

- 笔试多选题：激活函数公式判断 tant, relu, leakyrelu

### 2.6.3 其他方面

- 笔试简答题：处理训练数据集中类别不均衡问题？

## 3 大疆面经涉及项目知识点

第三节  
大疆面经  
项目知识点  
(整理：江大白)  
[www.jiangdabai.com](http://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

- Yolov3 的优缺点，相对于 ssd 的优点？
- Fpn 为什么要在多个尺度做 bounding box 回归，而不是只在 P2 这一个尺度上？
- 影响检测网络前向运行时间的因素？

### 3.1.2 目标追踪

- Siamese rpn 网络的优缺点？（和自己的研究相关）

### 3.1.3 多源融合算法

- 90 分钟：设计一个基于 EKF 的卫星与惯导组合导航滤波算法≈ 写一篇小论文
- 状态方程与观测方程的推导
- 系数矩阵的推导
- 观测量初始值如何设置
- 传感器校准算法
- 整体算法测试方案
- 滤波算法推导

### 3.1.4 机器人算法

- 笔试单选题：

(1) 机器人的坐标变换

(2) 路径和轨迹规划：一条从 A 到 B 的轨迹  $p(t)$ ，从 A 到 B 所需时间是 T，现在要在 1s 复现原轨迹，对应的速度和加速度与原先的速度与加速度的关系？若从 B 沿着之前的轨迹返回至 A，此时对应的速度与加速度与来时的速度与加速度之间的关系？还有考察 dijkstra 算法的实现细节；

(3) 概率论，状态估计，粒子滤波还是凯尔曼滤波？

(4) 凸优化

(5) 自动控制原理，传递函数的稳定性

笔试简答题：有经历过哪些不好的用户体验，怎么优化？

### 3.1.5 传感器算法

- 最优化算法，鲁棒优化的知识点
- 中间有问到我实际使用过哪些传感器？了解激光雷达吗？

### 3.1.6 感知算法

- 多线程内存安全
- 互斥锁
- 病态矩阵的判断方法
- 10 人按照生日排序，要求时间复杂度  $O(n)$
- 优化方法 LM 以及其他的优化方法顺带讲一下
- BA 解释
- ceres 的 autodiff 的原理
- C++11 特性 比较好的哪些
- shared\_ptr 和 unique\_ptr 作用与区别？

### 3.1.7 视频编码解码算法

- 硬件编码器的流水线设计，目前项目哪些地方有难度？
- 如果  $n$  级为重构， $n-1$  级位预测，怎样实现？
- 流水线设计需要考虑到什么问题？有哪些数据依赖？
- 如何提高编码速度？（大意就是用并行计算）
- 无人机的拍摄视频场景和相机拍摄场景的区别？
- 无人机拍摄需要考虑哪些视频编码性能？利用什么可以改善无人机视频编码？

## 3.2 深度学习：RNN 递归神经网络方面

无

## 3.3 强化学习

无

## 3.4 机器学习方面

无

# 4 数据结构与算法分析相关知识点

第四节  
大疆面经  
数据结构与算法分析  
(整理: 江大白)  
[www.jiangdabai.com](http://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 树

- 二叉树讲一下？
- 红黑树的原理？

## 4.2 算法思想实战及智力题

### 4.2.1 算法思想实战

- 笔试编程题：给出圆的半径、中心点坐标，矩形左下角和右上角坐标，判断是否相交？



## 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  
大疆面经  
编程高频问题  
(整理: 江大白)  
[www.jiangdabai.com](http://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

无

## 6 操作系统高频问题：数据库&线程&常用命令等

第六节  
大疆面经  
操作系统高频问题  
(整理: 江大白)  
[www.jiangdabai.com](http://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 线程和进程相关

#### 6.2.1.1 区别比较

- 线程、进程的原理？

#### 6.2.1.2 讲解原理

- 操作系统低调度的最小单位？

## 7 技术&产品&开放性问题

### 7.1 技术方面

- 如何对没见过数据有一个比较好的泛化能力？

### 7.2 产品方面

- 笔试简答题：有经历过哪些不好的用户体验，怎么优化？

## 34|蘑菇街算法岗武功秘籍

### 1 蘑菇街面经汇总资料

第一节  
蘑菇街面经  
汇总资料  
(整理: 江大白)  
[www.jiangdabai.com](http://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 蘑菇街面经整理心得

#### 1.1 面经汇总参考资料

##### ① 参考资料:

- (1) 牛客网: 蘑菇街面经-13 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&答案&目录&心得:

- (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) 实习岗位类

【图像算法实习生】

##### (2) 全职岗位类

【应用算法工程师】、【图像算法工程师】、【推荐算法工程师】、

## 1.3 面试流程时间安排

| 蘑菇街面试流程-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<br>+技术问答+算法编程 | /       |
| 第二面            | 技术面       | 自我介绍+项目/实习经验<br>+技术问答+算法编程 | /       |
| 第三面            | 技术Leader面 | 自我介绍+项目经验+公司发展             | /       |
| 第四面            | HR面       | 基础人力问题                     | /       |

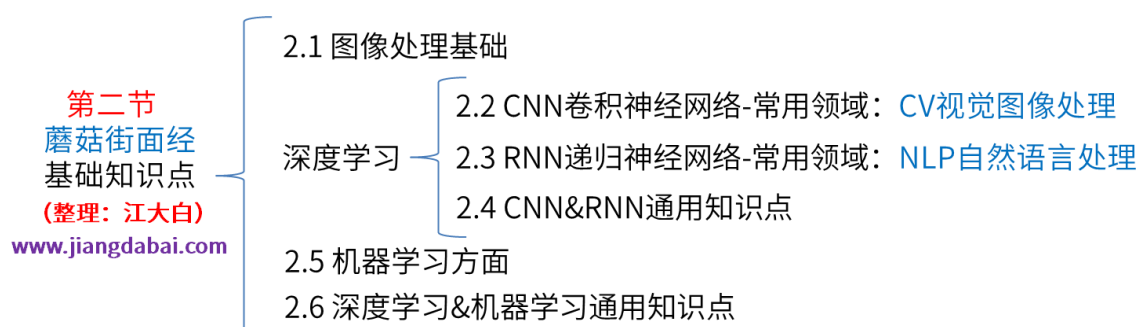
**PS：**以上流程为大白总结归纳所得，以供参考。

## 1.4 蘑菇街面试心得汇总

★ 面试官人很好，说什么都会给你回应。对项目会问很多细节。

★ 聊项目，问了类别不均衡问题，也问了模型可解释问题，好像蘑菇街喜欢可解释强的模型？  
然后当时引申的一个问题是，如果 nn 可以解决高维度离散问题，为啥还需要 gcn 来解决关联性的问题。其它不记得了。

## 2 蘑菇街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 网络结构方面

- 图像分类基础网络有何不同改进？

#### 2.2.1.2 其他方面

- BN 层的均值如何训练的？

## 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 和 RNN 有什么不同？
- LSTM 如何解决 RNN 中梯度消失的问题？
- 项目中分别使用了 LSTM 和 CNN 进行文本分类，效果比较，为什么项目中 LSTM 的分类效果好？问了 RNN、LSTM 和 CNN 使用场景？
- 什么是梯度消失？LSTM 为什么能解决 RNN 梯度消失问题？

### 2.3.2 手绘网络原理

- 手写 LSTM 门控公式？

## 2.4 深度学习：CNN&RNN 通用的问题

### 2.4.1 基础知识点

- 方差和偏差是怎么回事？

### 2.4.2 模型评价

- 知道 AUC 吗，不均衡会对 AUC 产生影响吗，为什么？

## 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

- 做过哪些特征工程，产生哪些类型特征(统计特征，交叉特征，转化特征都有什么)，异常值怎么处理，数据平滑问题？

#### 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）

- GBDT 调过哪些参数？
- 正负样本比对 lr, gbd 影响，为什么，用 lr 比例 N/P 比为 1 会怎么样？

##### ② 逻辑回归 LR

- LR 和 GBDT 有什么区别，为什么 GBDT 结果比 LR 好？
- 正负样本比对 LR, gbd 影响，为什么，用 LR 比例 N/P 比为 1 会怎么样？
- LR 原理、损失函数？
- LR 的详细过程怎么处理？
- 你说 LR 模型可解释性高，它的可解释性体现在哪里？是权重越大，可解释性就越强吗？

### ③ SVM（支持向量机）

- SVM 原理、损失函数？

### ④ 朴素贝叶斯（Naive Bayes）

- 贝叶斯公式的原理？

### ⑤ 决策树（DT）

- 你说一下信息增益的原理？

#### 2.5.1.4 无监督学习-聚类方面

- 详细说一下 K-means 聚类，你每个簇怎么选样本的，距离用什么算的？

#### 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

- 知道哪些损失函数？

### 2.6.2 网络优化梯度下降方面

- 知道熟悉什么优化算法，SGD 和 Adam 的公式是什么样的，怎么迭代计算的？
- Adam 的思想？Adam 能否适用于稀疏数据？
- SGD, Adam 等，以及他们的参数更新公式？

- 梯度下降如何跳出局部最优值？
- 如何解决陷入局部最小，自适应学习率方法？

### 2.6.3 压缩&剪枝&量化&加速

- 模型简化压缩怎么处理的？

### 2.6.4 过拟合&欠拟合方面

- 过拟合的原因？如何解决过拟合？

### 2.6.5 其他方面

- 机器学习和神经网络的选择问题？（什么情况下使用神经网络会取得较好的效果）

## 3 蘑菇街面经涉及项目知识点

第三节  
蘑菇街面经  
项目知识点  
(整理: 江大白)  
[www.jiangdabai.com](http://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

- Soft-NMS 和 NMS 的不同点？
- FPN 的原理？
- 项目细节，检测算法目前是怎么做的，跟踪问题经典的一些方法以及目前比较常用的方法，以及你的创新改进？



### 3.1.1.2 损失函数

- RetinaNet 的 loss function 是什么？
- 用到的 Focal loss 的公式，怎么用的，原理参数含义？
- Focal loss 是怎么缓解样本不均衡问题的，还有其他的方法吗？

### 3.1.1.3 手写代码

- 手写 NMS？

## 3.1.2 目标重识别

- ReID 用的网络是什么样的，损失函数是怎么构建的，公式是什么样的？

## 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 HMM 隐马尔科夫模型

- HMM 的维特比算法

#### ② Word2vec

- 介绍 Word2Vec, 我就把把 CBOW, skip-gram, hierarchical softmax, negative sampling, 介绍了一遍？

#### ③ CNN 方面

- 介绍 TextCNN？

## 3.3 强化学习

### 3.3.1 讲解原理

- 问 Q 方法，说一下强化学习中 DQN 算法？

- 强化学习的分类？
- 值函数方法有哪些？
- 随机策略相对非随机策略优缺点？
- 熟悉的策略梯度方法？
- 如何改进 ppo 算法？
- 对 V 函数和 Q 函数的看法，为什么要有 V 和 Q？

### 3.4 机器学习方面

无

## 4 数据结构与算法分析相关知识

第四节  
蘑菇街面经  
数据结构与算法分析  
(整理: 江大白)  
[www.jiangdabai.com](http://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

- 最大连续子数组和？
- 两个排序数组合并成一个数组？
- 给一个数组和一个数，求数组中有多少对和为该数的数？

##### 4.1.1.2 链表

- 实现链表反转？

- 两个单链表并集？

#### 4.1.1.3 字符串

- 大数相加，转换成字符串，对每个字符相加减，考虑进位？
- 字符串排列

## 4.2 算法思想实战及智力题

### 4.2.1 算法思想实战

- 分糖果问题？
- 10 个大文件返回词频 top10 的词？
- 统计 top100 词频？

## 4.3 其他方面

### 4.3.1 矩阵运算

- 写一下矩阵相乘代码，我有三列是数据，itemid category ctr，我想取每一类的 item 里面 ctrctr 最高的 100 个，写一下代码？

## 4.4 Leetcode&剑指 offer 原题

- Leetcode 135：分糖果问题

## 5 编程高频问题：Python&C/C++方面

第五节  
蘑菇街面经  
编程高频问题  
(整理：江大白)  
[www.jiangdabai.com](http://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

- Python 的装饰器解释一下？

## 5.2 C/C++方面

- C++动态绑定，数据结构了解不？

## 6 操作系统高频问题：数据库&线程&常用命令等

第六节  
蘑菇街面经  
操作系统高频问题  
(整理: 江大白)  
[www.jiangdabai.com](http://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

无

## 7 技术&产品&开放性问题

### 7.1 产品方面

- 场景题：比如用户点击了广告跳转进来，要展示一屏幕的物品给他，该怎么做？可以说一下整个流程吗？
- 一个商家。第一天看 A 商品和 B 商品的点击率。点击率公式是点击/曝光。然后，平均的是

a 的点击加 b 的点击/a 的曝光加 b 的曝光。问，第二天发现，a 的点击率和 b 的点击率上升了，但是平均的下降了。是什么原因，从哪些方面入手去解决这个问题？

# 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](http://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&答案&目录&心得:

- (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

| 云从科技面试流程整理：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<br>+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<br>+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。

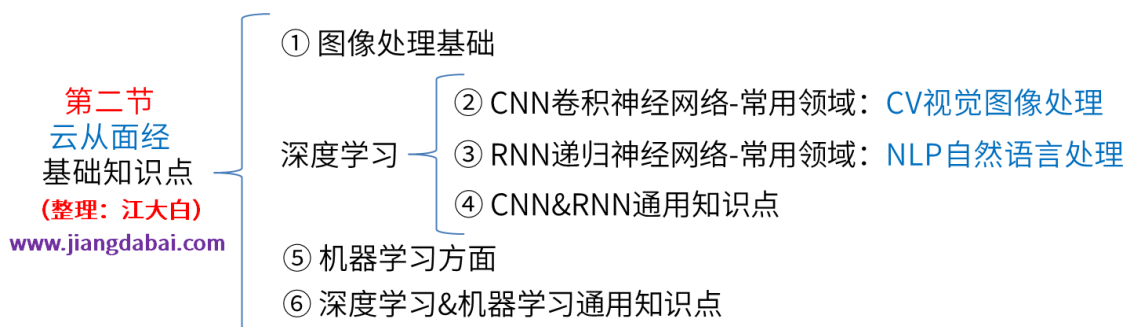
## 1.4 云从科技面试心得汇总

★ 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](http://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](http://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](http://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](http://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



# 36|追一科技算法岗武功秘籍

## 1 追一科技面经汇总资料

|                                                                                                  |                                                                |
|--------------------------------------------------------------------------------------------------|----------------------------------------------------------------|
| 第一节<br>追一科技面经<br>汇总资料<br>(整理: 江大白)<br><a href="http://www.jiangdabai.com">www.jiangdabai.com</a> | 1.1 面经汇总参考资料<br>1.2 面经涉及招聘岗位<br>1.3 面试流程时间安排<br>1.4 追一科技面经整理心得 |
|--------------------------------------------------------------------------------------------------|----------------------------------------------------------------|

### 1.1 面经汇总参考资料

#### ① 参考资料:

- (1) 牛客网: 追一科技面经-11 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&答案&目录&心得:

- (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

【图像算法工程师】、【NLP 算法工程师】、【3D 视觉算法工程师】、【机器学习算法工程师】

## 1.3 面试流程时间安排

| 追一科技面试流程整理：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<br>+技术问答+算法编程 | /       |
| 第二面            | 技术面  | 自我介绍+项目/实习经验<br>+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。

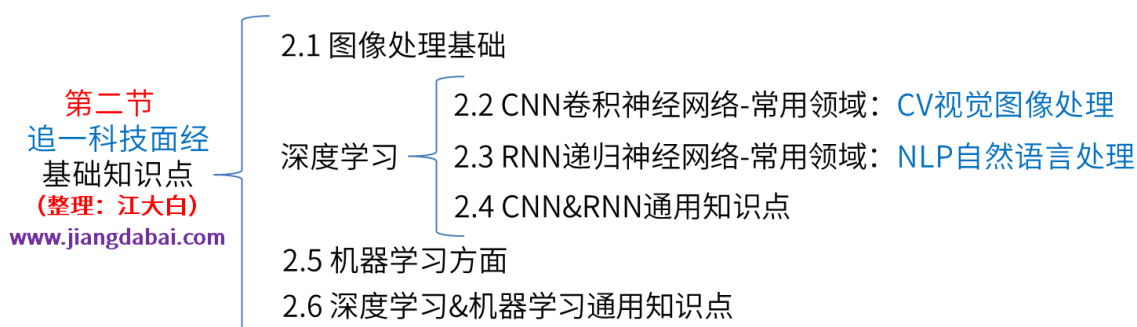
其他注意点：

- 第二面之后，有的人会有 CTO 面

## 1.4 追一科技面试心得汇总

- ★ 机器学习也会有，会从实习到比赛都问一遍，问得也会比较细，针对提出问题
- ★ 自己做的是图像,但面试官主要是做 nlp 的.所以自己觉得匹配度不是很高,但面试官人很好,一直说常用方法是差不多的,似乎没有计较这个
- ★ 简历上写的科研方向或者项目问的不多，主要针对于知识的广度与深度。

## 2 追一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

- 卡尔曼滤波讲一下？

## 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

- 卷积操作的详细介绍,以及对卷积和池化的理解？

#### 2.2.1.2 网络结构方面

- 讲一下 ResNet？

#### 2.2.1.3 其他方面

- BN 缺点，原理（参数）？

## 2.3 深度学习：RNN 递归神经网络方面

### 2.3.1 讲解相关原理

- 对 RNN 的理解？

## 2.4 深度学习：CNN&RNN 通用的问题

### 2.4.1 基础知识点

- 针对项目中样本不平衡问题，你是如何处理的？
- 分类中的数据不均衡，有什么办法可以缓解？（答了修改 loss，重采样）

## 2.4.2 模型评价

- 混淆矩阵，如何计算 recall 和 precision?
- AUC 的计算?
- AUC 跟准确度、F1 等指标相比有什么好处?

## 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

- 常用降维方法 PCA 和 SVD，重点谈了 PCA 的实现过程以及对其中协方差矩阵的理解?

##### ② 特征选择

- 如何处缺失数据?
- 特征筛选的方法有哪些?

#### 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）

- RF 跟 XGBoost 的区别?
- 谈一下 GBDT 的理解?
- RF 跟 XGBoost 的区别?

##### ② 逻辑回归 LR

- LR 原理讲一下?

- LR 和 SVM 区别（计算复杂度）？

### ③ SVM（支持向量机）

- SVM 原理及优化方法，如何体现支持向量？

### ④ 决策树（DT）

- 决策树的 ID3, C4.5, CART 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 网络优化梯度下降方面

- 常用的优化器有哪些？（adam,sgd,adagrad 等）
- 各种优化方法之间的区别？

### 2.6.2 其他方面

- 针对项目中样本不平衡问题，你是如何处理的？

## 3 追一科技面经涉及项目知识点

第三节  
追一科技面经  
项目知识点  
(整理: 江大白)  
[www.jiangdabai.com](http://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

- 常用检测网络的 one-stage 和 two-stage 方法中的网络框架简单介绍以及对比？

### 3.1.2 图像分割

- 选择用 mask-rcnn 做可行驶区域分割的原因,对实例分割的理解？(属于项目内容)
- 实例分割综述？

### 3.1.3 关键点检测

- 项目中使用的关键点检测算法介绍一下？

### 3.1.4 3D 图像

- 3D 检测方法，传感器。
- RGBD 相机原理，激光测量深度原理。点云相关技术。
- 3DMM 模型。
- 相机标定方法与流程，内外参矩阵求解。
- 相机标定精度优化。开放问题：现有标定和去畸变流程的精度估计以及如何优化标定流程减少精度损失。
- 3D 关键点检测商用设备与检测原理。如何提升 3D 关键点检测精度。

## 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

- 有没有试过 gpt、bert 这样的方法吗？

## ② CNN 方面

- 问 textCNN、transformer 结构，介绍的时候顺道说了模型的优缺点？

## ③ 其他

- 文本分类的模型了解有哪些？（textcnn,介绍 textcnn 的结构，和 rnn 的对比区别，应用场景？） 你了解的知识图谱构建有哪些过程？细节上需要解决哪些问题？（答了抽取三元组，实体链接，预测新实体，预测新关系等）
- 问了词向量的训练方式？
- 介绍 skip-gram、负采样？
- 我买了苹果，和苹果公司，苹果的一词多义如何解决？（回答了保存模型结构而且保存词向量）

## 3.3 强化学习

无

## 3.4 机器学习方面

无

## 4 数据结构与算法分析相关知识点

第四节  
追一科技面经  
数据结构与算法分析  
(整理: 江大白)  
[www.jiangdabai.com](http://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

无

## 4.2 算法思想实战及智力题

### 4.2.1 算法思想实战

- 1m 绳子剪两段，短的一段的期望？

### 4.2.2 智力题

- 一个二维矩阵，里面有数字，问从左上角走到右下角经过的数字的和最大是多少？

## 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  
追一科技面经  
编程高频问题  
(整理: 江大白)  
[www.jiangdabai.com](http://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

#### 5.2.1.1 区别比较

- 指针和引用区别，什么时候只能用指针或者引用。函数 return vector 变量的缺点，一般怎么操作？
- C++虚函数和纯虚函数？C++引用和指针？

#### 5.2.1.2 讲解原理

- C 11 的新特性？
- 多线程，线程安全，共享数据如何保证线程安全？如何进行数据共享？

## 6 操作系统高频问题：数据库&线程&常用命令等

第六节  
追一科技面经  
操作系统高频问题  
(整理: 江大白)  
[www.jiangdabai.com](http://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

- Linux 进程是怎么通信的？

## 7 技术&产品&开放性问题

### 7.1 技术方面

- 给定 100 万个 ip 段组成的数据库，每个 ip 段互相不重叠，各自对应一个城市，问给一个 ip 地址，返回对应的城市，有什么方法？
- 一些处理数据的技巧，比如用户年龄集中在一个年龄段，如何处理这种情况？

### 7.2 产品方面

- 目前有十个领域用户突然开始闲聊，如何判断是闲聊，如何生成回复 ？
- 用户输入“北京”，搜索框下的下拉框会出现多个联想词，如北京天气，北京旅游等，问：现在给你做这个下拉框，你怎么做？

# 37|爱奇艺算法岗武功秘籍

## 1 爱奇艺面经汇总资料

第一节  
爱奇艺面经  
汇总资料  
(整理: 江大白)  
[www.jiangdabai.com](http://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 爱奇艺面经整理心得

### 1.1 面经汇总参考资料

#### ① 参考资料:

- (1) 牛客网: 爱奇艺面经-13 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&答案&目录&心得:

- (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) 实习岗位类

【算法实习生】

#### (2) 全职岗位类

【机器学习算法工程师】、【广告算法工程师】

## 1.3 面试流程时间安排

| 爱奇艺面试流程 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<br>+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<br>+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：**

- 有些人两年之后才是笔试

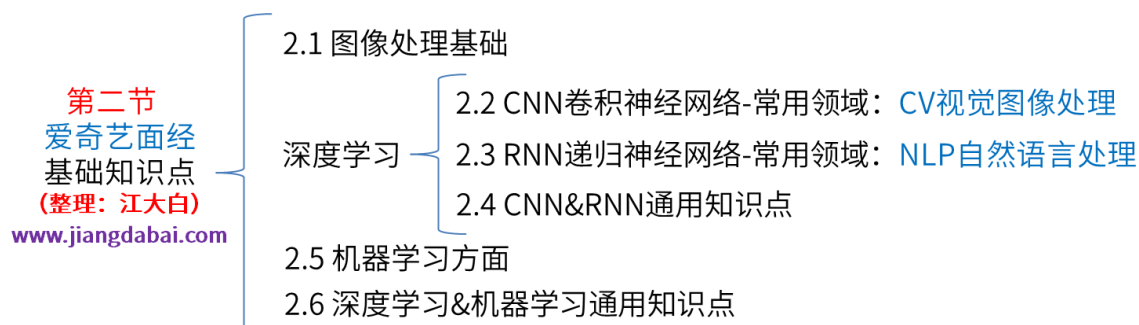
## 1.4 爱奇艺面试心得汇总

★ 以推荐算法为主，主要是做异常检测的，包括数据安全，网络安全等等，主要用的是统计机器学习偏多，也会用到深度学习，结合着具体的业务来看。

★ 感觉主要是自我介绍+项目介绍+问一下使用到的网络或者基础的细节，而且感觉经常会问到做 A/B 测试，效果如何？

★ 第一面时，职位对应的需要人的面试官，具体做事情的人，聊得很投机，对该职位具体要做些什么也说得明白，对于转领域和有一些技术没接触也表示去了之后现学是完全没问题。

## 2 爱奇艺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 为什么能够避免过拟合? 讲一下原因?
- 介绍一下 CNN 以及它的各个部件、作用?

##### 2.2.1.2 池化方面

- numpy 实现三维的 max pooling

##### 2.2.1.3 其他方面

- 梯度消失问题是怎么产生的?
- 讲讲 BN 的原理?
- 介绍一下 CNN 以及它的各个部件、作用?

## 2.3 深度学习：RNN 递归神经网络方面

### 2.3.1 讲解相关原理

- RNN, LSTM, GRU 区别与联系?
- LSTM 是如何解决 RNN 的梯度消失问题?
- LSTM, BERT, transform 中的位置编码是如何表示?
- LSTM 的 cell 中 tanh 是否可以替换成 sigmoid?

## 2.4 深度学习：CNN&RNN 通用的问题

### 2.4.1 模型评价

- 常用的模型评价指标分别说一下
- AUC 来说一下? 它表示什么意义? AUC 有哪些缺点? 有哪些前提条件?
- 项目中召回率你是怎么来进行衡量的?
- 你的目的是为了提高召回率, 有木有想着通过牺牲一定的准确率来进行提高召回率?

## 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

- 问了特征工程怎么做的, 具体的流程说一下?

##### ② 特征选择

- SVD 的原理讲一下？

### 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）

- 详细的来讲一下 GBDT 和随机森林，并且讲一下它们的区别？
- LR 和 lightgbm 的优缺点说一下？
- lightgbm 为什么不适合高维稀疏的样本？

#### ② 逻辑回归 LR

- LR 损失函数？
- LR 与 SVM 的区别？
- FM 算法了解么，跟 LR 算法相比做了哪些改进？

#### ③ SVM（支持向量机）

- SVM 的原理？
- 为什么要转化为对偶问题,为什么对偶问题就好求解，原始问题就不能求解？

#### ④ 朴素贝叶斯（Naive Bayes）

- 朴素贝叶斯的原理讲一下？
- 朴素贝叶斯，先验概率，后验概率，调整因子？

#### ⑤ 决策树（DT）

- 详细的讲一下决策树(楼主讲了 ID3、C4.5、CART，分别从特征选择、决策树生成、决策树剪枝来介绍了一波)

### 2.5.1.4 无监督学习-聚类方面

- K-means 讲一下？K 值怎么来进行确定？

## 2.5.2 手推算法及代码

- 推导下 GBDT

- 推导一下 LR

## 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

- 交叉熵的物理意义是什么？

### 2.6.2 激活函数方面

- 激活函数知道哪些？

### 2.6.3 正则化方面

- 说下正则化的常用方法？
- L1 正则化和 L2 正则化分别有什么作用？(从数学角度解释)，有什么区别？
- 讲一下 L1 正则化、L2 正则化以及哪个可以起到稀疏的作用，为什么？
- 能不能从数学的角度来讲一下 L1 正则化可以起到特征选择的作用？
- L1 正则化和 L2 正则化的先验分布分别是啥？

### 2.6.4 过拟合&欠拟合方面

- 深度学习中解决过拟合的方法一般有哪些？详细说一下

## 3 爱奇艺面经涉及项目知识点

第三节  
爱奇艺面经  
项目知识点  
(整理: 江大白)  
[www.jiangdabai.com](http://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 RCNN 的损失函数，准确说出来？

## 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 Word2vec

- 简答介绍下 word2vec 及 fasttext？
- word2vec 的缺点？
- fasttext 比 word2vec 快在哪里？是否可以无监督进行向量训练？

##### ② 其他

- skip gram 和 cbow 的优缺点？

## 3.3 强化学习

无

## 3.4 机器学习方面

### 3.4.1 推荐系统

- 场景不同下，推荐系统的关注点有什么不同？
- 协同过滤的原理讲一下？
- FM 算法？（简要补充了 FFM）

- FM 算法了解么，跟 LR 算法相比做了哪些改进？
- deep model 在推荐中的应用是否了解，讲一个？

## 4 数据结构与算法分析相关知识点

第四节  
爱奇艺面经  
数据结构与算法分析  
(整理: 江大白)  
[www.jiangdabai.com](http://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 字符串

- 判断一个字符串是不是有重复字符？

#### 4.1.2 树

##### 4.1.2.1 二叉树

- 求二叉树深度和宽度？
- 二叉树深度优先遍历(非递归)？

##### 4.1.2.2 堆

- 实现下最小堆

### 4.2 算法思想实战及智力题

#### 4.2.1 算法思想实战

- 上台阶，一步、两步、或三步，求上到 n 级台阶的可能数？

- 无限大的棋盘，随意给定两个点，起点和终点，从起点出发每一步只能按照“日”字的走法，问到达终点的最少步数？（注意两个点可能再一条直线上）

### 4.2.2 智力题

- 已知 A、B 两个人在一个圆里放硬币，谁放不下谁输，已知 A 先放，求有没有 A 必赢的策略？

## 4.3 其他方面

### 4.3.1 数论

- 已知  $x, y$  的期望和方差，且  $x, y$  独立，求  $xy$  的方差？
- 已知  $x, y$  独立同分布，求  $\max(x, y)$  的期望？

## 4.4 Leetcode&剑指 offer 原题

无

## 5 编程高频问题：Python&C/C++方面

第五节

爱奇艺面经

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 网络框架方面

- 你常用的深度学习框架有哪些？

## 5.1.2 基础知识

### 5.1.2.1 线程相关

- Python 里面的多线程有什么不一样的地方？

### 5.1.2.2 讲解原理

- 分布式处理了解么？
- 了解哪些常用数据结构？

## 5.2 C/C++方面

### 5.2.1 基础知识

- new, delete, malloc, free 的区别？

## 6 操作系统高频问题：数据库&线程&常用命令等

第六节  
爱奇艺面经  
操作系统高频问题  
(整理: 江大白)  
[www.jiangdabai.com](http://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

无

## 7 技术&产品&开放性问题

### 7.1 技术方面

- 讲一下信用卡欺诈检测的项目，怎么做优化的？
- 给定一个文件夹，文件夹里可能包含图片和子文件夹，子文件夹里还可能继续包含子文件夹和图片，问怎么获取所有图片的信息或者所有图片的路径等？

### 7.2 产品方面

- 有木有做相关的对照组实验？有上线进行 A/B test 吗？效果如何？

### 7.3 开放性问题

- 你觉得算法工程师这个职业的优势与劣势是什么？

# 38| 搜狗算法岗武功秘籍

## 1 搜狗面经汇总资料

第一节  
搜狗面经  
汇总资料  
(整理: 江大白)  
[www.jiangdabai.com](http://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 搜狗面经整理心得

### 1.1 面经汇总参考资料

#### ① 参考资料:

- (1) 牛客网: 搜狗面经-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&答案&目录&心得:

- (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) 实习岗位类

【算法实习生】

#### (2) 全职岗位类

【图像算法工程师】、【推荐算法工程师】

## 1.3 面试流程时间安排

| 搜狗面试流程-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<br>+技术问答+算法编程 | /       |
| 第二面           | 技术面  | 自我介绍+项目/实习经验<br>+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 搜狗面试心得汇总

★ 感觉搜狗问得问题不是很深入，分为 NLP，机器学习，图像算法，除了一些专有的算法，通用的算法有的会问，不过最好还是做多方面的准备

★ 编程题和算法题问得数量不算很多

★ 面试套路大概归结为下面流程：

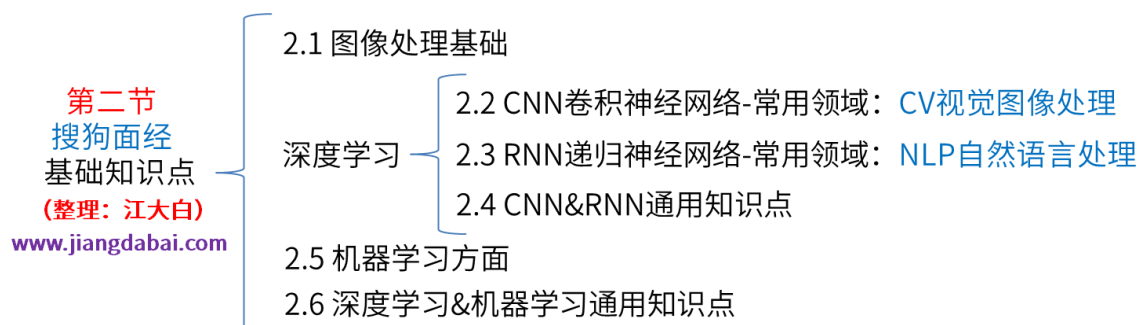
（1）先做一个简短的自我介绍

答：根据你项目的时间点结合你的学习历程简单介绍。比如研一上学的什么然后做了什么，研二学的什么然后做了什么。一定要简介别说细节，可以简单说效果提升。还有你获得过的国奖啊、校奖、acm 啊都要说一下。

（2）大体按照相关实习经历==发表过的论文（水准高的）>较水论文>top 比赛=项目的顺序进行发问。这个东西就看你对自己项目的理解了，没人能帮得了。

（3）手写代码，写不上来大概率 gg，不过 ai 公司问的算法都挺 easy。

## 2 搜狗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

- hog 细节
- 传统特征算子和 CNN 区别

### 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

- 卷积核和池化的细节?

##### 2.2.1.2 网络结构方面

- Resnet 的结构, 为什么是这样的?

##### 2.2.1.3 其他方面

- BN 解决的问题, 哪些优点, 测试和训练的模式是怎样的?
- 训练时出现 loss NAN 的可能因素, 分析原因并说如何解决?



## 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 的原理说一下？
- GRU 和 LSTM
- Multi-Head Attention

## 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

无

##### 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）

- 集成学习 bagging 和 boosting 的区别？
- gdbt 和 LR 的作用？

- xgboot 的原理？

## ② SVM（支持向量机）

- SVM 概述？
- SVM 说一下？有哪些核函数，核函数怎么选？
- SVM 和 lr 的异同

### 2.5.1.4 无监督学习-聚类方面

- k-mean 的原理
- DBSCAN 的原理

### 2.5.2 手推算法及代码

- 手推 gdbt？

## 2.6 深度学习&机器学习面经通用知识点

### 2.6.1 网络优化梯度下降方面

- 常用的优化方法？解释下 Adam 怎么好？

### 2.6.2 过拟合&欠拟合方面

- 如何防止过拟合？正则化为什么能防止过拟合效果？

### 2.6.3 其他方面

- 机器学习中“距离”的概念，知道哪几种距离？

## 3 搜狗面经涉及项目知识点

第三节  
搜狗面经  
项目知识点  
(整理: 江大白)  
[www.jiangdabai.com](http://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

- RCNN 和 fast-rcnn、faster-RCNN 的区别, 其中 RPN 的作用?
- 重点问了 nms、ssd、faster rcnn

### 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

- Bert 的 Multi-Head Attention

### 3.3 强化学习

无

### 3.4 机器学习方面

#### 3.4.1 推荐系统

- ftrl、frm、deepfm 的原理都讲一下?
- 深度学习在推荐系统的意义?

## 4 数据结构与算法分析相关知识点

第四节  
搜狗面经  
数据结构与算法分析  
(整理: 江大白)  
[www.jiangdabai.com](http://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 树

- 平衡二叉树的插入查找的时间复杂度？
- 最大公共子串？（区别于最大公共子序列）

#### 4.1.2 排序

- 每个排序算法讲了一遍，稳定的排序算法有哪些？

### 4.2 算法思想实战及智力题

#### 4.2.1 算法思想实战

- 海量数据下如何找到中位数？
- $m \times n$  矩阵吃苹果？

#### 4.2.2 智力题

- 8 个苹果，一个秤，挑出一个最重的苹果？

### 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  
搜狗面经  
编程高频问题  
(整理: 江大白)  
[www.jiangdabai.com](http://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 网络框架方面

- 熟悉的深度学习模型，有什么区别？

#### 5.1.2 基础知识

- python 深拷贝和浅拷贝

### 5.2 C/C++方面

#### 5.2.1.1 区别比较

- C++的类和继承？
- C++公有类和私有类？
- C++ 指针和引用？
- C++ malloc free 和 new delete 区别？

#### 5.2.1.2 讲解原理

- static 静态函数 静态变量解释解释？

- const 放在不同的位置，意义是什么？
- C++多态的解释？

## 6 操作系统高频问题：数据库&线程&常用命令等

第六节  
搜狗面经  
操作系统高频问题  
(整理: 江大白)  
[www.jiangdabai.com](http://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

- 进程和线程，如何实现多线程？

#### 6.2.2 常用命令

- 对 Linux 中 ubuntu 的熟悉程度？

## 7 技术&产品&开放性问题

### 7.1 技术方面

- 特征工程和算法谁更重要？

# 39|地平线机器人算法岗武功秘籍

## 1 地平线面经汇总资料

第一节  
地平线面经  
汇总资料  
(整理: 江大白)  
[www.jiangdabai.com](http://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 地平线面经整理心得

### 1.1 面经汇总参考资料

#### ① 参考资料:

- (1) 牛客网: 地平线机器人面经-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&答案&目录&心得:

- (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) 实习岗位类

【自动驾驶感知算法日常实习】

#### (2) 全职岗位类

【计算机视觉算法工程师】、【助理研究员】

## 1.3 面试流程时间安排

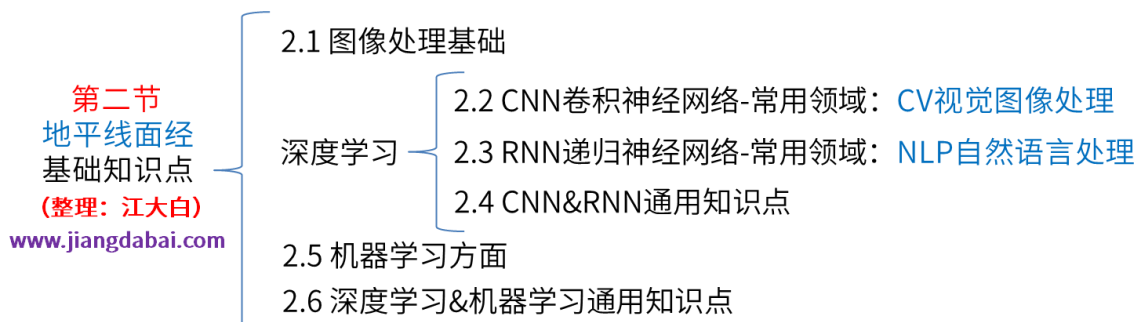
| 地平线面试流程-整理：江大白 |       |                            |         |
|----------------|-------|----------------------------|---------|
|                | 面试类型  | 面试流程                       | 备注（侧重点） |
| 第一面            | 技术面   | 自我介绍+项目/实习经验<br>+技术问答+算法编程 | 项目为主    |
| 第二面            | 技术面   | 自我介绍+项目/实习经验<br>+技术问答+算法编程 | /       |
| 第三面            | 技术综合面 | 自我介绍+项目/实习经验<br>+技术问答+算法编程 | /       |
| 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 地平线面试心得汇总

- ★ 问的不是很难，每一面都会问项目，讲解为什么这么做？
- ★ 问的基础不难，项目为主
- ★ 感觉基础 CV 领域和自动驾驶相关的都有做
- ★ 二面就是简历细致化面试，为什么这么做？为什么不这么做？你们没有再做优化吗？

## 2 地平线面试涉及基础知识点





## 2.1 图像处理基础

### 2.1.1 讲解相关原理

- 二值化图像中，给定一个函数可以返回某个点的值为 1 或者 0，现在判断两个点是否在同一个连通域内？
- 光流原理？

## 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

- 1\*1 卷积的作用，对同一个特征图使用 1\*1 卷积和 maxpooling 有什么区别，stride 为 2 的卷积操作和 maxpooling 是否有区别（除降维外的区别），区别在哪？

#### 2.2.1.2 网络结构方面

- Resnet 和 densenet 的原理和区别？
- shufflenet, mobilenet 讲一下？
- MobileNet V1,V2 原理，优化的方向？

#### 2.2.1.3 其他方面

- BN 原理，有几个参数，训练和测试的区别，多卡时 BN 如何处理？了解其他的归一化操作吗？（例如群组归一化？）

## 2.3 深度学习：RNN 递归神经网络方面

### 2.3.1 讲解相关原理

- 一些模型理解，比如为什么 cnn 可以用来做文本任务？

## 2.4 深度学习：CNN&RNN 通用的问题

### 2.4.1 基础知识点

- 讲一下 attention 的原理。你认为为什么有效或者什么情况下有效？
- 讲一下 SVM 的原理

## 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

- PCA 的原理、PCA 与 Autoencoder 的区别？
- 项目中 Auto-encoder 实际中是怎么采集数据确保有噪声和无噪声的？
- autoencoder 除了去噪还有什么功能？
- 如何实现 autoencoder？

##### 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）

- bagging 和 boosting 的方差和偏差比较？

##### ② SVM（支持向量机）

- 解释一下 SVM？

- SVM 如何实现多分类、SVM 核如何避免维度爆炸？
- 感知机与线性 SVM 的区别？
- SVM 的 loss 是什么？

#### 2.5.1.4 无监督学习-聚类方面

- kmeans 聚类如何选择初始点？

## 2.6 深度学习&机器学习面经通用知识点

无

## 3 地平线面经涉及项目知识点

第三节  
地平线面经  
项目知识点  
(整理: 江大白)  
[www.jiangdabai.com](http://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

- 重点探讨了小目标检测的 trick 和类别不均衡问题？
- 讲一下 faster rcnn
- 讲讲 RCNN 系列的演变过程。
- 讲讲 YOLO, SSD?
- 讲讲 RFCN, 重点阐述 position sensitive score map?
- 讲讲 FPN?

### 3.1.1.2 手写代码

- 写一下 NMS

### 3.1.2 图像分割

- 讲一讲语义分割的 ASPP 和 PSP 区别？
- 手写 MIOU

### 3.1.3 人脸识别

- Triplet Loss 用过吗？原理讲一下？

## 3.2 深度学习：RNN 递归神经网络方面

无

## 3.3 强化学习

无

## 3.4 机器学习方面

无

## 4 数据结构与算法分析相关知识点

第四节  
地平线面经  
数据结构与算法分析  
(整理: 江大白)  
[www.jiangdabai.com](http://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 树

- 手撕一个简单算法，递归？
- 二叉树层序遍历

### 4.1.2 排序

- 写一个体现你水平的排序算法？

## 4.2 算法思想实战及智力题

### 4.2.1 算法思想实战

- 有  $n$  个长度为  $m$  的由小到大的有序数组，合并为一个由小到大有序数组？（说了小根堆的思路，然后问建堆的时间复杂度，用我说的方法处理这个合并问题的时间复杂度）
- 有一次旅程用一个二维 list 存着，【(北京，上海) (上海，广州) (广州，东京) ...】，这个旅程是一个没有环路的，然后现在打乱这个 list 的顺序，最后输出一个【北京，上海，广州，东京，...】这样的一个 list？

## 4.3 其他方面

### 4.3.1 矩阵运算

- 全 0 (含有部分 1) 的矩阵中有一些 1 构成一个封闭曲线，把曲线内部的 0 变成 1？
- 用随机产生 7 以内数字的函数去写随机产生 10 以内数字的函数

## 4.4 Leetcode&剑指 offer 原题

无

## 5 编程高频问题：Python&C/C++方面

第五节  
地平线面经  
编程高频问题  
(整理: 江大白)  
[www.jiangdabai.com](http://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 网络框架方面

- Tensorflow 训练的整体流程？

#### 5.1.2 基础知识

- python 多继承？

### 5.2 C/C++方面

#### 5.2.1 基础知识

##### 5.2.1.1 内存相关

- C++如何高效管理动态内存？

##### 5.2.1.2 区别比较

- 深复制、浅复制
- array 和数组有什么区别？

##### 5.2.1.3 讲解原理

- 函数的形参 \*arg \*\*karg？
- C++11 如何有效防止内存泄\*\*r /> - 智能指针，weak\_ptr 讲一下？

## 6 操作系统高频问题：数据库&线程&常用命令等

第六节  
地平线面经  
操作系统高频问题  
(整理: 江大白)  
[www.jiangdabai.com](http://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

- 对 linux 的熟练程度，用过什么命令，什么作用？

## 7 技术&产品&开放性问题

### 7.1 技术方面

- Autoencoder 除了去噪还有什么功能、如何实现 autoencoder 的？
- 打开百度图片搜索，可以看到虽然同一行所有图片的高度是一样的，但他们的原始高度肯定是不一样的，这个后面可能是怎么实现的？

答：以每一行的第一张图片高度为基准，random 一个偏差，以基准加偏差作为这一行的高度，所有图片 resize 到这个高度，然后往固定的行宽里填充，如果变形太多，就取一个放到下一行。

- 浮点数在计算机里的表示，为什么会损失精度？具体表示是本科学的记不清了，大概是由符号位，有效数字位和幂次位组成，所以超过有效数字位的精度会被忽略掉。

怎么解决这个精度问题？

- 有一个图像分类的问题，类别有 2000 万类，不考虑计算资源和网络结构合理性的问题，暴

力地在最后做 2000 万的全连接。只看对前 900 万类的分类，效果很好，如果加上后面的 1100 万类，网络无法收敛，可能是什么原因？

答：既然不考虑网络结构和计算资源，那么只可能是数据的问题，有可能是类别不均衡导致的，可以用重采样或者部分类别做数据增广的方式解决。



# 40|58 集团算法岗武功秘籍

## 1 58 集团面经汇总资料

第一节  
58集团面经  
汇总资料  
(整理: 江大白)  
[www.jiangdabai.com](http://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 58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: 58 集团面经-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 面经框架&答案&目录&心得:

- (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

【NLP 算法工程师】、【计算机视觉算法工程师】

## 1.3 面试流程时间安排

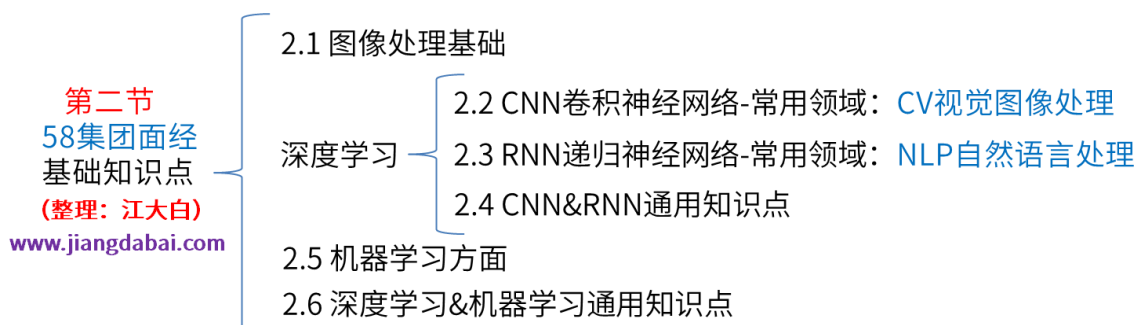
| 58集团面试流程 整理：江大白 |      |                            |         |
|-----------------|------|----------------------------|---------|
|                 | 面试类型 | 面试流程                       | 备注（侧重点） |
| 第一面             | 技术面  | 自我介绍+项目/实习经验<br>+技术问答+算法编程 | /       |
| 第二面             | 技术面  | 自我介绍+项目/实习经验<br>+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 58 集团面试心得汇总

- ★ 感觉算法研发的比较看的轻，更重要的是工程实现或者算法落地
- ★ 感觉推荐算法用的比较多
- ★ nlp 算法为主，还有少量的 cv 算法

## 2 58 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

- Sift 特征点匹配讲一下？

## 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

- DCN 可变形卷积用过吗？

#### 2.2.1.2 其他方面

- BN 在原理上怎么解决过拟合？

## 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

- 介绍 attention 机制？
- 注意力机制有哪些种类，本身原理上起了什么作用？

## 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 中的主题矩阵如何计算？

##### 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）

- 用了很多树模型嘛，说说树模型比线性模型的优势在哪里？
- 工业界也经常用线性模型，说说线性模型比树模型的优势？
- gbdt 和随机森林的区别？

##### A.基于 bagging：随机森林

- RF 的时间复杂度？

##### B.基于 boosting：Adaboost、GDBT、XGBoost

- gbdt 和 xgboost 原理和区别？

##### ② 逻辑回归 LR

- LR 和 gbdt 的时间复杂度？
- LR 为什么用 sigmoid 函数？
- LR 导数的取值范围？

### ③ SVM (支持向量机)

- 说点 SVM，什么是支持向量？
- 核函数有哪些？写一下？
- 这些核函数都是什么情况使用的？具体问什么情况用高斯。
- 你觉得 SVM 与 LR 的对比呢？
- SVM 的时间复杂度？

### ④ 决策树 (DT)

- 用了很多树模型嘛，说说树模型比线性模型的优势在哪里？
- 工业界也经常用线性模型，说说线性模型比树模型的优势？

#### 2.5.1.4 无监督学习-聚类方面

- kmeans 中 k 值如何确定。手肘法的纵坐标？
- DBSCAN 的计算原理，如何确定噪声点？

#### 2.5.2 手推算法及代码

- gbd 损失函数？
- xgboost 二阶泰勒展开表达式？
- 对 LR 求导，写一下，计算过程？
- LR 公式，损失函数，写一下？

## 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

- 常用损失函数有哪些？
- 回归问题主要用哪些损失函数？
- 数据不平衡怎么处理？

- 数据不平衡的损失函数有哪些？
- 交叉熵是什么原理？

## 2.6.2 压缩&剪枝&量化&加速

- 知识蒸馏是什么，通过什么方式来简化，比如 albert，具体原理是什么？

## 2.6.3 过拟合&欠拟合方面

- 怎么解决过拟合问题？

# 3 58 集团面经涉及项目知识点

第三节  
58集团面经  
项目知识点  
(整理: 江大白)  
[www.jiangdabai.com](http://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

- roi align 和 roi pooling 的区别？
- faster rcnn 和 yolo 的区别？

#### 3.1.1.2 损失函数

- focal loss 公式和解释？

#### 3.1.1.3 手写代码

- 手撕 nms

## 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

- BERT 的优缺点？
- PTM 都了解哪些，BERT 与 GPT 区别？
- 单项与双向在实际训练时有差别吗？
- bert 的 mask 会带来什么缺点吗？

##### ② Transformer

- Transformer 原理？

##### ③ HMM 隐马尔科夫模型

- 隐马尔可夫了解么？

##### ④ Word2vec

- 介绍 word2vec？
- word2vec 相关，什么情况用 CBOW，什么情况用 Skip-gram？
- 用了 word2vec，详细说说 softmax 和分层 softmax 还有负采样吧？怎么负采样？

##### ⑤ CNN 方面

- CNN 和 RNN 的区别？

## 3.3 强化学习

无

### 3.4 机器学习方面

无

## 4 数据结构与算法分析相关知识点

第四节  
58集团面经  
数据结构与算法分析  
(整理: 江大白)  
[www.jiangdabai.com](http://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

- 数组和链表的区别？

##### 4.1.1.2 链表

- 考了一个判断链表是否对称的，只用说出思想就行？

##### 4.1.1.3 字符串

- 字符串转成数字

#### 4.1.2 排序

- 问什么时候时间复杂度为  $O(n^2)$  ？
- 排序算法的时间复杂度？
- 快排相关，问什么时候时间复杂度为  $on^2$ ？



## 4.2 算法思想实战及智力题

### 4.2.1 算法思想实战

- 2000 个 1-2000 范围内的正整数，找第 k 大的，只说思想就行？
- 数组中出现最多的元素？
- 经典 DP
- 判断两个链表是否相交

## 4.3 其他方面

无

## 4.4 Leetcode&剑指 offer 原题

- 剑指原题：数组中出现最多的元素

## 5 编程高频问题：Python&C/C++方面

第五节

58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 网络框架方面

- Tensorflow 计算过程？
- 看简历上有 Keras 和 TensorFlow，面试官问了一下常用的哪些，都怎么用的，评价一下两个框架？

## 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  
58集团面经  
操作系统高频问题  
(整理: 江大白)  
[www.jiangdabai.com](http://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 TCP 协议相关

- TCP/IP 的三次握手四次挥手

#### 6.2.2 线程和进程相关

##### 6.2.2.1 讲解原理

- 多线程的原理讲一下？
- 什么是线程安全？

## 7 技术&产品&开放性问题

### 7.1 技术方面

- 分辨招聘信息是不是垃圾信息，比方说我现在已经把所有的数据都转成特征向量给你了，

怎么做？

# 41|Keep 算法岗武功秘籍

## 1 Keep 面经汇总资料

第一节  
keep面经  
汇总资料  
(整理: 江大白)  
[www.jiangdabai.com](http://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 keep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: Keep 面经-7 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&答案&目录&心得:

- (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】

## 1.3 面试流程时间安排

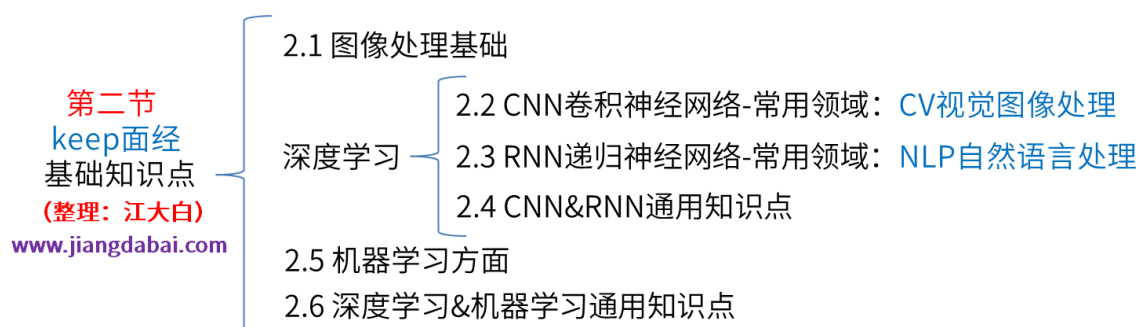
| Keep面试流程 整理：江大白 |      |                            |         |
|-----------------|------|----------------------------|---------|
|                 | 面试类型 | 面试流程                       | 备注（侧重点） |
| 第一面             | 技术面  | 自我介绍+项目/实习经验<br>+技术问答+算法编程 | /       |
| 第二面             | 技术面  | 自我介绍+项目/实习经验<br>+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 Keep 面试心得汇总

- ★ 算法岗好像没有图像方向的，未来可能会成立
- ★ 面试时不会问很多开放性的问题

## 2 Keep 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

无

## 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

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

无

##### 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）

- GBDT 的原理，流程，时间复杂度？
- GBDT 拟合负梯度，为什么拟合负梯度？
- T 棵树，H 高度，N 样本，P 特征，GBDT 训练的时间复杂度

## ② 逻辑回归 LR

- LR 的损失函数？
- LR 是不是凸函数，为什么？

## ③ SVM（支持向量机）

- SVM 的相关原理，比如损失函数，knn 的求距离方法？
- SVM 的损失函数？

## 2.6 深度学习&机器学习面经通用知识点

### 2.6.1 网络优化梯度下降方面

- 梯度下降法的效率和什么有关？

### 2.6.2 其他方面

- 梯度消失，梯度爆炸怎么解决？

## 3 Keep 面经涉及项目知识点

第三节  
keep面经  
项目知识点  
(整理: 江大白)  
[www.jiangdabai.com](http://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

无

### 3.2 深度学习：RNN 递归神经网络方面

无

### 3.3 强化学习

无

### 3.4 机器学习方面

无

## 4 数据结构与算法分析相关知识点

第四节  
keep面经  
数据结构与算法分析  
(整理: 江大白)  
[www.jiangdabai.com](http://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

- 求连续数组最大和？

##### 4.1.1.2 链表

- 两个有序链表合并，N 个呢？

##### 4.1.1.3 字符串

- 两个字符串找最长公共子串？
- 比如 12345,45123 两个字符串,判断 str2 能否顺时针旋转变成 str1,要求时间复杂度  $O(n)$ , 空间复杂度  $O(n)$



### 4.1.2 排序

- 介绍一下快排？
- 排序算法有哪些？

### 4.1.3 搜索

- 使用 python 对图的 bfs, dfs 实现？

## 4.2 算法思想实战及智力题

### 4.2.1 算法思想实战

- 比如12345,45123两个字符串,判断str2能否顺时针旋转变成str1,要求时间复杂度 $O(n)$ ,空间复杂度 $O(n)$ ？
- 图的 bfs, dfs 实现

## 4.3 其他方面

### 4.3.1 矩阵运算

- 矩阵中最长的连续1的个数？
- 矩阵(0,0)到右下的最短路径？

## 4.4 Leetcode&剑指 offer 原题

无

## 5 编程高频问题：Python&C/C++方面

第五节  
keep面经  
编程高频问题  
(整理: 江大白)  
[www.jiangdabai.com](http://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

无

## 6 操作系统高频问题：数据库&线程&常用命令等

第六节  
keep面经  
操作系统高频问题  
(整理: 江大白)  
[www.jiangdabai.com](http://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

无

## 7 开放性问题&场景问题

### 7.2 产品方面

- 你觉得 Keep 哪里能用到图像算法？

# 42|寒武纪算法岗武功秘籍

## 1 寒武纪面经汇总资料

第一节  
寒武纪面经  
汇总资料  
(整理: 江大白)  
[www.jiangdabai.com](http://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 寒武纪面经整理心得

### 1.1 面经汇总参考资料

#### ① 参考资料:

- (1) 牛客网: 寒武纪面经-13 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&答案&目录&心得:

- (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

【算法工程师 (深度学习算法推理引擎/高性能计算库方向)】、【深度学习算法研究员】、【计算机视觉算法工程师】、【深度学习框架研发工程师】

## 1.3 面试流程时间安排

| 寒武纪面试流程 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<br>+技术问答+算法编程 | /       |
| 第二面            | 总监面  | 自我介绍+项目/实习经验<br>+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。

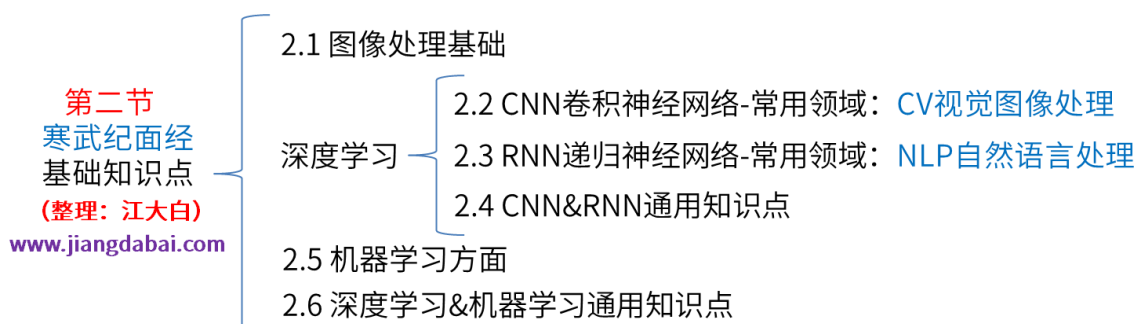
**其他注意点：**

- HR 面试之后，可能会有性格测试，通过之后就等通知了

## 1.4 寒武纪面试心得汇总

- ★ 感觉对 c++ 需求很大，面试中很多都是这方面的问题
- ★ 平日里准备的更多的是深度学习训练的细节，不过面试中会问框架方面的问题

## 2 寒武纪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

- 传统图像处理，比如腐蚀膨胀的原理？
- 描述一下空洞卷积？

## 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

- 卷积层的计算细节、激活函数如何引入非线性？
- 为什么说引入非线性，网络就可以表示几乎所有的连续函数？
- 有没有自己从头实现过 Conv2D？
- $5 \times 5$  卷积核和 2 个  $3 \times 3$  卷积核哪个好？
- 空洞卷积怎么实现，什么作用？

#### 2.2.1.2 其他方面

- 两张卡同时训练的时候每张卡上的 bn 的参数是否一样？

### 2.2.2 公式推导

- CNN 的反向推导计算？

## 2.3 深度学习：RNN 递归神经网络方面

### 2.3.1 讲解相关原理

- 讲讲 embedding 吧，讲讲 LSTM 吧？

## 2.4 深度学习：CNN&RNN 通用的问题

### 2.4.1 基础知识点

- 介绍卷积网络、RNN、bilstm?

### 2.4.2 模型评价

- 在给定场景下（好像是关于神经网络的层级结构的确定问题），如何给出最优化判据（就是采用什么样的方法来判断结果是好的，这样就能够提供反馈信息给神经网络，并以此来优化结构和参数。

## 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

- PCA 的原理?

##### ② 特征选择

无

## 2.6 深度学习&机器学习面经通用知识点

### 2.6.1 激活函数方面

- 激活函数如何引入非线性?

- 证明引入非线性，为何就可以说这个网络可以表示几乎所有的连续函数？

## 2.6.2 压缩&剪枝&量化&加速

- 多核 CPU 上矩阵乘法怎么加速？
- 卷积怎么加速？

## 2.6.3 过拟合&欠拟合方面

- 防止过拟合的方法？

# 3 寒武纪面经涉及项目知识点

第三节  
寒武纪面经  
项目知识点  
(整理: 江大白)  
[www.jiangdabai.com](http://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

- 对 Yolo 算法的了解程度？相关细节？
- 遥感目标检测，你做了那些工作，遇到哪些难点？

## 3.2 深度学习：RNN 递归神经网络方面

无

## 3.3 强化学习

无



### 3.4 机器学习方面

无

## 4 数据结构与算法分析相关知识点

第四节  
寒武纪面经  
数据结构与算法分析  
(整理: 江大白)  
[www.jiangdabai.com](http://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

- 给定一个无序数组，找出里面数据和为 0 的组合，分析时间和空间复杂度？
- 数组和链表区别

##### 4.1.1.2 链表

- 对一个链表进行排序，要求时间复杂度  $O(n\log n)$ ，空间复杂度为  $O(1)$ ？（归并排序）
- 环形链表

##### 4.1.1.3 字符串

- 字符串排序（按字母顺序）？
- 0-n-1：中找到一个重复的数字

#### 4.1.2 树

- 什么是平衡二叉树？红黑树？

### 4.1.3 排序

- 在  $n$  个值里找最小的  $k$  个数，分析时间复杂度？
- 快速排序时间复杂度？
- 堆排序原理

## 4.2 算法思想实战及智力题

### 4.2.1 算法思想实战

- 如果给定一个正整数  $N$ ，对于一个最小位是 2 的  $s$  次幂的数，需要多少位才能表示这个数？  
如何确定数字系统中的参数位数？
- 找出数组中的重复数字（哈希表）
- 两个鸡蛋掉落问题（动态规划）

## 4.3 其他方面

### 4.3.1 数论

- （线性代数）特征值与奇异值
- 如果给定一个正整数数  $N$ ，对于一个最小位是 2 的  $s$  次幂的数，需要多少位才能表示这个数？如何确定数字系统中的参数位数。

## 4.4 Leetcode&剑指 offer 原题

无

## 5 编程高频问题：Python&C/C++方面

第五节  
寒武纪面经  
编程高频问题  
(整理: 江大白)  
[www.jiangdabai.com](http://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 网络框架方面

##### 5.1.1.1 Tensorflow 相关

- 从构建模型到训练的过程中，tf 会用到的特性，想到多少说多少？
- 比较一下 tf 2 的 eager execution 和 PyTorch 的动态图的优劣？

##### 5.1.1.2 Caffe 相关

- 对 caffe 的框架理解得怎么样？知道 caffe 的框架组成吗？

##### 5.1.1.3 其他

- 使用什么深度学习框架，如何设计一个神经网络结构？
- 熟悉框架有哪些，tensorflow 或者 pytorch 做扩展开发吗？
- pytorch 与 tensorflow 之间的优缺点？
- 静态图与动态图的区别？

#### 5.1.2 基础知识

##### 5.1.2.1 区别比较

- tuple 和 list 的区别？

### 5.1.2.2 讲解原理

- python 或者 C++是否熟悉，python 如何包装 C++的库函数？
- python2 转 python3 项目如何向下兼容？
- python 中为何要使用 numpy？

### 5.1.2.3 讲解应用

- python 中的 self 像是 C++的什么？
- list 插入的操作，三种？

## 5.2 C/C++方面

### 5.2.1 基础知识

#### 5.2.1.1 内存相关

- C++中内存越界、内存泄漏、内存溢出？
- C++和 C 的指针有何区别？

#### 5.2.1.2 区别比较

- map 和 unordered\_map 的区别
- new、malloc 的区别，它们用在哪些存储区？
- 三个关键字的访问权限与继承权限的区别
- 继承和多态
- 堆和栈的区别？
- C++ const 和 volatile 什么作用？编译器会怎么处理？

#### 5.2.1.3 讲解原理

- std::vector 内存的分配？
- C++11 的新特性？

- 智能指针有几种，介绍每个的特性是什么？指针内存释放的机制？
- weak\_ptr 的使用场景
- 类型转换方法有几种，cast 方法，介绍一下
- 优先队列底层数据结构，堆的数据结构有什么特点
- 堆定义的头文件在哪，用过哪几种队列，优先队列自定义排序方法
- 介绍 multi\_map 的数据结构
- 如何实现多态？
- 判断一个链表有环
- 内存管理单元的作用
- 什么是指令流水？
- C++中的虚函数，各种类型转换方式？
- 析构函数的作用，什么时候作用？
- 虚函数的作用，将基类里的 virtual 关键字去掉会怎么样？
- 将基类的虚函数 的 virtual 关键字去掉又会怎么样？
- long 有几个字节，long long 跟之前的 C++版本有什么不同？会超过 8 个字节吗？
- 右值方便了哪些写法，有什么优点？左右值转换的函数是哪个？
- 右值引用在模板里面的作用，写模板用 std::move()来传参的作用，forward 的作用
- define
- 宏定义
- malloc
- static 的作用
- 继承种类？虚继承解决什么问题？
- 讲解一下虚函数的原理，析构函数可以写成虚函数吗？哪些情况需要写成虚函数？

#### 5.2.1.4 讲解应用

- C++中利用 opencv 存储图像的结构体？

## 5.2.2 手写代码相关

- C++里用过 lambda 表达式嘛？说说？它对开发效率和运行效率的提高。

# 6 操作系统高频问题：数据库&线程&常用命令等

第六节  
寒武纪面经  
操作系统高频问题  
(整理：江大白)  
[www.jiangdabai.com](http://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

#### 6.2.1.1 区别比较

- 线程和进程的区别？

#### 6.2.2 常用命令

- 对操作系统的了解有多深？
- linux 常用命令有哪些？

#### 6.2.3 其他问题

- 生产者消费者模式，以及临界资源有哪些？
- 造成死锁的四个必要条件？

- 软硬链接的区别？

## 7 技术&产品&开放性问题

### 7.1 技术方面

- 在给定场景下（好像是关于神经网络的层级结构的确定问题），如何给出最优化判据（就是采用什么样的方法来判断结果是好的，这样就能够提供反馈信息给神经网络，并以此来优化结构和参数？
- 浅谈关于人工智能与自动控制结合方法，就是人工智能方法该怎么样应用到自动控制中去？
- git 中如何恢复/撤销之前已经 push 的操作？

### 7.2 产品方面

- 做一个项目大概的周期和时间分布是怎样的？
- 数据、模型、优化等各占多少？

## 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](http://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 面经框架&答案&目录&心得:

- (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) 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

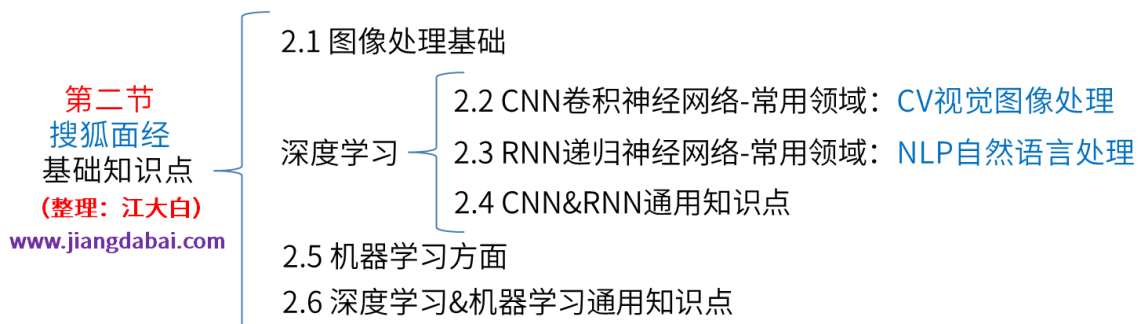
| 搜狐面试流程-整理：江大白 |      |                            |             |
|---------------|------|----------------------------|-------------|
|               | 面试类型 | 面试流程                       | 备注（侧重点）     |
| 第一面           | 技术面  | 自我介绍+项目/实习经验<br>+技术问答+算法编程 | 项目问的很细      |
| 第二面           | 技术面  | 自我介绍+项目/实习经验<br>+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 gbdt

## 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](http://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](http://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](http://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](http://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？



## 44|有赞算法岗武功秘籍

### 1 有赞面经汇总资料

第一节  
有赞面经  
汇总资料  
(整理: 江大白)  
[www.jiangdabai.com](http://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 有赞面经整理心得

#### 1.1 面经汇总参考资料

##### ① 参考资料:

- (1) 牛客网: 有赞面经-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&答案&目录&心得:

- (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】

## 1.3 面试流程时间安排

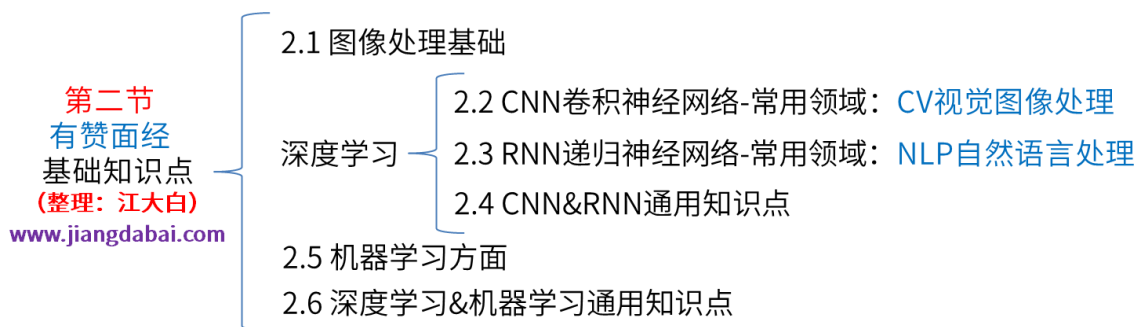
| 有赞面试流程-整理：江大白 |      |                            |         |
|---------------|------|----------------------------|---------|
|               | 面试类型 | 面试流程                       | 备注（侧重点） |
| 第一面           | 技术面  | 自我介绍+项目/实习经验<br>+技术问答+算法编程 | /       |
| 第二面           | 技术面  | 自我介绍+项目/实习经验<br>+技术问答+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 有赞面试心得汇总

- ★ 简历上写的项目及知识点一定要会，主要还是围绕项目来问，不要瞎写
- ★ 最好有一个你比较熟悉的项目，一定要熟悉，面试官问了好多项目细节的东西

## 2 有赞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 的原理？

#### 2.2.1.2 其他方面

- Batch normalization 的原理？
- 为什么会发生梯度爆炸、梯度消失？如何缓解梯度爆炸和梯度消失？

## 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 的原理？
- BLSTM 和 LSTM 的区别？（解决上下文语义问题）
- GRU 和 LSTM？
- 是否了解 GRU？
- 说下 rnn，rnn 有什么问题，说下 lstm，lstm 有什么缺点，gru 了解吗？
- LSTM 和 RNN 区别是什么（LSTM 缓解梯度消失问题）
- Batch normalization 作用（1.使得输入规范化到有梯度的区间，防止梯度消失。 2.改变输入数据的均值和方差，起到数据增强的作用，防止过拟合）
- CRF 可以做实体识别吗（可以），那为什么要用 BLSTM-CRF(没答上来)，CRF 和 HMM 的区别，CRF 原理（项目中只是用到 CRF，具体原理只懂一点）

## 2.4 深度学习：CNN&RNN 通用的问题

### 2.4.1 基础知识点

- attention 的原理了解吗？

### 2.4.2 模型评价

- 给了正负类样本的得分，如何计算 AUC？

## 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

- 介绍一下什么是 PCA、如果不是方阵怎么做？

##### 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）

- 随机森林和 gbdt、xgbost 和 gdbt 的改进？
- LR, xgboost 的原理说一下？

##### ② 逻辑回归 LR

- 逻辑回归的损失函数是什么？

- 逻辑回归的求解时，梯度的维度是多少？
- LR 介绍一下？

### ③ SVM（支持向量机）

- SVM 介绍一下，核函数有什么要求？

#### 2.5.1.4 无监督学习-聚类方面

- 问了 kmeans，具体步骤是怎样的？kmeans++

## 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

- 逻辑回归的损失函数是什么？（交叉熵损失）为什么？（因为引入 sigmoid 函数，如果使用 MSE 损失会使目标函数非凸，无法求解）
- 逻辑回归的求解时，梯度的维度是多少？（二维？）
- 逻辑回归为什么用 sigmoid 函数？（为了求出分类概率）
- 为什么不采用其他函数（sigmoid 函数对偏远值不敏感？）

### 2.6.2 正则化方面

- L1 和 L2 正则化的优缺点

### 2.6.3 过拟合&欠拟合方面

- 过拟合解决的方法（L1，L2 正则化，增加训练样本，特征选择）
- L1，L2 正则化原理（L1 是假设参数服从拉普拉斯分布，L2 是假设参数服从高斯分布）
- L1 为什么能使特征稀疏（L1 范数，使权重为 0，对应的特征则不起作用，使特征稀疏）

## 3 有赞面经涉及项目知识点

第三节  
有赞面经  
项目知识点  
(整理: 江大白)  
[www.jiangdabai.com](http://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

无

### 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

- Bert 的原理?

### 3.3 强化学习

#### 3.3.1 讲解原理

- DQN 的损失函数是什么?
- Double Q-Learning 的缺点是什么?
- 策略梯度损失函数是什么?
- 解释一下 MC、TD、DP 的区别?

## 3.4 机器学习方面

### 3.4.1 推荐系统

- 什么是 CB、CF 算法？

## 4 数据结构与算法分析相关知识

第四节  
有赞面经  
数据结构与算法分析  
(整理: 江大白)  
[www.jiangdabai.com](http://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 树

- 什么是二叉树，平衡二叉树？红黑树？
- 二分查找的时间复杂度？

#### 4.1.2 排序

- 稳定的排序方法有哪些？不稳定的有哪些？
- 哪个效率最高？快排时间复杂度？快排的最快情况？如何解决？
- 考虑空间复杂度，如何排序？
- 排序，稳定性和空间复杂度
- 归并排序的时间复杂度，空间复杂度，归并排序的思想是什么？
- 给一堆数字，如何找出 topk 个数？

## 4.2 算法思想实战及智力题

### 4.2.1 算法思想实战

- 如何求一个数组中前 5 个最小的数？
- 一个场景，100 万个商家，ID 号是顺序的，但是其中有两个 ID 丢失了，问怎样快速的找到丢失的 ID？（我说了因为顺序，假设 ID 是从 1 开始，那么对应的下标为 0，使用二分查找就可以了）

### 4.3 其他方面

无

### 4.4 Leetcode&剑指 offer 原题

- Leetcode64：不同路径

## 5 编程高频问题：Python&C/C++方面

第五节  
有赞面经  
编程高频问题  
(整理: 江大白)  
[www.jiangdabai.com](http://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

- python 的内存管理机制（垃圾回收）

### 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  
有赞面经  
操作系统高频问题  
(整理: 江大白)  
[www.jiangdabai.com](http://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

#### 6.1.1 基础问题

##### 6.1.1.1 区别比较

- 数据库的连接方式，有什么区别？
- 数据库中删除的关键字，他们三个之间有什么区别？

##### 6.1.1.2 讲解原理

- 数据库的范式，对照着自己说的范式，举三个例子说明一下违反了三范式的实例

### 6.2 操作系统方面

#### 6.2.1 TCP 协议相关

- TCP\UDP 是什么？区别？场景？对应协议？ TCP 的三次握手、四次挥手过程

## 7 技术&产品&开放性问题

无

# 45|知乎算法岗武功秘籍

## 1 知乎面经汇总资料

第一节  
知乎面经  
汇总资料  
(整理: 江大白)  
[www.jiangdabai.com](http://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 知乎面经整理心得

### 1.1 面经汇总参考资料

#### ① 参考资料:

- (1) 牛客网: 知乎面经-5 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&答案&目录&心得:

- (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】

## 1.3 面试流程时间安排

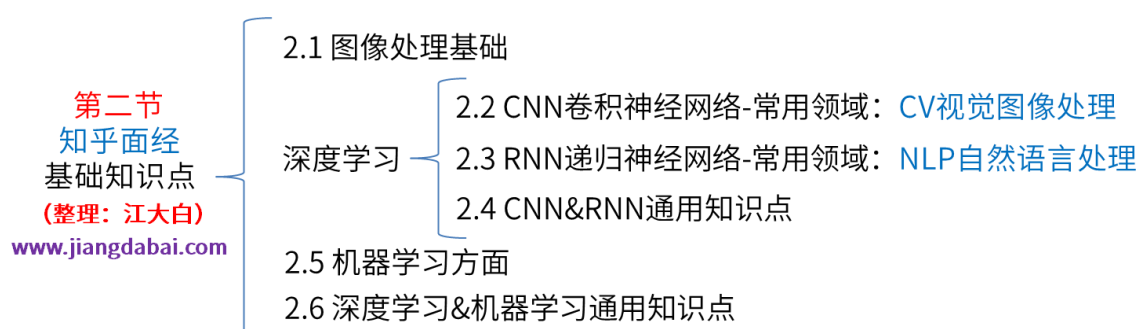
| 知乎面试流程-整理：江大白 |      |                            |         |
|---------------|------|----------------------------|---------|
|               | 面试类型 | 面试流程                       | 备注（侧重点） |
| 第一面           | 技术面  | 自我介绍+项目/实习经验<br>+技术问答+算法编程 | /       |
| 第二面           | 技术综合 | 自我介绍+项目/实习经验<br>+技术问答+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 知乎面试心得汇总

★ 感觉问得不是很深，会考算法题和概率题

## 2 知乎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

无

## 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 模型评价

- 精确率，召回率的原理？

## 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

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）

- boosting 和 bagging 的区别？bagging 为什么能减小方差？

##### ② 逻辑回归 LR

- LR 的损失函数？

##### ③ SVM（支持向量机）

- SVM 损失函数？

##### ④ 决策树（DT）

- 决策树分裂准则？

- 信息熵与信息增益公式推导？

## 2.5.2 手推算法及代码

- 手推 LR

## 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

- 交叉熵损失函数，0-1 分类的交叉熵损失函数的？
- 0-1 分类如果用平方损失为什么用交叉熵而不是平方损失？

## 3 知乎面经涉及项目知识点

第三节  
知乎面经  
项目知识点  
(整理: 江大白)  
[www.jiangdabai.com](http://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

无

### 3.2 深度学习：RNN 递归神经网络方面

#### 3.2.3 自然语言处理 NLP

- Word2vec 的原理简单介绍下？
- word2vec 里面的层次索引？

### 3.3 强化学习

无

### 3.4 机器学习方面

#### 3.4.1 推荐系统

- DeepFM 介绍下?
- FM 推导?

## 4 数据结构与算法分析相关知识点

第四节  
知乎面经  
数据结构与算法分析  
(整理: 江大白)  
[www.jiangdabai.com](http://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

- 数组排序，假设数组排序后的位次和排序前的位次绝对值差值小于 K，有什么比快排好的算法?
- 一个数组，所有数组都出现了两次，只有一个数出现了一次，返回这个数?
- 一个数组，一个数出现了超过一半次数，返回这个数?

##### 4.1.1.2 链表

- 链表转换 (abcde -> aebdc)

- 两个单链表，找出相同值的位置？

#### 4.1.1.3 字符串

- 不占用额外空间去除字符串中的空格  $O(1)$  写法？
- 将除法的结果用字符串返回，如果能够除尽，则返回相除的结果，如果不能除尽，则无限循环部分用[]标记。-参考：<https://www.jianshu.com/p/40259bb05357>

#### 4.1.2 树

- 树中两个节点的第一个的公共祖先？

### 4.2 算法思想实战及智力题

无

### 4.3 其他方面

#### 4.3.1 数论

- 什么是凸函数？

### 4.4 Leetcode&剑指 offer 原题

无

## 5 编程高频问题：Python&C/C++方面

第五节  
知乎面经  
编程高频问题  
(整理: 江大白)  
[www.jiangdabai.com](http://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

- python 里的 gc 和多线程说一下?

## 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  
知乎面经  
操作系统高频问题  
(整理: 江大白)  
[www.jiangdabai.com](http://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

无

## 7 技术&产品&开放性问题

无