

曲 硕

北京市海淀区学院路 37 号

☎ 18810960119 | ✉ m18810960119@163.com | 📷 qslublu

Education

北京航空航天大学计算机学院

计算机科学与技术

GPA: 3.45

北京, 中国

9. 2017 -

Experience

商汤

见习深度学习工程师

深圳, 中国

3. 2020 - 9. 2020

- 参与设计并实现深度学习 pipeline 框架, 调用各个功能组件实现数据处理, 训练, 评测, 模型转换等 stage 的 pipeline。
 - 实现了各个框架分布式训练接口的统一, 包括 pytorch, 单机训练, 以及公司内部框架。
 - 设计并实现了各个 stage 输入参数的规范, 支持单个 string 类型, path 类型, list 类型。
 - 完成了框架的 sphinx 文档。
 - 完成了框架的单元测试代码。
- 完成组内 docker 的基本构建。
 - 完成 docker 镜像的构建, 支持深度学习的基本环境, 包括 gcc, pytorch, cuda, 和其他依赖。
 - 在服务器上配置了 nvidia-docker, 支持 container 内使用 gpu。
 - 为分类, 检测等各个组件打包了 docker image, 支持从 osg 读取训练数据。
- 完成组内 gitlab-CI 的基本构建。
 - 在服务器上配置 gitlab-runner。
 - 完成 CI 的配置文件, 支持在 CI 中进行, 风格检查, 调用组件打包好的 docker 单元测试, 文档部署, docker 镜像打包。
- 阅读 mimic 相关 paper, 并在组件中进行复现和支持。
- 工具的对接, 使用 mimic 相关组件为 tracking 组件提高准确率。

北航 ACT 实验室-糖尿病药物推荐

本科实习生

北京, 中国

12. 2018 - 5. 2019

- 阅读相关 paper, 了解相关方法进行整理。
- 实现医院数据 parse 的脚本, 对于部分数据采用处理表格等库进行处理, 对于难以处理的数据采用 ocr 进行辅助处理, 将数据处理为血糖-时间点-药物。
- 对于数据采用 xgb 等机器学习模型进行训练, 实现给定时间点, 给定药物的血糖的预测。

datafountain 乘用车销量预测比赛

组长, 初赛成绩 35/2500, 进入复赛

10. 2019 - 12. 2019

- 对分省, 分车型汽车销量的数据进行特征工程。
 - 采用最小二乘法计算各个特征与预测值的相关性, 去掉不相关或相关性小的特征。
 - 采用滑窗构建多个时序特征。
- 采用多种模型进行训练, 包括 xgb, lgb, lstm 等。
- 对多种模型进行融合。

- 在 github 上爬去 java 代码。
- 对爬去的代码进行处理为方法-注释对, 同时把单词进行词性归一, 降低词典大小。
- 使用 pytorch 实现 seq2seq 模型。
- 对代码和注释采用 embedding 编码进行训练, 实现代码到注释的翻译。
- 将模型封装为 IDEA 的插件, 能够实现在 IDEA 内对选中的方法生成注释。

Honors & Awards

- 2017 国家励志奖学金,
- 2018 二等奖, 北航学习优秀奖学金
- 2018 二等奖, 北京市数学竞赛
- 2018 二等奖, 北航社会工作奖学金
- 2019 二等奖, 北航学科竞赛奖学金
- 2018 三等奖, 北航冯如杯科技竞赛
- 2018 北航优秀学生干部,

Skills

DevOps Docker, CI

Back-end Django, REST API

Front-end Vue, HTML5, Javascript

Programming Python, C, C++, Java, Typescript, LaTeX