

19|蚂蚁金服算法岗武功秘籍

1 蚂蚁金服面经汇总资料

第一节
蚂蚁金服面经
汇总资料
(整理: 江大白)
www.jiangdabai.com

- 1.1 面经汇总参考资料
- 1.2 面经涉及招聘岗位
- 1.3 面试流程时间安排
- 1.4 蚂蚁金服面经整理心得

1.1 面经汇总参考资料

① 参考资料:

- (1) 牛客网: 蚂蚁金服面经-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 面经框架及参考答案:

- (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】、【机器学习算法工程师】、【人工智能算法工程师】

1.3 面试流程时间安排

蚂蚁金服面试流程-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+延伸的基础知识
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+延伸的基础知识
第三面	技术面	自我介绍+项目/实习经验 +技术问答+算法编程	简历上的项目 也会问开放性的问题
第四面	交叉面	自我介绍+项目/实习经验 +技术问答+算法编程	简历上的项目 也会问开放性的问题
第五面	HR面	基础人力问题	/
第六面	杭州终面	ppt+问答	/

PS：以上流程为大白总结归纳所得，以供参考。

其他注意点：

- 感觉实习是四面：技术+技术+技术交叉+技术&HR 混合
- 真正工作是六面：技术+技术+技术+技术交叉+HR+杭州终面
- 不过有的人也和实习那个一样，是四面，最后技术大佬+HR 是压力面
- 第一面技术面，有的人不是现场面，而是电话面
- 最后的杭州终面，一般是两个技术 leader (P9) 和 一个 HR，加上两个工作人员，半小时左右。

1.4 蚂蚁金服面试心得汇总

- ★ 项目的熟悉程度，其中涉及的知识点，基础知识，比如机器学习，深度学习问得很多。
- ★ 整个流程下来感觉蚂蚁非常看重的是你对于自己研究领域的思考和实践，比较注重

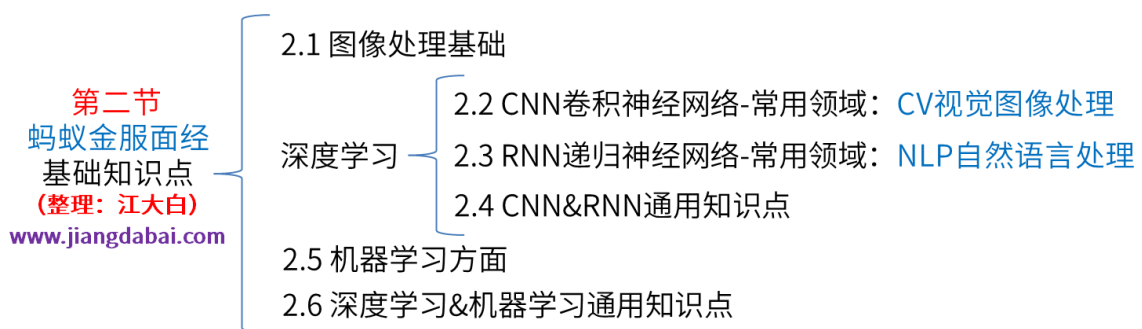
你是否具有较高的学习热情和个人潜力

★ 蚂蚁金服面试下来的心得就是，一定要梳理好自己以往的项目，不要随意跳槽，每份工作如果不是特殊情况，尽量干满两年，简历太花可能连简历关都过不了。

★ 面试总体来说，我感觉阿里的面试包容性比较强，原意去挖掘你的优势，目前我还没遇到手撕代码，更多的是对基础的考察。

★ 很多同学交流反馈的不完全统计，阿里面试一面难度最大，之后的面试好多都是只聊简历项目。阿里的 HR 权利很大，即使面到了最后一面 HR 面还是有很大可能会被挂掉。

2 蚂蚁金服面经涉及基础知识点



2.1 图像处理基础

- Canny，关键点是如何提取的，关键点是如何匹配起来的？
- SIFT 原理讲一下？如何保持尺度不变性的？

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

- 知道感知野吗？什么作用？
- 卷积的作用吗？

- 你知道 depthwise-CNN 吗？讲讲具体原理？那 1×1 的 kernel 的作用是什么呢？对网络 model 有什么影响？

- Dropout 怎么实现？

- Dropout 使得训练和测试输出不一致，怎么处理？

2.2.1.2 池化方面

- Pooling 的原理？

- 有哪些下采样方式？max pooling, average pooling, global average pooling。再问分别有什么用？

答：max pooling 我蠢到说提取最有特征的特征，其实就是最具有代表性的特征；average pooling 提取的是比较 general 的特征；global average pooling 用来分类的，因为后面网络的加深，full connected layer 参数太多了，不容易训练，为了快速准确得到结果，采用 global average pooling，没有参数，但是得到的分类效果跟 FC 差不多。

2.2.1.3 网络结构方面

- 为什么说 SENet 泛化性好？SENet 为什么效果好？

- 讲一下神经网络的发展？vgg, resnet, densenet 和 inception v1-4

2.2.1.4 其他方面

- BN 原理？BN 是解决什么问题的？BN 为什么有效？

- 说一下视觉的 attention？

- 梯度消失怎么造成？什么导致变得平缓，哪些函数会导致梯度消失，为什么？

- 解决 Overfitting、regulation 的方法？

(1) Dropout

介绍 dropout 的概念，问了下 train 和 test 阶段的不一样过程细节

主要讲了下 test 把 active function 的输出都乘以 P，这样就把 train 和 test 的输出期

望都 scale 到了一个 range，这样才是更加准确的

(2) Batch Normalisation

BN 的原理，问什么好，好在哪里？

降低了样本之间的差异，scale 到 (0, 1)

降低了层层之间的依赖，主要体现在前一层的输出是下一层的输入，那么我们把所有数据都 scale 到了 (0, 1) 的 distribution，那么降低了层层之间的依赖关系，从而使数据更加准确

● 说一下过拟合怎么造成和解决办法？ dropout 的原理，怎么抑制过拟合？预测阶段能不能 dropout？

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 模型, LSTM 模型讲一下？

● Transformer 与 LSTM 区别？

2.3.2 手绘网络原理

● 画出 LSTM 模型细节以及四个门计算公式？

2.4 深度学习：CNN&RNN 通用的问题

2.4.1 基础知识点

● 不平衡的 label 怎么办？

2.4.2 模型评价

● 问怎么衡量模型优劣？答 testing error 和 F1 Score。再问讲一下 F1 score 的计算方法？

● AUC、ROC 曲线了解吗？这两个曲线的应用场景是？

- AUC 曲线、precision 和 recall 分别在什么情况下使用？
- AUC 一般用于哪些地方，map 和 AUC 区别在哪里什么时候用？
- F1-SCORE 和 Auc 的区别，F1-SCORE 比较好或者 AUC 比较低是什么情况？

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

- 处理/清洗过数据吗？

2.5.1.2 特征工程

① 特征降维

- 为什么决策树之前用 PCA 会好一点？
- 具体讲一下 SVD 原理，除了降维还有什么作用？SVD 分解后怎么用于电影评分？
- 度量分解矩阵的原理？

② 特征选择

- 对于一般的问题你是怎么样处理特征的？
- 高维数据，其中有一维是时间，有缺失，如何处理？

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）

- 随机森林和 GBDT 的区别？

A. 基于 bagging：随机森林

- 随机森林的原理？
- 问 random forest 随机在哪里？
- 为什么随机森林要取 2/3 的样本构建树？

- 随机森林和 GBDT 的区别?
- 谈谈随机森林? 为什么项目选择随机森林, 不选择 GBDT?
- 为什么随机森林的树比 gbdt 的要深一点?
- 怎么使用随机森林做特征多样性选择?

B.基于 boosting: Adaboost、GDBT、XGBoost

- 问了解 XGBOOST 吗? 讲一下它和 GBDT 的区别?
- 说一下 GBDT 的详细过程? GDBT 如何做分类和回归?
- 从方差偏差的角度解释 bagging?为什么随机森林泛化能力强?
- 讲一下 boosting 算法? 说一下 adaboost 是怎么更新的?
- XGBoost 或 GBDT 中怎么控制模型复杂度? XGBoost 的参数选择 (怎么调参)

② 逻辑回归 LR

- LR 中参数的意义?
- 如何检验, p-value?LR 与线性回归比较? LR 的前提假设?
- 逻辑回归原理、手写损失函数?

③ SVM (支持向量机)

- SVM 和 Logistic Regression 对比?
- SVM 和 LR 的区别?
- SVM 适合什么样的场景?或者说有什么限制?
- 说下 SVM, 核函数?
- 介绍一下 hinge 损失函数?

④ 朴素贝叶斯 (Naive Bayes)

- 贝叶斯公式知道吗, 什么含义?
- 朴素贝叶斯和决策树的差别, 各有什么缺点? 再加上 SVM 呢?

⑤ 决策树 (DT)

- 决策树知道吗？
- 决策树分裂节点的选择？
- 讲一下决策树的优缺点？（优点说了构造比较简单，比较快。然后聊了 RF,GBDT,XGBoost）

2.5.1.4 无监督学习-聚类方面

- 如何确定 k-means 的 k？
- 聚类算法？ k-means 的缺点？

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

- 使用过哪些损失函数？
- 损失函数怎么设计？平方差损失优缺点？什么时候用交叉熵和平方差区别什么时候用？
- 做损失函数的时候，欧式距离和余弦距离的区别？
- Cross entropy loss 交叉熵损失与均方差损失（Mean Squared Error Loss）的差别？

2.6.2 激活函数方面

- Relu 小于 0 的时候也会导致梯度消失怎么办？怎么保证输入不是负数？

2.6.3 网络优化梯度下降方面

- 用什么优化方法，梯度下降的种类，各有什么优点？
- 神经网络为啥不用拟牛顿法而是用梯度下降？
- 说下从 sgd 到 adam 的思路？
- Adam 优化器是怎么工作的？

可以看做是 RMSProp+Momentum

具体说来，就是它使用了一个梯度平方的滑动平均（同 RMSProp）

然后使用的是平滑版的梯度 m ，而不是原始梯度 dx ，积攒历史梯度（同 Momentum）。

2.6.4 正则化方面

- L1 和 L2 的区别？从贝叶斯估计的角度看，他们先验分布是什么？
- L1、L2 正则化原理，为什么 L1 产生的权重稀疏？
- 什么时候用 L1，什么时候用 L2 正则，有什么区别？

2.6.5 过拟合&欠拟合方面

- 有没有遇到过拟合？说一下什么是过拟合，解决方案有哪些？

2.6.6 其他方面

- 距离公式的局限性和改进？什么时候用，优缺点？
- 各种距离公式适用性？欧式距离，余弦距离什么时候用？

3 蚂蚁金服面经涉及项目知识点

第三节
蚂蚁金服面经
项目知识点
(整理: 江大白)
www.jiangdabai.com

- 3.1 深度学习：CNN卷积神经网络方面
- 3.2 深度学习：RNN递归神经网络方面
- 3.3 强化学习方面
- 3.4 机器学习方面

3.1 深度学习：CNN 卷积神经网络方面

3.1.1 目标检测方面

3.1.1.1 讲解原理

- RoI Pooling 如何操作？
- 如果有很长，很小，或者很宽的目标，应该如何处理？

- FPN 具体是怎么操作的?
- FPN 的特征融合具体是怎么做的?
- FPN 的特征融合为什么是相加操作呢?
- Soft-NMS 是如何操作的?
- Softmax 是什么，公式是什么样的?
- Softmax 的梯度是什么?
- 对 SSD 和 YOLO 有没有什么了解?
- 对 SSD 做过什么实验吗?
- 介绍一下目标检测有哪些方向，最近的一些新的进展？、
- FPN 是怎么提升小目标检出率的?
- 大目标如果有 2 个候选框和 gt 重合应该怎么处理?
- 分类和定位的不一致具体是什么?
- Faster-RCNN 比起 RCNN 有什么改进的地方?

3.1.1.2 损失函数

- Focal loss 怎么操作的?

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 Bert

- Bert 与 Word2vec 区别?
- Bert 和 Elmo 在工程中存在的一些 Trick?

② Transformer

- 说一下 Transformer 的结构
- Transformer 中的前馈神经网络的构建的两种方式有什么区别(一个传统的 FFNN,

一个是基于卷积的，第二种这个根本没有了解过)

③ CRF

- CRF 原理?
- CRF 怎么接在 LSTM 后面?

④ HMM 隐马尔科夫模型

- 介绍一下 HMM?

⑤ Word2vec

- 讲一下 Node2vec?
- Word2vec 的 Loss 及原理?

⑥ 其他

- 了解 seq2seq 模型? 谈谈这个模型是怎么做的? attention 是如何实现的? self-attention 了解过吗?
- 你知道 sequence to sequence 吗?

3.3 强化学习

无

3.4 机器学习方面

3.4.1 推荐系统

- FM 模型, FFM 模型, WIDE&DEEP 模型&DEEP FM 模型的区别?
- Deep FM 重点询问了细节, 画出模型结构图?
- 讲一下 FM 算法和矩阵分解的区别? 具体怎么组合? FFM 和 FM 的区别?

4 数据结构与算法分析相关知识点

第四节
蚂蚁金服面经
数据结构与算法分析
(整理: 江大白)
www.jiangdabai.com

- 4.1 数据结构与算法分析：线性表、属、散列表、图等
- 4.2 算法思想实战及智力题
- 4.3 其他方面：数论、计算几何、矩阵运算等
- 4.4 Leetcode&剑指offer原题

4.1 数据结构与算法分析

4.1.1 线性表

4.1.1.1 数组

- 给定一组超大数据的数组，该数组内的数据无法完全存入计算机中，问如何得到第 k 个小的数？
- 给定一个二维数组，每一行递增、每一列递增，判断目标值 target 是否存在？

4.1.1.2 链表

- 升序链表找中位数？

4.1.1.3 字符串

- 如果一个 list 里有空的字符串，怎样最快的速度把空字符串筛去？

4.1.2 树

4.1.2.1 二叉树

- 打印从根结点到叶子节点的所有路径，每条路径一行打印，最后递归方式写完？
- Z 字形打印二叉树？

4.1.2.2 堆

- N 个有序数组的排序，最后输出一个数组，采用最小堆算法解决？

- 一个文件（大小未知），每一行是一个数值，如何找出数值最大的前 100 个数？要求用大顶堆或小顶堆实现，讲实现过程？

4.1.5 排序

- 快排的时间复杂度和空间复杂度？平均多少？最差多少？
- 还有那些排序的时间复杂度是 $O(n\log n)$ ？知道排序中的稳定性吗？
- 三路快排？

4.2 算法思想实战及智力题

无

4.3 其他方面

- 不用 math 中的取平法差，判断一个自然数是不是可以开方（时间复杂度尽量低）？
- 动态规划：最长公共子序列？

4.4 Leetcode&剑指 offer 原题

- Leetcode139：单词拆分

5 编程高频问题：Python&C/C++方面

第五节
蚂蚁金服面经
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

- Python 多线程有了解吗？
- Pytorch 的卷积是如何实现的？

5.2 C/C++方面

- 写过多线程吗？

6 操作系统高频问题：数据库&线程&常用命令等

第六节
蚂蚁金服面经
操作系统高频问题
(整理: 江大白)
www.jiangdabai.com

6.1 数据库方面：基础知识、手写代码相关

6.2 操作系统方面：TCP、线程&进程、常用命令相关

6.1 数据库方面

无

6.2 操作系统方面

- 多线程和多进程的区别？

7 技术&产品&开放性问题

7.1 技术方面

- 开放问题：高维数据，每个特征有标签，如何对特征进行聚类/类别划分？
- 场景题：人脸识别如果人脸图像比较少，如何用小样本学习的方法去做？
- 场景题（基于字幕文本，如何把长视频剪切成 N 个短视频）

7.2 产品方面

- 开放性问题：给你一些用户每天的相对位置信息，怎么区分他们的职业？怎么判断上线的现金贷产品的盈利能力？