

张美伟 (男)

电话: 18581554051

Email: zhangmw_play@163.com

GitHub: <https://github.com/chriszhangmw>

领英地址: <https://www.linkedin.com/in/chris-zhang-8a644913a/>

居住地址: 龙泉驿区, 成都, 四川

教育背景

阿斯隆理工	硕士	软件工程	2019-2020
电子科技大学	学士	微电子学	2011-2015

基本技能

- **开发语言:** 主要熟悉 Python, 熟悉 Linux 开发环境, 了解 Java 等常用编程语言。
- **开发工具:** 熟悉使用常见的工具 Git, Agile, Scrum, Pycharm, RESTful, Postman, 熟悉 Linux 和 docker 常用命令; 了解 Sanic, flask 等常用 web 框架
- **NLP 算法技能:** 熟悉传统机器学习算法: LR, SVM, Decision Tree, K-means, HMM 原理和使用; 熟悉常见深度学习网络机构和 NLP 相关结构: RNN, LSTM, Bi-LSTM, Seq2Seq, TextCNN, Graph CN, Transformer PGN, BERT, ELECTRA, Skip-Gram model. 熟悉 tensorflow, keras, sklearn, rasa 等
- **技能领域:** 文本分类, 文本摘要与生成, 命名体识别, 关系抽取, 聊天机器人, 知识图谱等

工作经历

华为爱尔兰研究所	知识图谱研究员	06/2020-至今
----------	---------	------------

- 负责AIops的chatbot开发工作, 将知识图谱, chatbot, stackstorm (工作流) 结合用到智能运维模块
- 负责根据知识图谱做意图识别的设计和实体槽位设计, 并结合rasa框架开发NLU模块, 结合图谱动态设计form action的槽位, 增强了chatbot与图谱的对接功能, 增强了rasa与后端图谱对接的可扩展性, 并负责了rasa action与stackstorm工作流对接模块开发。

Avaintec	NLP工程师	03/2018-08/2019
----------	--------	-----------------

- 负责开发基于RASA框架的检索和任务型多轮对话应用开发, 根据不同业务场景开发form action, 结合NER和意图识别开发NLU模块。
- 开发datachief平台工具, 将不同算法功能模块接口集成到该平台, 模块包括常见的分词, 分类, 聚类, NER等

华通科技公司	NLP工程师	04/2016-02/2018
--------	--------	-----------------

- 负责结构化医院内的数据清洗, 分析, 主要根据现有数据构建专科辅助诊断系统以及NER。

IBM	系统工程师	06/2015-03/2016
-----	-------	-----------------

- 参与政府, 社保等行业的IT咨询与技术支持, 负责常见的IT健康检查, 并在售前阶段协助销售完成方案设计。
- 负责POC环境搭建以及售中环境搭建, 并根据项目交付阶段定期组织项目复盘, 持续优化Z系列主机产品方案。

项目经历

➤ 高频词对于文本生成的影响研究

10/2019 - 01/2020

项目描述: 生成式文本摘要或者文本推理应用广泛, 但在解码阶段均会收到超出词表, 无法准确解码和重复解码的问题。结合tf-idf原理, 提出高频词对于文本生成任务的研究与评估。

项目贡献: 提出了词采样的方案, 平滑高频词在解码阶段的权重, 即调整attention矩阵的值; 同时与embedding

阶段平滑高频词的词向量对比，发现attention阶段的高频词平滑会导致过拟合。作为第一作者，将该论文发表于期刊Journal of Machine Learning and Computing (IJMLC, ISSN: 2010-3700)。

➤ 数据增强对于多轮对话任务的提高

12/2019 – 02/2020

项目描述: 结合知识图谱与文本数据增强提高了多轮对话主题切换的平滑性。

项目贡献: 结合命名体识别，短文本匹配，损失正则构建鲁棒性强的知识选择模块，提高了知识图谱在对话系统的应用。作为第一作者，将该论文发表于期刊：International Journal of Mechanical Engineering and Robotics Research。

➤ 移动医疗问答系统

06/2018 – 08/2019

项目描述: 根据知识库/谱结合rasa框架开发检索式和任务型聊天机器人，提高患者在对于常见问题咨询的便捷性，并根据不同的form定义药物咨询，专家咨询，诊断咨询等应用

项目贡献: 负责NLU模块功能开发，结合chatito构建意图样本，利用Bert等框架训练实体提取模块；根据业务场景与爬虫团队一期构建RDF数据库，结合互信息，tfidf等算法做召回排序。

➤ Datachief平台开发

12/2018 – 08/2019

项目描述: 该平台供提供医学文本的分词，分类，聚类，摘要，NER，相似性工具。

项目贡献: 利用HMM，维特比算法构建医疗领域的分词工具，结合bert as service，SIF算法等完成文本向量化并基于此开发了文本聚类，相似性等工具；利用BERTSUM等算法完成了摘要应用的升级。利用Sanic部署restful

➤ 急腹症辅助诊断系统

06/2016 – 12/2017

项目描述: 根据急腹症科室，系统提供多个器官与多个症状的级联系统，加快科室快速分诊能力。并结合datachief平台，与感染科，传染科等兄弟科室构建数据湖平台。

项目贡献: 系统构建初期，负责算法设计与实现，提出了多标签文本分类在有有限数据集下的模型级联方案，同时负责了以循环神经网络，图神经网络，预训练神经网络等模型的设计与实践。最终在14个二分类的级联模型中将8个模型的找回了（敏感度）提升到商用水准。

技术认证与论文发表

-
- Journal of Machine Learning and Computing (IJMLC, ISSN: 2010-3700)
 - Journal of Mechanical Engineering and Robotics Research
 - PMP – 项目管理师认证
 - H3CNE - H3C 网络工程师认证
 - IBM System z Technical Support V4 – 培训
 - IBM Global Sales School (Best Team Award) – 培训

其他信息

-
- 爱好广泛，阅读，篮球，吉他。
 - 性格比较友善，易沟通，协作能力强。