

徐恒

Web前端开发工程师

男 / 1992.10
中国科学技术大学 · 软件学院
硕士 / 2020年毕业

我的技术博客
Github - xuoutput
847438919@qq.com
15167514153

经历

闪马智能

2019.04 ~ 至今

在闪马智能主要参与标注系统的开发

▪ 由于系统规模越来越大，不同功能模块间耦合度越来越高，开发难度不断增加，参与考虑项目重构的讨论。确定按照开闭原则和约定大于配置原则，在业务上划分为基础业务：用于管理用户，数据集和运营系统；标注业务：用于管理标注类型，标注任务，沙子库和考试。将代码按所属业务模块重新组织，模块间代码尽量解耦，但模块数据层暂时不变。按照以上模块划分对Controller，Service，Model和Router等文件制定命名规范，例如controller/[业务类型]/[具体业务].js。同时对API进行调整，按HTTP Status Code来作为请求成功与否的标志，一个请求的结果分为三大类（成功、客户端错误、服务器错误），如果是客户端错误，应该明确告诉用户该如何操作。

七牛云

2018.09 ~ 2019.03

在上海七牛云参与深度学习平台AVA和标注系统labelx的开发，包括：

- 深度学习平台AVA是七牛人工智能实验室打造的一款为深度学习提供定制服务的平台，包括数据集管理，镜像管理，工作台管理，训练任务管理，工作流等，支持MXNet，TensorFlow，Caffe，PyTorch等深度学习框架，主要目的是优化资源管理。把算法人员从繁琐的资源管理和环境管理中解放出来，让他们能聚焦于算法和模型本身的设计和实现。
- AVA是构建在七牛云存储服务（KODO）之上的，实现所有数据管理都在云端；有效管理GPU资源，通过Kubernetes合理划分和调度，将调试代码与模型训练分开进行，减少GPU闲置时间；最关键的是实现了存储和计算分离，使得各自可以独立平行扩展。
- 我主要负责AVA前端界面的设计与展示，在后端完善单独更新工作台/训练的资源API后，利用React + antd + MobX对创建工作台/训练的流程重新进行规划，保证在缺少资源的情况下也能暂时创建工作台/训练，等有资源空余的情况下可以按需进行获取，方便算法人员灵活利用同一个工作台/训练进行开发，不必每次重新创建工作台/训练，使用go来转发前后端交互信息，充当中间层，提示各种错误信息。
- 新增工作流功能，目的是方便算法人员对数据进行自动化处理，选择基于kubernetes的argo来作为实现工作流的工具，通过选择镜像并设置输入输出参数和执行入口来创建一个任务定义，然后排列多个任务定义创建一个工作流模板，最后通过选择工作流模板并传入具体参数和资源后运行。但功能还不成熟，仍需要继续优化。
- LabelX是服务于AVA平台的标注系统。拥有图片，视频标注功能，用于提供数据便于AVA平台中的训练进行学习。可单独使用作为标注平台。

技能

Web

- 熟悉JS基本语法与知识，最新JS语法与API，基本用TypeScript开发
- 熟悉网页重构所需的CSS、LESS
- 了解前端流行框架React、Vue、Angular

React

- 用React参与开发大型复杂应用

Webpack

- 了解Webpack的基本配置

Golang

- 了解golang基本语法

网络协议

- 熟悉HTTP协议

其它

- 了解Git、Docker

- LabelX 使用MySQL来存储数据标注信息, 使用 egg.js 来实现后台项目逻辑, 前端使用 React + MobX + antd 来设计与实现界面和业务逻辑。
- 在原LabelX标注系统的基础上参与设计与开发了运营系统, 方便对人员, 对数据进行集中管理。增加标注人员管理: 划分组来解决单一管理员管理所有人员情况以及增加查看人员数据标注情况; 任务类型管理: 从原来新建标注数据集就需要重新选择标注任务类型, 尤其是新建标签类别的情况, 到把任务类型单独分离开, 用时只需要下拉选择, 简化创建标注数据集操作; 标注任务管理: 即新版的创建标注数据集, 又在此基础上增加可以拆分到小数据集, 以及增加沙子的功能, 有利于管理员管理同类型标注任务; 沙子库管理: 主要是用来创建沙子供创建标注任务以及考试使用; 考试管理: 从沙子库中抽取数据用来考核标注人员, 培训用; 业务量统计功能: 按单位来查看所有任务的完成情况, 人员的标注任务完成情况, 便于结算。
- 遇到的最大问题是在修改权限上, 需要替换原来系统中鉴权。由于后端在路由上, 数据操作上都做了不同方式权限控制, 例如用http authorization header来验证用户信息和从session中获得用户组身份来进行对应操作。以及前端对每个页面也做了验证用户身份, 同时需要调整。优化从旧鉴权到新鉴权切换时, 若本地存在localStorage不满足新鉴权信息时就要删除并重新登录。
- 运营系统上线后, 在不影响标注人员标注的情况下, 提高管理员管理人员效率, 掌握具体人员标注数据数量, 数据合格率, 历史完成量, 考试情况。