孙焱

电话：183-3189-1019 | 邮箱：sunyanhust@163.com
求职意向：自然语言处理、搜索推荐

教育经历

华中科技大学 - 计算机技术 硕士 2018年9月 - 2020年6月

- 华中科技大学张量计算实验室，导师是莫益军教授，主要研究方向为任务型对话系统，有模型压缩相关经验
- 获2018年秋季硕士新生奖学金，2019年优秀班级团支书

中北大学 - 通信工程 本科 2014年9月 - 2018年6月

- GPA： 4.2/ 5.0（专业前5%）| 荣誉/奖项：2016年国家励志奖学金，2017年航天公益奖学金

专业技能

- 编程技能：熟练使用Python，C/C++, Git，Markdown，了解Java, Matlab等编程语言和工具
- 机器学习和深度学习：了解常见的机器学习模型如逻辑回归、朴素贝叶斯、SVM，较为熟练的掌握Tensorflow、Keras、Sklearn等深度学习和机器学习框架
- 传统统计自然语言处理：词袋模型、TFIDF、N-Gram，以及潜在语义分析等主题模型方法。
- 深度学习自然语言处理：熟悉RNN以及LSTM、GRU，CNN以及Transformer的特征提取器，掌握Word2Vec、SeqSseq、Attention等网络结构。熟悉TextCNN、RCNN、HAN等文本分类模型，特别是样本不均衡分类、多分类和多标签的分类任务。熟悉BiDAF、QANet阅读理解模型，以及ELMo、GPT、BERT等语境化词表征模型。

实习经历

阿里巴巴 - 算法实习生 阿里云智能事业群 2019年5月 - 2019年8月

主要工作内容：在AliOS Things物联网实时操作系统上实现了语音和语言的智能识别，完成了在移动端部署语音和语言深度学习模型的任务，包括基于CNN的声纹识别和孤立词语识别、基于INT8量化和Tensor Network表示实现模型参数压缩，以及基于KD树的向量相似度高效计算。

- 使用CNN对语音的特征图谱进行训练，使用Triplet Loss拉大类间距离，得到语音信号的向量表示，在测试集上的正确率达到91%。之后对模型的参数进行INT8量化表示和模型压缩，模型大小减少到原来的1/3，推断速度提高20%。
- 将训练好的模型转换为嵌入式C可以调用的模型，校验模型输出结果。在MCU上录音并完成音频频谱特征提取，调用C模型进行预测，根据模型输出向量和模板向量的相似度进行声纹识别和孤立词语识别。

比赛项目经历

论文：联合Multitask Learning阅读理解的任务型对话系统 - 第一作者 2019年3月 - 2019年6月

为提高任务型对话系统的流畅性和准确性，提出一种Chatbot与基于Multitask Learning阅读理解结合的对话系统框架。首先使用Chatbot的NLU模块处理用户问题，根据命名实体识别和意图分类的结果启动不同的Action进行应答。在闲聊型问题中，采用基于Transformer的Seq2Seq模型生成答案；在任务型问题中，采用问答和分类同时进行的Multitask Learning阅读理解模型生成答案。最终模型在测试集上的综合F1值达到0.91，远远超过原来Pipline模型的效果。

论文：基于Tensor Train分解的Transformer模型压缩 - 第一作者 2018年11月 2019年3月

通过对Transformer模型的Embedding层和全连接层进行张量化表示，使用Tensor Train(张量链)分解对模型参数进行了压缩。在不降低模型精度的情况下，张量分解的方法能有效压缩模型参数，6层8头512隐藏单元的Transformer压缩后的参数仅为原来的0.168倍，显著降低了模型训练时的GPU占用。压缩后的模型在文本分类，命名实体识别、阅读理解以及对话系统任务上均取得了略低于甚至超过原始模型的效果。

文本分类比赛：Quora Insincere Questions Classification 2018年12月 - 2019年1月

Kaggle Quora英文二分类比赛，主要工作为通过词形还原、词语纠错等减少了词向量的OOV，分别对常用的文本分类模型如TextCNN，RCNN，Transformer Encoder等进行了尝试和对比，使用Focal loss作为损失函数和更换F1值作为评价指标来缓解正负样本不均衡，并使用Self Attention改进模型，效果提升了6个百分点。最后对多个模型进行了10折交叉验证和Bagging集成，最佳排名为53/1350。