

个人陈述

很久以来, 我一直被一种信念驱动——如何利用世界的信息网络,去改善我们的生活. 这个话题很宽泛, 确实,人类已经挖掘各个维度的信息, 并将他们应用在不同的 AI 领域中, 第一步就是让机器去认识这个世界. 我感兴趣的领域是 NLP, 知识获取, 知识表示等, 过去的经验给了我充分的背景知识和决心去寻求更远的研究, 基于我强大的自我推动力.

我是在 2016 年进入研究生阶段的学习,在这个时间段,我培养了我对人工智能的兴趣,并将其作为事业的方向.在 2019 年加入一家 AI 公司, 作为一名语音工程师. 到目前为止,我主要参与 NLP 和语音合成的研究和工作的, 以下我会基于领域,讲述我做的相关的一些研究.

关于语义 NLP 的研究: 在 Mobvoi,基于跨语言(中文/台湾/粤语/英文/四川 etc.), 主导了几个模型的构建或者优化,包括,语义分段,文本归一,多音字消歧,韵律分析等(语音合成前端). 这些联合模型通过一系列的文本分析,将原始文本转换为包含准确发音、韵律、重读、情感等信息, 这些信息直接影响了语音合成的效果, 比如可懂性和自然度. 首先, 我们基于规则和简单模型建立了每个模块, 比如 CRF, FST, LSTM 等. 接着, 我发现语义分段/多音字/韵律都是基于文本的语义, 同时我们的系统能力有限. 因此, 我提出建立一个 unified 的模型架构, 旨在将复杂联合模型转化为端到端模型. 我们构建一个 pre-trained backbone 去获取这个语义信息,并将其提供给 unified 的架构.最后,使用多任务的模式获得最后的模型. 这个过程中, 我验证了每个模型的最大能力, 最后设计一个合理的多任务训练方式. 这个方案可以很好解决 OOV 和歧义以及模型复杂度的问题.这个部分我们涉及到了:CRF,FST, 语义分析,数据挖掘,序列模型,预训练模型,知识蒸馏,多任务训练,模型、数据鲁棒性等方面的研究.并基于以上撰写多个专利和相关论文;

关于 Emotional NLP 的研究: 在 Mobvoi, 基于文本研究了影响 TTS 自然度的因素,并着力于有声书等场景. 研究主要分了几个阶段: 首先, 基于预训练模型, 引入文本语义信息, 可以有效提高句子合成的自然度,比如获得感叹、疑问等语气的表现. 其次, 引入对话上下文并没有用, 分析原因——对话是多风格的, 引入不相关或者不明显的信息, 对整句话的情感不能起到很好的指导作用. 最后, 引入了 style 的概念, 通过研究大量文献以及专业播音人员的沟通, 一起制定了数据基准. 同时, 在少量数据以及半监督方式下, 得到了一个很好的风格提取模型——提取句子的 style,将它作为整个文本合成的风格指导, 并且不会被整个句子大量的信息影响. 最后, 基于 style 的研究, 提出了渐进式粒度的重读预测. 这部分基于数据分析和语言学知识, 两阶段建立由粗到细的重读模型, 在不同粒度信息监督下, 化繁为简, 减少模型的方差. 结果令人非常满意, TTS 获得更好的表现力. 以上的研究得到了用户和其他产业的关注.这个部分我们涉及到: 语言学知识, 情感分析, 对话分析, 对比学习,预训练模型,知识蒸馏,数据鲁棒性,并基于以上撰写多个专利和相关论文. (one paper + one patents).

关于 Knowledge Tracking 的研究, 在 TAL Lab, 我建立了一个 DKT 模型, 去于学生的历史答题信息进行建模,监控学生的知识状态或技能获取水平. 我们不仅

考虑了问题的序列依赖，还考虑了问题内在的联系。首先，我们基于语义 (embedding) 和统计信息 (共同连接，共同 words 等维度) 建立了问题的 graph。然后，我们使用了图嵌入方法去获取问题 graph 的表示。此外，我们建立问题的拉普拉斯矩阵，以便捕捉到问题之间的关联性，并将矩阵加入 loss 优化目标。最后我们的最优模型较 baseline 都有 4%-5% 的提升。该方法证明了 side information 的有效性，并提出基于 graph embedding 来表示这种序列记忆问题的思路，可以很方便的扩展到类似的问题，我们在 movielens 数据集上验证了这个结论。该方法也可以应用在答题情况预测以及规划自适应的学习路径上。以上研究是我同一位 DR 一起完成，并完成一篇论文。

其他：在我本科期间，学习了计算机科学以及数学的基础知识，这为我的研究打下了基础。在研究生期间，我自学并主动参与一些企业以及高校组织的学术活动中，参与过很多方向的研究和实验，包括计算机视觉、推荐系统、知识图谱，数据挖掘等，加强了我对机器学习、深度学习的认知。同样我遇到了很多优秀的人也增加了我的工程经验。在工作的这几年，我跟很多人一起共事，提高了我解决问题的能力。作为 mentor，我跟 5 个实习生一起合作过。期间，我建立了自我驱动力和管理能力。我有良好的代码风格和记录习惯，最重要的是我的初心未变。期待在感兴趣的领域有个傲人的突破。那也是我正在做的！

未来工作：我相信 USC 可以很好满足我的愿望，因为**。特别的是，我将会非常高兴，如果我们跟**一起工作。我研究过他最近的一些工作，包括在 EMNLP 和 NAACL 发布的工作。我对知识获取，数据鲁棒性表示，以及如何利用这些信息，有着非常大的兴趣。我认为人类处理世界信息的过程是非常神奇和令人着迷的。因为我已经在相关领域 (NLP) 有一些研究并发表过一些工作，所以，我非常有信心，在未来我可以跟我的同事和教授一起解决更多实际的问题，同时在一个领域贡献自己的力量。