

# 8|网易算法岗武功秘籍

## 1 网易面经汇总资料

第一节  
网易面经  
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 网易面经整理心得

### 1.1 面经汇总参考资料

#### ① 参考资料:

- (1) 牛客网: 网易面经-83 篇, [网页链接](#)
- (2) 知乎面经: [点击进入查看](#)
- (3) 面试圈: [点击进入查看](#)

#### ② 面经参考答案:

- (1) 面经答案: [点击进入查看](#)

### 1.2 面经涉及招聘岗位

#### (1) 实习岗位类

【网易机器学习实习岗】

#### (2) 全职岗位类

【网易游戏研发工程师】、【网易考拉机器学习工程师】、【网易考拉 NLP 工程师】、【数据挖掘工程师】、【网易有道云算法工程师】、【网易雷火机器学习】、【网易互娱 NLP】、

【网易云音乐机器学习】、【网易严选算法工程师】、【网易云音乐机器学习】、【网易云音乐音频算法工程师】

## 1.3 面试流程时间安排

网易面试流程整理：江大白			
	面试类型	面试流程	备注（侧重点）
第一面	技术面	自我介绍+项目/实习经验 +技术问答	项目为主
第二面	技术面	自我介绍+项目/实习经验 +技术问答+算法编程	项目为主
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。

**其他注意点：**

- 第三面，技术 Leader 总监面，有的人有，有的人没有

## 1.4 网易面经面试心得汇总

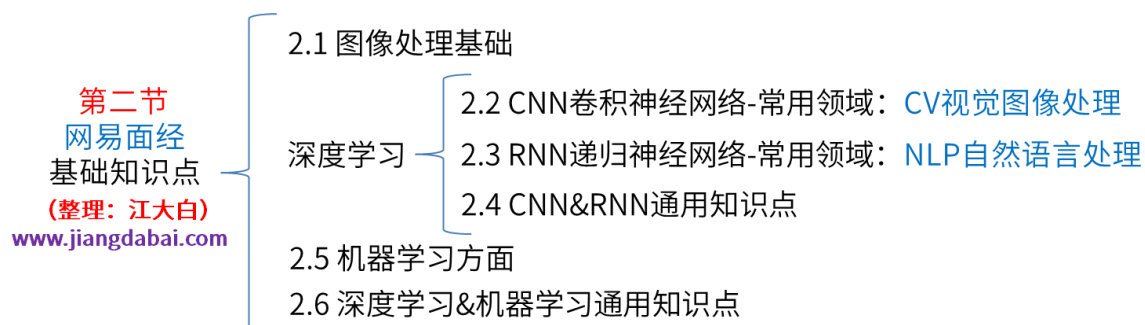
★ 先准备简历，简历干净凝练，最好所有内容都直指一个方向，最重要的是，简历上的任何东西一定要搞的明明白白，并且可以侃侃而谈，这很关键。

★ 相当重视基础，问的比较偏特征工程。先聊项目论文和实习，但是没细问。

★ 主要是问项目，然后根据项目里问一些细的技术点，每个项目都仔细问，确认是不是本人做的。

★ 面试官的风格是连环追问，你说一个答案他会在你基础上再问，为什么用这些？

## 2 网易面经涉及基础知识点



### 2.1 图像处理基础

- 传统图像处理: 边缘检测、均值滤波、霍夫变换

### 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?
- CNN 卷积核的计算, 参数量怎么算?
- 传统卷积核 depthwise conv 的区别?
- 是否了解空洞卷积?
- 卷积核参数怎么算, DW 卷积参数怎么算?
- 简历中提到了 dropout, 那具体说下 dropout 操作细节、怎么解决 dropout 之后输出值减小的问题
- Dropout、BN 的原理、训练过程以及测试过程的具体做法?
- CNN 可视化特征?
- CNN 原理, 池化的目的?

### 2.2.1.2 池化方面

- 介绍 CNN 为什么要用池化层？

### 2.2.1.3 网络结构方面

- shuffleNet v2, mobileNet v2 的结构？
- senet 结构？senet 为什么是这样的？为什么非要加池化呢？
- ResNet50 的 bottleneck 结构 (conv1+bn+ReLU+conv2....), 与 ResNet18 差别？
- 了解哪些 backbone, MobileNet v1 v2 v3 的差别？
- resnet 比 alex vgg 一类有什么好处？
- 为什么 resnet 可以深一些？
- VGG 优点, resnet 优点特点？

### 2.2.1.4 其他方面

- 介绍下 CNN 原理？CNN 之所以成功的原因是什么？CNN 的参数共享是什么？
- BN 的原理, BN 的训练和测试的区别, 讲一下过程？
- BN 层介绍, 作用, 为什么？
- BN 层参数有几个, forward 怎么计算？
- 说一下 BN 的归一化操作放在激活前还是激活后, 两种有什么区别？
- 梯度爆炸和梯度消失原因, 解决策略？
- 深度学习权重初始化方法有哪些？
- relu 为什么可以解决梯度消除, 除了这个好处还有什么好处？

## 2.2.2 数学计算

- 输入图像  $N \times N$ , 卷积核  $k \times k$ , 问计算方式, 时间复杂度？补充：单通道情况  $O(N \times N \times k \times k)$ , 如何优化？提示了重复计算？
- 输入  $n \times n \times 1$ , 卷积核  $k \times k \times 6$ , 问输出尺寸和参数量？

### 2.2.3 公式推导

- Softmax 公式？

## 2.3 深度学习：RNN 递归神经网络方面

### 2.3.1 讲解相关原理

- 介绍下 RNN？
- RNN 的改进有哪些？讲 lstm 和 gru，对于更长的序列怎么处理
- LSTM 和 RNN 的区别？lstm 为什么好？
- LSTM 的思想，原理，结构和公式？

### 2.3.2 手绘网络原理

- 推一遍 LSTM？
- RNN 原理，画一下？
- 写 RNN 公式

## 2.4 深度学习：CNN&RNN 通用的问题

### 2.4.1 基础知识点

- Attention 机制及它是如何聚焦的？
- 交叉验证的原理，作用，与直接划分训练集和测试集相比的优点？
- 数据增强方法？

### 2.4.2 模型评价

- 如何衡量模型性能，指标如何计算？常用的评价指标及含义？
- AUC 的意义，ROC 的绘制方式，AUC 的优势（不平衡数据集的情况）
- 如何评价一个分类器，AUC 的工程计算方式，ROC 曲线？

## 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

- 数据清洗如何做？
- MCMC 如何实现抽样，MH 和 Gibbs 抽样的区别，详细讲一下他们都是怎么实现的？

#### 2.5.1.2 特征工程

##### ① 特征降维

- 降维方法有哪些？PCA 的原理，作用？
- SVD 的原理？

##### ② 特征选择

- 常用特征选择的方法？如何把类别型数据转为数值型？数据缺失值处理？

#### 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）

- 为什么两个重分类分支融合的不做 bagging 和 boosting，两个概念的本质是什么意思，为什么效果会变好呢？，boosting 为什么多个分类器合在一起效果就变好了？
- xgb、gbdt、RF 区别与联系、xgb 的优势？RF/AdaBoost/GBDT 怎么做推断？

##### A.基于 bagging: 随机森林

- 随机森林了解吗？随机森林是什么，每棵树有啥区别？对比了一下 RF 和 GBDT？
- 随机森林/adaboost/gbdt 的原理，相同点不同点。在训练和测试阶段各是如何进行的？
- 随机森林的树会不会限制它的生长（不会），gbdt 的树呢（会），为什么？
- 随机森林和 GBDT 的差异描述一下？

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

- 介绍一下 GBDT？GBDT 有哪些可调的参数呢？一般怎么选取参数呢？

- GBDT 如何进行多分类？
- 讲一下 xgboost 的原理？xgboost 的过程、损失函数？为啥泰勒展开成二阶，作用？
- 为什么 XGBoost 效果好于随机森林？RF 怎么解决的过拟合问题？
- 如何做特征选择，xgboost 做特征选择的时候，重要性是如何确定的，信息增益做特征选择如何做的，特征选择方法？xgboost 如何一步步构建分类树的？

## ② 线性回归

- 线性回归和逻辑回归区别？

## ③ 逻辑回归 LR

- 介绍下逻辑回归原理？逻辑回归推导？
- LR, SVM 怎么算的，损失函数是什么？
- LR 处理的特征是离散的还是连续的？离散化，会有什么影响吗，比如一个特征取值 0-1，需要离散化吗？
- 写 LR 的公式和 loss function
- LR 的 loss 是什么？
- 把 LR 的 loss 改成平方损失可以么？为什么？
- LR 的梯度下降有几种优化器？
- LR 中存在相同特征的话，对模型预估有影响吗？

## ④ SVM（支持向量机）

- SVM 了解吗？（SVM 是通过最小间隔最大化寻找超平面）；为什么么要最大化最小间隔呢？
- SVM 需不需要做 normalization？
- SVM 和 LR 的区别？
- SVM 有哪些核函数，对应有哪些使用场景和特点？
- 对于 SVM，假如先把数据映射到高维，然后不使用核函数，如何？为什么大多数人选择使用核函数？

## ⑤ 决策树 (DT)

- 说一下决策树的原理?
- 决策树的分裂策略: ID3, C4.5, Gini 指数, 选一个讲一下?
- ID3、C4.5、CART 的区别
- 写信息增益、信息增益率、基尼系数的公式, 讲解原理
- 树有几种剪枝的方式, 各有什么优缺点?

### 2.5.1.4 无监督学习-聚类方面

- 问了 kmeans 的计算过程?
- 哪些条件对 kmeans 的影响最大? k 值选取?
- kmeans 时间复杂度和空间复杂度?
- kmeans 和 EM 算法的关系, 哪一步是 E 步, 哪一步是 M 步?

## 2.5.2 手推算法及代码

### 2.5.2.1 手推公式

- 完整推导 SVM?
- LR 公式写一下?
- GBDT 推导一下?

### 2.5.2.2 手写代码

- 手写 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

- 常用 loss 函数?
- 写一下多分类交叉熵的公式?
- 解释 triplet loss, 详解 triplet loss, triplet loss 反向传播?



- 分类为什么用交叉熵？

## 2.6.2 激活函数方面

- 各种激活函数介绍一下，优缺点及适用场景？
- 激活函数的作用？
- sigmoid 和 softmax 的区别？
- 为什么使用 sigmoid 激活函数会导致梯度消失？

## 2.6.3 网络优化梯度下降方面

- 用过那些优化器，SGD 和 Adam 分别在什么情况下使用？
- 牛顿法和拟牛顿法讲一下过程？
- 有哪些二阶优化方法，牛顿法存在什么问题？
- sgd 和 adam 的区别，有啥优点？
- bgd, sgd, mini-bgd 的区别？
- 手推：adam, sgd, momentum, rmsprop，通过公式解释一下为什么他们能优化？
- 最优化方法，ADAM？
- FTRL、牛顿法懂吗？
- SGD 与牛顿法的区别在哪？

## 2.6.4 正则化方面

- 各种正则化原理、方式及各自优缺点？L1 和 L2 正则的区别？
- 从数学角度讲一下正则项为什么能防止过拟合？
- 写一下 LR 的损失函数，加上 L1 /L2 正则化；然后解释原理，分析不同点，怎么用？

## 2.6.5 压缩&剪枝&量化&加速

- 剪枝用了哪几种方法，怎么训练的？
- 用过量化吗？，模型训练完成后 32bit 怎么量化为 8bit？

## 2.6.6 过拟合&欠拟合方面

- 描述下过拟合和欠拟合？
- 如何判断过拟合，如何解决过拟合？
- 为什么会减小过拟合的风险？（惩罚非线性参数的力度较大，减小模型非线性的程度）

## 3 网易面经涉及项目知识点

第三节  
网易面经  
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

- 目标检测的整个流程，包括数据处理、模型训练、模型选择？
- 说说 Faster-RCNN, YOLO, SSD, FPN?
- Faster-rcnn 的训练过程是怎样的？
- 目标检测中的 mAP?
- 介绍你用过的一个目标检测算法？

#### 3.1.1.2 损失函数

- focal loss 介绍？

### 3.1.2 OCR

- CTPN 原理，文字识别的实现，遇到的问题，数据集，做的改进？

- 车牌识别实现，遇到的问题，数据集，数据增强？
- 文字检测的一些 trick 说一下？
- 最新的文字检测方法了解吗？

### 3.1.3 超分辨

- 超分算法的缺陷及改进方向？
- 超分落地要考虑哪些问题？
- 如何将超分算法与视频压缩相结合？
- 超分算法中不同采样方式的对比？

### 3.1.4 目标重识别

- ReID 常用的方法？三元组损失怎么训练？

### 3.1.5 音频算法

- 用到的模型，以及对 Kaldi 和端到端的了解？
- 在语音降噪过程中，怎样实现噪声估计？
- 使用模型降噪的原理？
- 语音识别用到的模型？
- OMLSA？MCRA 用过吗？
- FIR 和 IIR？
- 16K 重采样到 8K，怎么做？
- MFCC 的原理？
- 频率分辨率
- 如何做卷积？

## 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、GPT、ELMO 之间的区别？（模型结构、训练方式）

#### ② CRF

- CRF 的原理讲一下？

#### ③ Word2vec

- 讲一下 word2vec 怎么实现？
- Word2vec 和 bert 区别？
- word2vec 的缺点，word2vec 的输出是什么？
- Word2vec，负采样，层次归一化？

#### ④ 其他

- 画出 fasttext 的网络结构，描述其在分类和 embedding 时的区别。详述训练过程
- 介绍方面级的情感分析模型？情感分析任务用哪个数据集？
- tfidf 的计算公式？
- 如何进行句子编码，提取句子的特征向量，有哪几种方式（CNN，LSTM，Attention），各种方式的优缺点？

## 3.3 强化学习

### 3.3.1 讲解原理

- 用过的 CycleGAN 介绍一下
- 了解 GAN 吗，简单介绍一下？
- GAN 在文本生成中如何应用？
- 如何解决 GAN 中文本离散的问题？

- 如何解决 GAN 中生成器与判别器训练不平衡的问题？
- 为什么用 PPO 算法？
- 讲一下 A3C 异步效果为什么可能不收敛？
- 提升样本使用效率还有什么方案？

### 3.3.2 损失函数

- 介绍 GAN，实际实现中的 loss 是什么？

## 3.4 机器学习方面

### 3.4.1 推荐系统

- 讲一下协同过滤的原理？
- FM 公式？
- deepFM 的原理？

## 4 数据结构与算法分析相关知识点

第四节  
网易面经  
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

- 在乱序数组当中找出中位数？
- 求数组的最长子集，整个子集满足所有的元素两两模除为 0？
- 两个一样的数组，其中一个数组插入了一个数，如何找到它的位置？

- 一个二维数组从左到右，从上到下都是增大的，找一个数，又问了时间复杂度？
- 数组有  $n$  个整数，每次对  $n-1$  个数全部加 1,求最少几次让所有的数相等？
- 16 进制转 10 进制，最好用 C 语言来写，实在不行也可以用 python。

#### 4.1.1.2 链表

- 环形链表 判断有无环
- 找环的入口
- 两条链表求第一个公共节点？
- 两个不等长链表的公共节点？
- 合并两个有序链表？

#### 4.1.1.3 字符串

- 翻转字符串
- 给个字符串，返回最长无重复的子串？
- 找出字符串的所有全排列？
- 求最长匹配括号的长度？
- 给定一个字符串，只保留  $k$  位，不改变字符间的顺序，使得字符串字典序最小？
- 给定字符串，求长度为  $k$  的字典序最小的子序列？

#### 4.1.2 树

- 哈夫曼树相关

#### 4.1.3 排序

- 排序算法了解吗？快速排序的时间复杂度怎么样呢？快速排序是稳定的吗？有哪些稳定排序算法呢？
- 解释一下排序的稳定性，冒泡排序是否稳定，复杂度多少？
- Topk 问题
- 有足够多的数据（内存无法一次性装下），如何获得最大的  $k$  个数？

- 写一下快排，讲一下最好和最坏的情况？
- 写快排里的 partition 函数？
- 介绍堆排序？
- K 个最大值：堆排序

## 4.2 算法思想实战及智力题

### 4.2.1 算法思想实战

- 洗牌算法，有多少种可能性以及数学证明（很简单，Knuth 洗牌算法）
- 图找环方法，有向图和无向图找环区别？

### 4.2.2 智力题

- 问三个囚犯一碗汤,每次都分的不均不开心，设计分法让大家都满意？

## 4.3 其他方面

### 4.3.1 数论

- 两枚硬币,依次掷，两枚掷完算一次，到第五次两枚第一次同时出现相同的面的期望值？
- 如何在一堆数里确定是否存在某几个数？

### 4.3.2 计算几何

- 数轴上某些位置有点，每个点都有一个速度和方向（左或右），在零时刻他们开始运动，求第一次有两点相碰的时间？如果只有相反方向的相碰才算，如何求解？

### 4.3.3 概率分析

- 10000 个黑球、10000 个白球，混合在一个桶里，无放回的取两个球，异色放白球，同色放黑球，求最后一次是黑球的概率？
- 54 张扑克牌，分三堆，其中 4 张 A 在同一堆的概率？

- 流数据  $n$  个中随机取  $k$  个数，每次只能取一个，怎么使取到每个的概率相等？
- 27 个球，有一个轻的，找出来最少需要几次？
- 一条绳子切两刀得到的三段线组成三角形的概率？

#### 4.3.4 其他

- 给出  $n$  个点的坐标  $(x_1, y_1) \dots (x_n, y_n)$ ，找出其中离  $(a, b)$  最近的点，要求  $x_i, y_i, a, b$  的数字动态变化，实现高频查找？
- 给定一个长度为  $n$  的序列，将其分割成若干连续子序列，若这些序列构成的数能整除  $m$ ，输出相应结果？
- 有序列表合并？
- $N$  的阶乘后面有几个零？
- 求  $0 - N-1$  的全排列输出？

### 4.4 Leetcode&剑指 offer 原题

- Leetcode 33：搜索旋转排序数组
- Leetcode 518
- Leetcode 原题：射气球
- 剑指 offer 原题：二维向下向右递增的矩阵查找

## 5 编程高频问题：Python&C/C++方面

第五节  
网易面经  
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

- Pytorch 和 TensorFlow 有什么区别？

#### 5.1.1.2 Tensorflow 相关

- 手写了 tensorflow 的图像分类代码？
- tensorflow 有哪些缺点？

### 5.1.2 基础知识

#### 5.1.2.1 线程相关

- python 多线程有多少了解？和 c 的多线程最大区别是什么？
- python 多线程缺点？

#### 5.1.2.2 内存相关

- Python 怎么做内存回收？

#### 5.1.2.3 区别比较

- Python xrange 和 range 差别？
- 装饰器及多进程和多线程区别？
- Python 浅拷贝和深拷贝有什么区别？

#### 5.1.2.4 讲解原理

- Python yield 关键字是什么用的？
- 问了 Python 怎么加速？(np 矩阵乘法替代循环)
- Python 的命名规则、self、lambda、with、循环引用

## 5.2 C/C++方面

### 5.2.1 基础知识

#### 5.2.1.1 内存相关

- C++的栈区和堆区知道吗？分别是干什么用的？（栈区是存储函数内部变量的内存区，堆区是存动态申请的内存）
- 什么时候要进行动态内存申请？（以前没思考过，没答上来，后来查了一下，当无法事先确定对象需要使用多少内存（这些对象所需的内存大小只有在程序运行的时候才能确定）时就要申请动态内存，比如维护一个动态增长的链表或树）
- 栈和函数调用的关系？

#### 5.2.1.2 区别比较

- C++和 Python 的区别？
- 栈和堆的区别？

#### 5.2.1.3 讲解原理

- 多态，虚函数
- 为什么 C++比 Python 快？
- Python 比 C++好在哪里（自动能实现内存回收机制）
- C++虚函数（如何实现，有什么功能）、虚函数表
- 类实例化的机制，具体实现是什么？
- stl 容器分类？
- 析构函数的基类为什么是虚函数？

### 5.2.2 手写代码相关

- 16 进制转 10 进制，最好用 C 语言来写？
- vector 排序

## 6 操作系统高频问题：数据库&线程&常用命令等

第六节  
网易面经  
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

- 进程和线程区别？

#### 6.2.2 常用命令

- 如何查看某进程关联的相关文件有哪些？

## 7 技术&产品&开放性问题

### 7.1 技术方面

- 不平衡数据的解决方式，数据分布改变了怎么办？
- 直播中如何判断人眼关注点区域？
- 视频通话场景如何估计噪声？
- h.264 压缩优化？
- 图像复原与图像增强的区别与联系？
- 如果有一百万个游戏片段，仅有少数有标记，如何利用这些数据？

- 如果数据的维度很高（3 万），如何完成聚类？
- 100W 个起始结束 IP 段以及对应中文名，建一个系统，让它可以很快查找出某个 IP 对应的中文名？
- 开放题：如何设计一个人脸检索系统（从数据、模型、loss 等考虑）
- 开放题：对 CV 类前景的看法

## 7.2 产品方面

- 场景题：根据游戏用户反馈的问题，进行信息分类与关键词提取，给出一个比较详细的综合方案（从技术到实际运行）

## 7.3 开放性问题

- 假如面试官是一个零基础的深度学习学习者，那么有什么建议，如何入门？
- 如果某个事业部想要销量提升 20%，作为数据分析人员，给什么方案？