

# 11|招银网络算法岗武功秘籍

## 1 招银网络面经汇总资料

第一节  
招银网络面经  
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 招银网络面经整理心得

### 1.1 面经汇总参考资料

#### ① 参考资料:

- (1) 牛客网: 招银网络面经-49 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

【机器学习算法实习岗】

- (2) 全职岗位类

【数据分析工程师】、【大数据研发工程师】、【NLP 算法工程师 (客服机器人、金融风控相关)】

### 1.3 面试流程时间安排

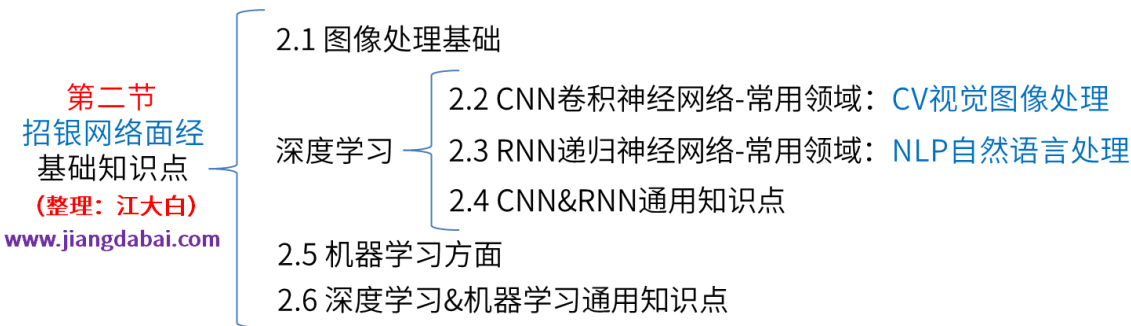
招银网络面试流程-整理：江大白			
	面试类型	面试流程	备注（侧重点）
第一面	技术面	自我介绍+项目/实习经验 +技术问答+算法编程	项目技术为主
第二面	技术面	自我介绍+项目/实习经验 +技术问答+算法编程	专注项目和领域方向
第三面	HR面	基础人力问题	/

**PS：**以上流程为大白总结归纳所得，以供参考。

### 1.4 招银网络面试心得汇总

- ★ 总体来说，感觉题目都比较基础，不过因为是开发的小哥面试的我，所以机器学习和深度学习的问题没怎么问。
- ★ 针对项目问，问得比较宽泛，算法题有的人会考

## 2 招银网络面经涉及基础知识点



### 2.1 图像处理基础

边缘检测的原理说一下？

## 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

- 讲讲 CNN 的各个层，卷积层是怎么实现的？

- 卷积核大小如何选择？

答：大卷积核和小卷积核：大卷积核可分解为多个小卷积核，减少计算量，同时 inception 将  $n \times n$  小卷积核又分解为  $1 \times n$  和  $n \times 1$  的卷积，最后 concatenate。面试官第一次追问：大卷积核为什么不用于高层。这里没答好。只说了高层大卷积核可能没用，因为没有那么大的特征图来提取，多数是 padding 区域。面试官：还有嘛？第二次追问，卷积核能否偶数。答：不能，会使特征图产生偏移。

- CNN 卷积层的作用？

- CNN 的 padding 有什么用？

#### 2.2.1.2 池化方面

- 解释一下池化层？池化层作用？

- max pool 反向传播？

#### 2.2.1.3 网络结构方面

- 对残差网络的理解，为什么有效？

#### 2.2.1.4 其他方面

- 解释一下 BN？

- 梯度消失、梯度爆炸原因和解决方法？

## 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 的结构讲一下？为啥比 RNN 要快？
- LSTM 长短期记忆如何实现？

### 2.3.2 手绘网络原理

- 手写 LSTM 的公式（手画 LSTM 图）

## 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

- 项目中的数据处理这部分用的啥工具（楼主答 numpy 和 pandas）
- numpy 和 pandas 的实际应用场景以及哪个速度更快一些？

#### 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）

- Bagging 和 boosting 的区别？
- 随机森林和 gbdt 谁主要降低偏差和方差？
- 解释下 RF 和 GDBT？RF 的属性采样时有放回还是不放回？

##### A.基于 bagging：随机森林

- 讲讲随机森林的原理？介绍一下随机森林？
- 随机森林的重要性排序原理？

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

- xgboost 和 gbd 区别？
- boosting 了解吗？讲讲流程
- Lightgbm 对缺失值的处理方法？
- Xgboost 为何能并行化？

### ② 线性回归

- 线性回归为什么用均方误差作为损失函数？

### ③ 逻辑回归 LR

- 讲一下 LR 的原理？
- LR 和 SVM 的区别？
- LR，从二项分步，指数族开始讲了下原理，以及为什么使用交叉熵。
- LR 的缺点？

### ④ SVM（支持向量机）

- 讲讲 SVM，通过什么解原始问题？KKT？
- 核函数，你怎么理解？有哪些核函数？
- SVM 做多分类怎么做？
- SVM 的原理和目标 、SVM 如何防止过拟合？
- SVM 有什么损失函数？
- 问 SVM 与贝叶斯的优缺点？

### ⑤ 朴素贝叶斯 (Naive Bayes)

- 贝叶斯分类器的原理？

## ⑥ 决策树 (DT)

- 决策树原理，常用哪几种？
- 决策树相关问题，影响决策树效果的有哪些，开放性问题？
- Adaboost 和决策树的区别？
- 决策树的构建，决策树怎么防止过拟合(前向剪枝，后向剪枝) ？
- 说一下决策树的 c4.5？

### 2.5.1.4 无监督学习-聚类方面

- 聚类算法有哪些，密度聚类，怎么评价聚类的效果？
- KMean 了解吗？讲一下，KMean 的缺点，以及改进？
- KMeans 和 dbscan 原理，区别，如何调参？

### 2.5.2 手推算法及代码

- LR 的推导以及应用细节？
- 写出 LR 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 损失函数方面

- 常用损失函数/激活函数？

### 2.6.2 激活函数方面

- 常用的激活函数说一下？
- Sigmoid 优缺点，对比 Relu 优缺点，怎么改进？
- 激活函数的作用，加在卷积层还是池化层？

### 2.6.3 网络优化梯度下降方面

- 讲一下梯度下降具体是怎么做的(梯度下降的过程)？

- 问梯度下降算法与牛顿算法的优缺点?

## 2.6.4 正则化方面

- 正则化有哪些?
- 解释 L1、L2? L1 和 L2 正则化的区别?
- L1 和 L2 为什么一个稀疏一个平滑?

## 2.6.5 过拟合&欠拟合方面

- 什么样的情况容易过拟合?
- 除了 dropout 和正则化还有什么方法可以处理过拟合 (降低网络复杂度)?
- 过拟合的解决方案? 1.数据层面:增加数据和数据增强 2.模型层面:换用小的模型, Dropout, L1L2 正则化, 以及模型剪枝 3.模型集成以多任务学习

# 3 招银网络面经涉及项目知识点

第三节  
招银网络面经  
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

- yolov3 与 yolov4 的区别?
- 小目标检测有哪些改进方式?

#### 3.1.1.2 手写代码

- NMS 算法实现

- 为什么会需要用到 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

#### ① Word2vec

- Word2vec 如何实现，实现方法有什么区别？

#### ② 其他

- 词袋模型是什么？
- 讲解算法模型（seq2seq）以及 TextRank 算法？
- tfidf 公式？
- 基于匹配的问答系统的关键技术是什么（文本相似度匹配）
- 如何判断两个词相似？

## 3.3 强化学习

无

## 3.4 机器学习方面

### 3.4.1 推荐系统

- 基于物品的协同过滤讲一下？
- 介绍推荐系统的召回和排序系统，召回系统的输出是什么？



## 4 数据结构与算法分析相关知识点

第四节  
招银网络面经  
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

- 数组相邻元素之差绝对值都为 1，找给定元素 t 在数组中位置？
- 找出数组中超过一半的那个数？
- 给你两个有序数组，合并排序一下？
- 给定一个数组，找出能构成差为 1 的最长的等差数列的元素。除了先排序后暴力之外，有没有其他方法？
- 给一个数组，已知第一个输出的数，输出栈可能的情况？
- 数组和链表的区别？

##### 4.1.1.2 链表

- 单向链表的倒序输出（vector 再 reverse，或者存在栈里一个个 push 即可）
- 链表反转、链表判断有没有环
- 数组链表的区别？

##### 4.1.1.3 字符串

- 阿拉伯数字汉字转数字，开始以为是 int，被提醒了一下用 string
- 找出一个字符串中第 M 个出现 N 次的字符？

- 字符串查重

## 4.1.2 树

### 4.1.2.1 二叉树

- 完全二叉树和满二叉树的区别？
- 二叉树的种类，介绍平衡二叉树
- 搜索二叉树的特性？（左子树比根节点小，右子树比根节点大）
- B 树和 B+树了解吗？讲一下？
- 红黑树讲一下？
- 怎么获取二叉树的深度？我先反问二叉树有没有什么约束，比如完全二叉树之类的，他说没有。之后我就把暴力解法，平衡二叉树，完全二叉树所有的都说了一遍。
- 说一下树的遍历
- 树的深度遍历、广度遍历
- 中序遍历二叉树
- 怎么查到第 k 个大的数（中序遍历二叉搜索树）

### 4.1.2.2 堆

- 堆和栈的介绍以及它们之间有什么区别？

## 4.1.3 排序

- 常用排序算法、哪几种排序算法时间复杂度比较低、讲堆排序原理？
- 介绍一下排序算法，选择一种(除了冒泡)进行具体介绍，它是怎么操作的？
- 针对你具体讲的那种排序算法，追问你时间复杂度分析以及空间复杂度分析？
- 几种排序复杂度稳定性
- 讲一下归并排序
- 海量数据排序，我写的堆排序

- 写快排和堆排，然后比较

## 4.2 算法思想实战及智力题

### 4.2.1 算法思想实战

- 青蛙跳台阶

### 4.2.2 智力题

- 8 个砝码，一个超重，用天平最少几次称出来？（2 次）

## 4.3 其他方面

### 4.3.1 矩阵运算

- 手撕螺旋矩阵
- 给定矩形内有若干点，求距离最小的两个点的距离，不能用穷举法？
- 给一个二维矩阵和一个值，怎么判断这个值在矩阵中，矩阵很大，写出你能想到的最优方法？

### 4.3.2 其他

- 贪心算法的原理

## 4.4 Leetcode&剑指 offer 原题

- Leetcode 原题：怎样实现一个 LRU，说思路以及用到的集合框架
- Leetcode 原题：求一个字符串里面最长的回文串
- 剑指 offer 第 33 题：数组排成最大的数字

## 5 编程高频问题：Python&C/C++方面

第五节  
招银网络面经  
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

- 描述一下 TensorFlow 计算图，工作原理、用的什么数据结构？
- Tensorflow 数据流图？

#### 5.1.2 基础知识

##### 5.1.2.1 内存相关

- Python 分配内存方法

##### 5.1.2.2 区别比较

- tuple 与 list 的区别？
- Python 元组和列表的区别？
- not exists 和 not in 区别？
- 生成器和迭代器区别？

##### 5.1.2.3 讲解原理

- 知道 Python 的匿名函数吗？

##### 5.1.2.4 讲解应用

- 说一下 Python 的 yield：生成器，惰性机制
- 讲一下 Python 装饰器

- Python 的几种运算符介绍?
- Python new & init 方法
- 常用的 Python 包

## 5.2 C/C++方面

### 5.2.1 基础知识

#### 5.2.1.1 线程相关

- 多线程的切换是怎么实现的?

#### 5.2.1.2 内存相关

- 讲述一下 C++内存泄漏和内存溢出?
- 内存对齐讲一下?

#### 5.2.1.3 区别比较

- 引用和指针的区别?(会深入问为什么)
- 讲述 C++中重载和重写及区别?

#### 5.2.1.4 讲解原理

- 介绍一下构造函数，析构函数，函数重载?
- 多态怎么实现?
- 介绍一下 C++编译过程
- 索引的优缺点
- 面向对象了解吗? 讲讲主要特性?

## 6 操作系统高频问题：数据库&线程&常用命令等

第六节  
招银网络面经  
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

- 数据库索引的基本原理

#### 6.1.2 手写代码

- 一个三个字段的数据表，id,name,age，写一个 where 查询，写一个按照年龄分组求每组有多少人。直接用 case when 分组再 group by 后 count?
- sql 模糊匹配
- sql 里面用过的 join 语句

### 6.2 操作系统方面

#### 6.2.1 线程和进程相关

- 进程与线程的区别？
- 深拷贝和浅拷贝？

#### 6.2.2 常用命令

- Linux 复制命令是什么？查询帮助用什么命令？

## 7 技术&产品&开放性问题

### 7.1 技术方面

- 深挖任务型对话系统的细节？
- 多轮对话的场景保存和跳转怎么做？可以通过什么数据结构来实现？
- 怎么用知识图谱做 qa？
- 遇到脏数据怎么处理？
- 给定一段文本，如何识别出其中的省市区等地址并正确拼接
- 场景题：类似于反欺诈识别---给定用户画像，如住址（文本）--关键、年龄、职业、消费金额（num-feature）、其他信息（文本）
- 如何精确识别欺诈营销？（对文本处理不是特别了解，大致说了 kmeans 的方案，有待改进）
- 大数据处理中的数据不平衡和数据倾斜讲一下？
- 说一下信用分预测用的什么模型，说说模型的原理？（说用了 lightgbm，讲了下原理）
- 场景题：交易检测场景，拿到所有用户的交易数据，我们知道一些数据是正常的操作数据，一些是非正常操作的数据，会引发投诉，问我们怎么检测非正常操作数据？（面试官想问的是机器学习的一般步骤）
- 深度学习和机器学习，讲讲你的理解？深度学习好在哪？缺点在哪？

### 7.2 产品方面

- 银行怎么预测用户的流失率，你从哪些角度去考虑？我说的是用户持有卡的收入支出和交易频率与以往记录相比是否有下降趋势？