

2018 年度个人工作总结与展望

——2018 annual personal work summary and outlook

胡伟龙

huweilong@whu.edu.cn

<https://github.com/knightBoy>

小米-武汉大学人工智能联合实验室
武汉大学语言与信息研究中心
武汉大学计算机学院

2019 年 1 月 11 日



提纲

研究总结

项目总结

研究展望

项目展望



2018 年研究工作总结

● 论文投稿

- Distant Supervision for Relation Extraction Combining Sentence-Level Attention with Pattern Filtering, (EMNLP 2018, 2018.6)
- Distant Supervision for Relation Extraction Combining Sentence-Level Attention with Pattern Filtering, (AAAI 2019, 2018.9)
- Dilated Convolutional Networks Incorporating Soft Entity Type Constraints for Distant Supervised Relation Extraction, (IJCNN 2019, 2018.12)
- Distant Supervision for Relation Extraction Combining Sentence-Level Attention with Pattern Filtering, (WWWJ 2018, 2018.12)

● 参加 AI CHALLENGER 2018 英中文本机器翻译比赛 (A 榜 45/99, 2018.10)

● 论文、基金评审工作, 若干项目材料的撰写

● 命名实体识别、实体关系抽取、基于知识库的问答学习, 《深度学习》整理 <https://github.com/knightBoy/Knowledge-Notebook>

2018 年项目工作总结

- 小米精品知识图谱项目 (2018.3-2018.12)
 - 手动建立艺术领域本体库, <http://git.pmnlpplab.top/Knowledge-Graph/Art-DBpedia>
 - 基于 Scrapy 对百度百科、美术网等 7 个网站进行半结构化知识抽取
 - 初步建立命名实体识别、关系抽取标注平台, <http://111.6.102.18:7575/tagger>
- 北京小米实习 (2018.8)
 - 学习了小米智能问答组在知识图谱上的工作
 - 情感分析 (文本分类) 项目的需求对接
 - 其他 3 个项目的汇报、标注数据获取与沟通
- 实验室环境建设
 - 深度学习 GPU 环境管理, <http://wiki.pmnlpplab.top:8888/dokuwiki/>
 - 实验室主页、演示平台等服务迁移, <http://2b3292h001.imwork.net/>

2019 年研究工作展望

- Distant Supervised Relation Extraction
 - 现状：
 - Embedding+Encoder+Attention+Classifier 的方式行不通
 - New York Times 数据集指标变得不可超越
 - 计划：
 - 使用新数据集
 - 利用基于图方式建模实体与关系
 - Unsupervised Relation Extraction
- Question Answering
 - Question Answering base on Knowledge Base
 - Dialogue Generation



2019 年项目工作展望

- 艺术领域精品知识图谱 (Art-DBpedia)(-2019.7)
 - 通用图谱建设流程：本体建立、数据标注、数据爬取、模型训练、数据存储、平台展示
 - 完成第一年目标：数据覆盖率 70%(三元组数达到百万)，数据准确率 75%
 - 后两年计划：提高自动化水平(主动学习、远程监督、现有知识库、自动爬取)
- 实验室环境建设
 - 整合 Wiki, Gitlab, HomePage, 技术 (Intention recognition, Machine reading comprehension, etc.) 对外展示宣传平台
 - 形成规范并做技术交接
- Get a good job!

结语



祝大家猪年快乐，学习进步，心想事成！