

严晟嘉

✉ i@yanshengjia.com · ☎ (+86) 159-6188-7272 · 🌐 yanshengjia · 🏠 yanshengjia.com

教育经历

东南大学 南京, 中国
计算机科学与工程学院, 计算机科学学士 2013 – 2017

- 导师: 漆桂林教授, 方向: 自然语言处理
- 平均绩点: 3.56/4.0

工作经历

一起科技 北京, 中国
自然语言处理算法工程师 2017.06 – 现今

- 基于规则系统、特征工程、统计模型和深度学习模型, 开发英语作文自动批改系统
- 开发了拥有多种质量控制机制的众包标注平台, 包括登录系统和计费系统

论文

- Tianxing Wu, **Shengjia Yan**, Zhixin Piao, Liang Xu, Ruiming Wang, Guilin Qi, “Entity Linking in Web Tables with Multiple Linked Knowledge Bases”, Joint International Semantic Technology Conference (JIST), 2016. [PDF]

科研经历

知识科学与工程实验室 (由漆桂林教授领导) 南京, 中国
研究助理 2014.10 – 2017.06

- 学习自然语言处理领域的知识
- 参与了实体链接相关的研究项目

Entity Linking in Web Tables 南京, 中国
组长 2015.10 – 2016.10

- 提出了一种新的通用方法, 基于多知识库执行 Web 表格上的实体链接任务
- 基于实验成果的论文被第六届国际联合语义技术会议 (JIST 2016) 接收
- 设计了一个表格实体链接系统并且专利在中国发表

项目经历

Essay Rater 北京, 中国
主程序员 2017.10 – 现今

- 使用 LSTM 来提取文本特征从而预测作文整体得分, 基于统计模型预测作文各维度得分
- 实现了基于 Ngrams 语言模型的拼写修正模块
- 开发了基于规则和 Seq2Seq 的语法错误检测模块

Annotator [Link] 北京, 中国
主程序员 2018.02 – 2018.03

- 开发了带有登录系统、计费系统和质量控制机制的众包数据标注平台
- 前端使用 JS 和 Bootstrap, 后端使用异步 Web 框架 Tornado, 数据库使用 MongoDB

MiniC Compiler [Link] 南京, 中国
主程序员 2016.05 – 2016.06

- 用 Python 实现了“正则表达式 -> NFA”转换器
- 用 Python 实现了 LR(1) 分析器和相应的语义动作

AI Course Projects [Link]

南京, 中国

独立开发

2016.03 – 2016.06

- 使用了人工神经网络进行人脸识别
- 用 Matlab 实现了遗传算法来求解函数最值问题
- 用 C++ 实现了 A* 算法来解决 24 数码问题
- 用 C++ 实现 QS4 算法来解决百万皇后问题

技能与兴趣

兴趣: 自然语言处理, 深度学习, 机器学习

英语: 托福 100, GRE 317+4.0

编程: C++, Python, Shell, LaTeX, Markdown

Web 开发: 有 JS, HTML, CSS, Bootstrap 前端经验; 熟悉 Tornado 后端框架; 熟悉 MongoDB 操作

软件开发: Git 重度用户, 了解基本的设计模式, 能用 Qt 开发客户端, 能快速构建 Demo

机器学习框架: Tensorflow, Keras, Sklearn, Gensim

荣誉与奖项

第四名, 东南大学第十二届大学生程序设计竞赛

2016

省级三等奖, 江苏省程序设计大赛 (JSCPC)

2016

志愿者之星, 南京青奥会

2014

课外活动

加州州立理工大学国际学生领导力项目

加州, 美国

成员

2016.02

南京青年奥林匹克运动会志愿者

南京, 中国

国家奥委会助理

2014.07 – 2014.08