

## 阿里巴巴人工智能数据库招聘

阿里巴巴智能数据库招聘正式员工，博士后和实习生。机器学习与数据库系统方向，面向双十一的技术挑战，海量数据与大规模业务场景，研究与实际系统相结合，给你充分发挥的大舞台，博后解决北京户口。要求精通机器学习/优化算法。请联系 [tieying.zhang@alibaba-inc.com](mailto:tieying.zhang@alibaba-inc.com) 或者微信 18347624 欢迎大家转发推荐。

### 职位详情

机器学习 for 数据库：阿里巴巴大规模线上数据库存储了海量业务数据并和实际需求想结合，以机器学习与数据挖掘为基础面向 workload 预测、Cache 优化、异常检测、智能运维、调度优化、冷热分离、负载均衡等核心应用场景，使用人工智能的方法优化重构大规模数据生态系统。方向（但不限于）如下：

1. Workload 预测。根据历史 workload 预测未来负载情况，如预测双十一访问峰值、资源使用情况等以优化数据库系统。
2. 智能调参。基础软件系统（如数据库系统）通常有几百个参数作为系统输入，人工调参十分困难。我们的目标是使用机器学习的方法进行自动参数调优，转变以 DBA 为中心的数据库运维系统到自动化自学习的数据库运维系统。
3. 异常检测（预测）。自动预测诊断数据库系统的异常问题，并采取应对措施；针对采取的措施加以智能学习，迭代反馈更的优化手段。
4. 冷热分离。根据数据的访问频度（温度）对数据库进行冷热划分（预测），根据预测结果分离不同温度的数据到相应的存储层级。
5. 智能调度与负载均衡。针对该规模线上服务器集群，调度不同虚拟机到物理机上，增加资源利用率；同时设计高效的集群管理算法优化负载均衡，提高分布式系统效率。
6. Learned Index/LRU 等数据结构。使用人工智能模型替换传统索引（B+Tree、LSM 等）、LRU 等结构，减少存储开销，提升检索效率，并能够适应更新/写入操作频繁的应用场景。
7. SQL 优化器建模。传统 SQL 优化器方法适用场景较为简单，使用 histogram 优化器效果也并不理想。我们的目标是使用机器学习的方法针对不同硬件平台进行建模，更加准确的预估查询计划的代价，同时减少 CPU 开销。
- 8 数据库系统。典型去 IOE 场景，大规模数据库与存储系统，支持阿里巴巴双十一核心交易，构建自主可控的分布式数据库系统。

### 职位要求

- 精通人工智能和机器学习的理论和系统
- 熟悉 C++ 或 python，有较强的分析，定位和快速解决问题的能力
- 具备强大的自我驱动力，勇于接受挑战和拥抱变化，保持持续学习技术的热情，乐于团队合作和技术分享