

# 18|滴滴算法岗武功秘籍

## 1 滴滴面经汇总资料

第一节  
滴滴面经  
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 滴滴面经整理心得

### 1.1 面经汇总参考资料

#### ① 参考资料:

- (1) 牛客网: 滴滴面经-54 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

【滴滴地图事业部算法实习生】、【滴滴算法实习生】、

#### (2) 全职岗位类

【滴滴推荐算法工程师】、【滴滴大数据算法工程师】、【滴滴新锐算法工程师】、【滴滴机器学习算法工程师】、【滴滴地图事业部机器学习算法工程师】、【滴滴定价策略算法工程师】、【滴滴国际化事业部算法工程师】、【滴滴 R-Lab 算法工程师】、【滴滴金融事业部算法工程师】、【滴滴核心岗出行算法工程师】、【网约车供需策略算法工程师】、【滴滴出行规划算法工程师】

## 1.3 面试流程时间安排

滴滴面试流程-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	项目细节和基础知识点
第二面	技术面	自我介绍+项目/实习经验 +技术问答+算法编程	项目问题更开放、 偏实际应用
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 滴滴面试心得汇总

★ 整体感觉滴滴是面试体验很好，面试官是以一种平等的态度来交流技术，即使有时候卡壳也会慢慢提醒。

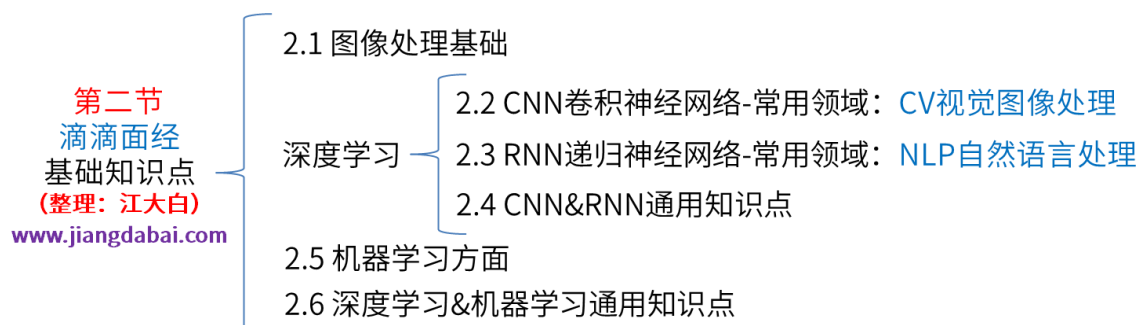
★ 有竞赛的会问竞赛项目，大部分都是围绕比赛项目问的，建议大家对自己的项目了如指掌

★ 主要机器学习为主，基础知识也会问，但不会问的特别深，计算机视觉很少

★ 总体还是不错的。除了一面面试官有点严肃，二三面都是边笑边面。尤其三面面试官，给人印象很不错，有种如沐春风的感觉。

★ 偏向于纯机器学习，我做图像方向的不太适合岗位

## 2 滴滴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

- CNN 和全连接的区别？（参数共享，降低运算量）
- 详细讲一下 CNN 怎么卷积多通道？

##### 2.2.1.2 网络结构方面

- 说一下 ResNet 和 DenseNet 的区别？为什么 DenseNet 比 ResNet 效果好？
- Autoencoder 的原理？

##### 2.2.1.3 其他方面

- 什么是梯度消失和梯度爆炸？如何解决？
- 梯度消失和梯度爆炸问了一下，缓解的方法（Relu,LSTM,BN），为什么能缓解，BN 的参数？
- Batch normalization 的原理？

## 2.2.2 公式推导

- 手推两层神经网络(input->FC->sigmoid->FC->sigmoid->output)的反向传播?

## 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 的区别?
- LSTM 原理讲一下? LSTM 的用法/具体应用?
- LSTM 有什么新的改进, 替代? (针对 rnn 来说 通过门电路把连乘转换成了连加, 一定程度上缓解了梯度消失问题, 梯度爆炸可以直接 clip 不需要考虑, 另外就是广度了, lstm 的应用和改进)
- LSTM、CNN、convLSTM 原理, 异同对比?
- LSTM 为什么可以解决梯度弥散问题?
- 连续性变量需要离散化的时候在神经网络里要怎么做?

### 2.3.2 手绘网络原理

- 手推 LSTM 公式

## 2.4 深度学习：CNN&RNN 通用的问题

### 2.4.1 基础知识点

- 样本类别不均衡你会怎么处理(重采样/Focal loss)?
- Transformer 结构是什么样的, self attention 公式是什么, 你怎么理解 self attention 的?
- 介绍 transformer?

### 2.4.2 模型评价

- 分类评价指标讲一下?

- 讲一讲常见评价指标，TPR/FPR/ROC/AUC?
- 数据不平衡用 accuracy、PR、ROC 指标会咋样?
- AUC 指标的含义，同时增大减小样本的预测概率（保持顺序不变）会对 AUC 指标产生什么影响?

## 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

- 如何做特征选择?
- 用了哪些技术做特征工程? (特征筛选+数据预处理)

#### 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)

- 模型融合方法，尝试了几种，都怎样做的?
- boosting 和 bagging 的区别? (bagging 就是投票策略，会降低方差，减小过拟合风险，boosting 是降低偏差的策略，可以描述一下 Adaboost 和 gradient boost 的区别)
- XGB 和随机森林的区别?

##### A.基于 bagging: 随机森林

- 随机森林算法讲一下?

- 随机森林的随机性体现在哪里？
- 说一下随机森林的缺点？
- RF 分裂时特征选择怎么选的，常用的还有什么选法？
- 随机森林是怎么采样的，随机森林的基本分类器都有啥？

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

- 为何使用 xgb, xgb 原理, xgb 并行实现？
- xgboost 与 gbdet、RF 的对比？
- gbdet 原理？xgboost 源码看过吗？
- xgboost 相对于 gbdet 改进了什么？
- 因为模型 xgb 用到了，所以问的很细，xgb 的原理？xgb 和 gbdet 的不同？为什么 xgb 用二阶导？xgb 怎么梯度下降的？xgb 的损失函数？xgb 的正则化？
- xgb 和 gbdet 区别？(基本是必考题目，从主要的优化点说起，xgb 是二阶泰勒展开，gbdet 是一阶，可以类比牛顿法和梯度下降法的区别，牛顿法收敛更快，但是由于更快逼近目标，会增大过拟合风险，因此在目标函数里有一个显示的惩罚项，与叶子节点数和叶子节点的权重有关，来控制模型复杂度；另外还有分裂节点的选择，xgb 怎么选取最优分裂节点，有哪些加速的优化之类的知识)
- lightgbm 相对 xgboost 的改进？
- xgboost 流程说一下？

## ② K 近邻 (KNN)

- KNN 你知道吧，里面搜索 k 近邻用的数据结构叫 kd 树，你知道吧，说一下 kd 树用在 KNN 的哪个步骤里？
- kd 树的作用是啥？
- kd 树是怎么构建的？它的时间复杂度是多少？为什么 kd 树能够比较快的找到 k 近邻？

## ③ 逻辑回归 LR

- LR 优缺点？
- 问了下 LR 的损失函数？
- 逻辑回归的损失函数，和极大似然的内在关系？
- LR 做特征重要性分析时，连续值会有影响吗，怎么处理，特征的重要是看 one\_hot 后的 sum 还是 mean？

#### ④ SVM（支持向量机）

- SVM 原理，损失函数是什么，软间隔怎么做的？
- 讲讲 SVM 吧，SVM 对错分点怎么处理的，SVM 是什么问题？
- SVM 训练完成后得到的模型是什么？

这题我简单说一下，一般人都会认为得到的模型是那个超平面，即  $y = w_1x_1 + \dots + w_nx_n$ ，我也是这样回答的。但是面试官反问如果是带核函数的模型，怎么进行预测。因为核函数是计算两个样本之间的高维内积的，因此只有超平面和另一个待预测样本的话就不知道怎么做预测了。后来面试官跟我说得到的模型其实是支持向量，预测时将待预测样本与支持向量计算距离。

- SVM 为什么自带泛化性？
- SVM 对偶化的条件是什么？具体说明每一条，并且回答为什么满足该条件就可以对偶化？
- SVM 与其他模型相比，在面对样本不均衡问题是更优还是更劣？
- SVM 的软间隔和硬间隔是什么？软间隔加入的松弛变量是如何求解出来的？

#### ⑤ 决策树（DT）

- 决策树有哪几种，平时用的多的是哪种，每种决策树有什么区别？
- 决策树如何选择特征？信息增益怎么计算？
- 决策树的流程？了解决策树剪枝么？
- 决策树和随机森林的区别？
- cart 和 c4.5 的区别？

- 离散特征经过 onehot 处理后，直接放入树模型有什么影响？

#### ⑥ 其他

- 介绍下极大似然？
- 遗传算法原理说一下、缺点、常见的编码方式？
- 树模型、NN、NB 的区别和应用场景？
- 让我介绍参加过的京东的比赛(要求按照一定的框架讲，比如做了什么，觉得自己做的好的有说什么，有什么做的不够好的)，在介绍的过程中针对性的提了一些问题，可以说都是一些常考的问题，记录如下：

- (1) 解释下过拟合；
- (2) 数据类别不均衡的处理方法；
- (3) XGBoost 相对于 GBDT 有什么不同；
- (4) 数据记录有多少？特征有多少？训练时间多久？

#### 2.5.1.4 无监督学习-聚类方面

- 介绍常用的聚类算法（KMeans、DBSCAN、Mean Shift）
- K-means 算法的过程讲一下？
- 看你项目中用到了 kmeans，这是分类算法还是聚类算法？请介绍 kmeans 与 knn 算法流程？

### 2.5.2 手推算法及代码

#### 2.5.2.1 手推公式

- 手写 LR？
- 手推梯度反向传播及 SVM 的过程

#### 2.5.2.2 手写代码

- 手写 K-means 伪代码(计算到中心点距离/判断迭代停止/更新中心点的坐标等核心函数，



串一起)

## 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

- 写了 sigmoid 函数

### 2.6.2 激活函数方面

- 有哪些常用的激活函数？
- sigmoid 函数做激活函数有哪些缺点？

### 2.6.3 网络优化梯度下降方面

- 常用的优化算法？
- 牛顿法和梯度下降法的区别？
- 牛顿法有什么优缺点？
- 梯度下降的更新公式，和梯度提升的区别？

### 2.6.4 正则化方面

- L1 正则和 L2 正则的目的和原理讲一下？
- 问了下 L2 正则会不会影响 adam 的学习率？

### 2.6.5 过拟合&欠拟合方面

- 你遇到过拟合么？什么是过拟合？有什么体现？
- 有哪些抑制过拟合的方法？
- 模型发生过拟合了，怎么判断它是数据量不够还是模型复杂了？

### 2.6.6 其他方面

- 样本不平衡怎么做？

## 3 滴滴面经涉及项目知识点

第三节  
滴滴面经  
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

- Fast-RCNN 和 Faster-RCNN 的区别是啥?

#### 3.1.2 图像分割

- 你说你是做医学图像分类与分割的, 说一下你用的经典的图像分割网络(回答得是 UNet 系列)
- 问为啥 UNet 在医学图像分割上面效果比较好?

### 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 的输入是什么?

##### ② Transformer

- 问了 textcnn、transformer 的细节?

##### ③ CRF

- 为何使用 biLSTM+CRF 的结构?

##### ④ HMM 隐马尔科夫模型

- HMM、MEHMM、CRF 原理以及对比?

- EM 算法的原理，在 HMM 中如何使用的？

### ⑤ Word2vec

- 介绍 word2vec 2 种实现模型、区别以及和 fasttext 的区别？
- word2vec 的原理讲一下？

## 3.3 强化学习

- 了解强化学习么？
- Q 值和 V 值的联系和含义？
- 强化学习 OFF POLICY 和 ON POLICY 的区别？

## 3.4 机器学习方面

### 3.4.1 推荐系统

- 解释了一下 FM、DeepFM 方面推荐的基础模型？
- node2vec 怎么做的？
- deepFM 相对 wide&deep 改进？
- fm, ffm, deepfm 区别？

## 4 数据结构与算法分析相关知识点

第四节  
滴滴面经  
数据结构与算法分析  
(整理: 江大白)  
[www.jiangedabai.com](http://www.jiangedabai.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，要求把奇数放在数组前半部分，偶数放在数组后半部分？并且奇数部分和偶数部分均有序。
- 把一个数组分成  $k$  组，返回所有可能的组合？
- 找到数组中每个数的右边比他大的第一个数，要求时间  $O(n)$ ？
- 两个有序数组的中位数
- $\log n$  找出不出现的最小正整数
- 二维数组最短路径，存在障碍的最短路径？
- 旋转数组中的查找
- 有序数组找和目标值相等的最左边的下标
- 两个有序数组，求第  $k$  大，时间复杂度尽可能低 ( $\log(m+n)$ )
- 二维数组的旋转打印，比如 1 2 3 8 9 4 7 6 5 从 1 到 9 打印出来？
- 多个有序数组求交集？
- 数组和链表的区别？
- 多个有序数组求并集？
- 数组逆序对
- 给一个数组的字符串，然后求 shape。如：“[[1,2,3],[1,2,3]]”，结果为(2,3)。

#### 4.1.1.2 链表

- 反转链表？
- 单链表的排序，要求时间要达到  $n \cdot \log(n)$
- 两个链表合并

- 链表题，奇数位升序，偶数位降序，重排成升序的链表

#### 4.1.1.3 字符串

- 给一个字符串和一个 k，要求找到不超过 k 个不同字符的最长子串的长度，on 时间（LeetCode 上的 medium 或者是偏简单的 hard 题目，用划窗的思路做）
- 求和大于 k 的子串的最小长度？
- 最长公共子串问题
- 给你若干个 list，从每个 list 中选出一个元素，求全排列？
- 动态规划问题，一个长的字符串中找到最长的没有重复字母的字符串？
- 括号匹配问题，给一个不匹配括号串，只包含左括号和右括号，这个串去掉一个字符可以变成匹配串，找到所有可能的匹配串

#### 4.1.2 树

##### 4.1.2.1 二叉树

- 求树深，时间复杂度（LeetCode easy，时间复杂度是  $O(N)$ ，因为要扫一遍所有的节点）
- 给出二叉树的先序遍历和中序遍历，输出后序遍历？
- 之字打印二叉树

##### 4.1.2.2 堆

- 堆排序算法过程？

#### 4.1.3 排序

- 归并排序求逆序对个数，并记录每个元素对应的逆序对个数
- topk
- 非递归快排
- 手写了一个快排

- 实现一个堆排序

## 4.2 算法思想实战及智力题

### 4.2.1 算法思想实战

- 一亿个结点，每个结点有个 score 和一个 class，按照 score 大小要把排名前 10% 的结点 class 变为 0, 10~20% 的变为 1，以此类推？
- 有  $m$  件物品，每件物品分别有不同的重量是一个  $m$  大小的数组， $n$  个背包 1~ $n$  编号，每个背包的容量为  $T$ ，把物品往背包里面装，一个背包在容量运行的范围内可以装多个物品，但是装背包必须按照背包的编号来，并且开始往第二个背包里面装物品之后就不能再往前面的背包里面装了，问最多可以装多少件物品。（类似买卖  $k$  次股票的问题）
- 数三减一，就是小朋友抱成圈，数到三删除一个人，考代码实现能力
- 国王挖金矿的变体
- 最长递增子序列
- 参数量估计：绳子剪成  $m$  段，最大乘积
- 因式分解
- 判断年份是否为闰年
- 开三次方根。其实跟开根号的题类似。牛顿法可能要推一下公式，可以用二分法。

### 4.2.2 智力题

- 一条路上，总长度为  $n$ ，有长度为 1、2、3、5 的砖，问总共有多少种铺法？DP 求解。
- 1000 桶水有一桶有毒，一头猪喝了有毒的水之后过 15 分钟会死掉，问最少需要多少头猪来找到有毒的水，在一个小时之内？

## 4.3 其他方面

### 4.3.1 数论

- 讲解泰勒级数，傅立叶变换，复变函数（因为主管是学数学出身）
- 判断质数？

### 4.3.2 概率分析

- 2 人抛硬币，正面赢，反面让另一个人抛，求先抛的人获胜的概率？
- 抛硬币，连续两次正面向上为止，求次数期望
- 三个门，门后一个有礼物，你选了一个，帮你排除了一个没有礼物的门，问你换不换：换，假设第一门有礼物，计算换和不换拿到礼物的概率？最后换的概率是  $2/3$ ，一开始我直觉上以为不用换，差点就说出来不换了，还好算了下概率
- $n$  个人，其中一个是明星其余为普通人，明星不认识普通人，普通人都认识明星，普通人之间是否认识未知。现给一个黑盒，可以询问一次 A 是否认识 B，问使用黑盒几次能找到明星？

面试官提示了下，最终算出了  $n-1$  次，推导如下：A 不认识 B，那么 B 不可能是明星；A 认识 B，A 不可能是明星，每次询问必能淘汰一个人，所以最后需要  $n-1$  次

### 4.3.3 矩阵运算

- 矩阵中给定两个点作为对角线组成的矩阵里的数的和（复杂度  $O(1)$ ，可简单预处理）？
- 序列【1, 0, 1, 0, 0, 1.....】1 可以变成 0，0 可以变成 1 求最少多少次操作可以让 0 后没有 1？
- 矩阵，只能向右或向下，求最大路径 时间空间复杂度（LeetCode easy 入门级的动态规划，时间复杂度  $O(M*N)$ ，空间复杂度  $O(N)$ ）
- 动态规划问题，一个二维矩阵，第一行一个数字，第二行两个，第  $N$  行  $N$  个，求从头到尾最大值为多少？

## 4.4 Leetcode&剑指 offer 原题

- Leetcode 64
- Leetcode 85
- Leetcode 104
- LeetCode 原题：铺方块（斐波拉契数列）
- Leetcode 原题：蓄水池抽样

## 5 编程高频问题：Python&C/C++方面

第五节  
滴滴面经  
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 的一些代码习惯(你是写 session.run 还是 estimator 啊, 有什么区别/优缺点)?
- 你常用的框架有哪些?

#### 5.1.2 基础知识

##### 5.1.2.1 内存相关

- python 深浅拷贝

##### 5.1.2.2 讲解原理

- python 传参
- python 垃圾回收, 如何分代



- python 魔术方法
- python 中的反射
- python 字典底层怎么实现的

### 5.1.2.3 讲解应用

- 怎么使用 map 函数？
- 说说 lambda？
- 怎么得到 list 中的 index、value 值？

### 5.1.3 手写代码相关

- Python 实现 one-hot 特征？

## 5.2 C/C++方面

### 5.2.1 基础知识

- 重载和重写的区别？

### 5.2.2 手写代码相关

- 两个线程同时访问同一段代码程序？

## 6 操作系统高频问题：数据库&线程&常用命令等

第六节  
滴滴面经  
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

- awk 实现 left join

## 6.2 操作系统方面

### 6.2.1 TCP 协议相关

- 三次握手四次挥手

### 6.2.2 线程和进程相关

- 线程和进程的区别？

## 7 技术&产品&开放性问题

### 7.1 技术方面

- 假如有  $m$  个司机和  $n$  个乘客，每个乘客有自己满意的若干司机，让你去配对司机与乘客，让配对权值最大。这个问题从你们控制专业常用方法角度如何解决？
- 如果一个模型最后效果不好，你会从哪些方面来考虑？
- 定价策略算法岗问题：
  - (1) 预测一个城市不同区域一天的发单量，和滴滴业务相关，设计一些调度策略
  - (2) 预测一个城市不同区域的降水量（这个可以作为前一道题的特征）
  - (3) 判断用户感知是否顺路，给出思路（怎么获取训练数据，怎么去判断用户的感知到底顺不顺路，找一些简单的规则）
- 特征工程的题目：假如要预测一段时间内某个地方的出租车需求量，写下你认为的影响因素，然后想象怎么影响？
- 离散型属性工作性和连续性属性的优缺点？连续性属性离散化的好处和坏处？（好处就是可以维度扩展，从而可以训练非线性模型，坏处就是容易过拟合，过拟合了之后怎么解决？计入正则项，或者去看离散化之后的特征，哪个特征所包含的样本太少的话就把这维特征去掉，去噪声，在特征的数量上进行制约）

## 7.2 产品方面

- 场景假设：预判滴滴乘客接单等待时间：需要考虑的那些因素，特殊时间天气特俗考虑
- 场景题，如何判断突发性异常事件（如追尾等）

## 7.3 开放性问题

- 给定北京周边各个地方的历史天气状况，来预测当天各地的天气状况？

我当时从这几方面答的：

（1）每个地方根据历史数据+当天的天气预报相关特征+时间点（早高峰）+周几（是否周末、节假日）+PM2.5+可见度+是否下雨等等训练预测模型，我当时考虑是采用朴素贝叶斯去做

（2）不同地方之间的天气状况会相互影响，我当时想的是根据历史数据构建不同地方的互相影响的贝叶斯网络

- 在学校建 4 个食堂 保证 4 个食堂午餐的流量基本相等 怎么选址？