

刘俊帅

☎ 电话: 13552169300 ✉ 邮箱: 1358780585@qq.com

📍 地址: 北京市海淀区西北旺镇夏霖园

🏠 教育背景

2016-2019, 硕士, 北京理工大学, 计算机科学与技术

2012-2016, 学士, 北京语言大学, 计算机科学与技术+金融 (第二专业) 英语专业四级

📅 工作经历

【1】2019.7 - 至今

联想研究院-AI Lab

(语音语言研究员)

主要工作内容

- 1, 负责 NLU 模块: 搭建标注数据集, 搭建模型 (BiLSTM、BERT、CNN) 训练以及优化; 在不同领域的意图分类 F1 值达到 0.9, 实体抽取 F1 值达到 0.9; 支持乐语音智能手机助手业务、show mode 平板电脑业务。
- 2, 负责语音识别后处理模块: 搭建标注数据集, 搭建模型 (序列标注) 训练以及优化; 数字转换、标点符号、语法纠错等任务的 F1 值达到 0.8, 多个任务集成于一个序列标注模型中, 将系统时延相对降低 30%; 并集成于线上语音识别引擎中, 支持乐语音智能手机助手业务、联想中国区电话销售业务。
- 3, 负责语言模型模块: 搭建数据集, 搭建模型 (RNN、LSTM、BERT) 训练以及优化; 将语音识别准确率相对提升 10%; 并集成于线上语音识别引擎中, 支持乐语音智能手机助手业务、联想中国区电话销售业务、show mode 平板电脑业务。

【2】2018.3 - 2019.5

腾讯科技有限公司

(企业实习)

主要工作内容

- 1, 新闻摘要, 采用等间隔采样的多层编码器来学习文本的语义结构信息; 将线上新闻摘要系统准确率相对提升 10%。
- 2, 客服对话摘要, 提取客服对话中的核心问答对; 采用 GAN 来生成对抗样本, 解决数据不足的问题; 构建问答对知识库, 服务于在线客服系统。
- 3, 多轮对话中 SQL 语句生成, 提出增强记忆的网络单元来记录对话信息; 提出单向信息流的注意力机制来对上下文进行表示; 服务于在线客服系统。
- 4, 多义词的向量化表示, 基于概率分布以及上下文信息来对多义词进行建模;
- 5, 新闻推荐, 采用自注意力机制根据新闻中的实体信息来进行细粒度推荐; 集成于线上新闻推荐系统, 准确率提升 10%。
- 6, 新闻推荐中的实体表示学习, 针对推荐任务的知识图谱表示方法; 集成于线上新闻推荐系统中, 准确率提升 20%。

【3】2016.1 - 2016.9

中金创新有限公司

(企业实习)

主要工作内容

- 1, 基于机器学习方法的期货的量化交易程序开发, 主要负责算法设计以及代码实现。

📅 项目经历

【1】2017.7 - 2018.2

腾讯科技有限公司

(校企合作)

主要工作内容

- 1, 客服对话摘要, 本项目面向客服领域的客服对话摘要, 通过对客服对话进行摘要提取其中核心问答对。构建客服问答知识库。针对客服领域标注数据不足的问题, 将自编码器任务与新闻摘要任务联合优化, 并将两个任务共享语义特征空间, 在该特征空间中将新闻领域的摘要知识迁移到客服领域。

【2】2017.3 - 2017.7

华为北京研究所

(校企合作)

主要工作内容

- 1, 构建用户在通话、位置、APP 使用等不同领域的社交网络, 形成融合多领域的用户社交关系网络。
- 2, 利用 k-means 算法划分社交圈; 利用随机游走算法计算用户影响力。

📅 论文专利

论文 1: A Multi-Level Encoder for Text Summarization. In Proceedings of the 2017 IEEE Symposium Series on Computational Intelligence. (IEEE 人工智能领域旗舰会议, EI)

论文 2: Extracting Company-Specific Keyphrases from News Media. In Proceedings of the 2017 International Conference on Computational Intelligence and Security. (EI)

论文 3: Fine-Grained Deep Knowledge-Aware Network for News Recommendation with Self-Attention. In Proceedings of the 2018 IEEE/WIC/ACM International Conference on Web Intelligence. (EI)

论文 4: Zero-Shot Extractive Text Summarization. IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING. (under review)

论文 5: Entity Representation Regularization for Recommender Systems with Knowledge Graphs. Neurocomputing. (under review)

论文 6: Fine-Grained Deep Multimodal Fusion with Mutual Attention for Fake News Detection on Microblogs. IEEE Internet of Things Journal. (under review)

专利 1: 一种基于文学文本的情感赋能有声资源合成方法。

专利 2: 基于脚步声的电梯控制系统。

专利 3: 一种防窃听与干扰耳机。

专利 4: 一种联合标点预测与纠错方法。