



陈璐

电话：153-5992-9898 | 邮箱：none954@163.com | 地址：清华大学FIT楼
生日：1995年8月2日 | 性别：女
当前状态：2020年毕业



教育经历

清华大学 - 计算机科学与技术 硕士 交叉信息研究院 2017年9月 - 2020年7月

- GPA : 3.32 / 4.0
- 清华之友-百度社会工作奖学金 (2018年)
- 相关课程：统计学习理论, 概率图模型、计算语言学、知识工程、最优化方法、应用随机过程

北京邮电大学 - 计算机科学与技术 本科 计算机学院 2013年9月 - 2017年7月

- 成绩：93 / 100 (排名 2 / 302)
- 国家奖学金 (2015年、2016年)、校一等奖学金 (2014年)、校三好学生 (2014年、2015年、2016年)
- 相关课程：操作系统、计算机网络、数据库原理、数据结构、算法设计、概率论、离散数学

项目经历

上市公司公告关系抽取 2017年4月 - 2018年4月

结合labeling function (LF) 的远程监督 (DS) 方法, 抽取上市公司公告中的董事、股东、监事、法人四种关系。为了快速构建大规模数据集, 使用DS为大量数据打上标签, 但是这些标签带有大量噪声, 利用LF降低噪声, 从而使用降噪过的数据训练关系抽取分类器。为了快速获取test数据, 搭建了快捷标注平台。在test集上多关系平均AUC达83%。

事件抽取调研 2018年9月 - 2018年11月

在上市公司公告中抽取股权转让, 股权质押, 股份增减持, 股份增发事件。

金融结构化信息平台 2017年1月 - 2017年5月

利用爬虫框架WebMagic爬取上市公司信息, 使用Java前端框架SpringMVC和ORM框架Mybatis搭建平台。通过图形数据库Neo4j和js库D3.js在页面展示公司的信息关系图。自动生成结构化公司信息报告, 其中包括公司工商信息, 资本信息等, 并且对公司各个主页自动截图, 使用TextRank算法在主页提取关键词和摘要, 生成公司的梗概信息。最后实现了简单的公司信息自动问答。

研究经历

基于规则诊断的远程监督关系抽取 2018年5月 - 2018年8月

DIAG-NRE: A Deep Pattern Diagnosis Framework for Distant Supervision Neural Relation Extraction 在投

为了通过小型数据集学习到较好的分类模型, 使用了远程监督 (DS) 的半监督学习方法。使用强化学习自动生成规则, 对生成的规则筛选出具有代表性的pattern, 并且对其人工校准, 将规则转化为LF对样本打上弱标签, 将远程监督 (DS) 生成的Label和LF生成的Label做weak label fusion生成新的soft label, 以此作为训练集denoised label训练更好关系抽取模型。在NYT数据上的10个关系分别得到了不同程度的提升, 不同关系上F1相比DS最好的提升了40个点。

图像的对攻击 Adversarial Attack in Sequential Recognition Model 2019年3月 - 至今

对攻击是对原始图片进行微小扰动生成对抗样本, 使被攻击模型错误识别对抗样本, 且识别结果是攻击者指定的。目前在中文数据集上训练 Densenet+CTC 模型作为端到端的 OCR 识别模型, 通过Momentum Iterative Fast Gradient Sign Method 生成targeted对抗样本进行识别, 仅仅进行 20 轮迭代即可达到 94%的攻击成功率。

实习经历

华控清交信息科技有限公司 - 软件开发 2018年12月 - 至今

公司的主营业务是多方安全计算, 我的主要负责通信模块(grpc), 测试模块(googletest|pytest)和增加调试模块(jupyter)。

专业技能

- 编程语言：熟练：Python, 熟悉：Shell, Java, C++, SQL, JavaScript, 了解：C, C#
- 模型：熟悉CNN、RNN、AttBiLSTM、word2vec、NER、CRF等自然语言处理模型
- 工具包：Numpy、Pandas、Scikit-Learn、PyTorch、XGBoost
- 课程项目：点击欺诈检测(python)、自动作诗(python)、中央空调系统(Qt/C++)、购物平台(Java)、HEXA机器人(Golang)
- 外语：英语CET-6 (536)

社团经历及兴趣

清华大学 交叉信息研究院 学生会 文体部部长 2017年9月 - 2018年7月

清华大学SSLP志愿支教 领队 2017年7月 - 2017年7月

兴趣爱好: 跑步, 瑜伽, 钢琴