

应聘职位

自然语言处理实习生

基本信息

姓名：李伟康 学校：北京大学 英语水平：CET 6
籍贯：河南洛阳 学历：研一在读（学硕） 出生年月：1993 年 10 月
性别：男 专业：计算语言学 爱好：篮球，乒乓球
联系电话：13146038034 邮箱：wave.jkd@pku.edu.cn 邮编：100871
通信地址：北京市海淀区颐和园路 5 号北京大学



学校经历

2012-2014 年，连续两年获得**专业一等奖学金**
2014-2015 年，获得**国家奖学金**
2012-2014 年，连续两年获得**国家励志奖学金**
2012-2013 年，获得**三菱日产奖学金**
2013-2014 年，获得院电子作品大赛**编程类二等奖**（获奖作品：基于 Heritrix 和 lucene 的手机领域的垂直搜索引擎）
2014-2015 年，参加学院 T.B2 **创新比赛**（作品名称：体型测试仪 APP），获得**三等奖**
2012-2014 年，连续两年获得校**数学建模比赛三等奖**
2015 年 9 月，本科时期成绩在所在专业位列第二，获得北京大学学术硕士研究生的保送资格
2016 年至今，北京大学信息科学技术学院研究生在读（计算语言学方向）

科研项目经历

2016 年 3-6 月，课题：问句理解的话题和焦点信息的识别。通过基于树切分的 MDL 方法将问题表示为包含话题信息和焦点信息的数据结构，然后利用语言模型框架对这些信息进行建模，从而解决实际问题中的问题检索问题。这个课题主要通过传统的机器学习方法进行，过程中实现了一些简单模型（VSM、LMIR、BM25、Lucene）等。
2016 年 10-12 月，课题：通过深度学习方法进行问答系统中的答案选择工作（Answer Selection）的研究。这个过程主要借助 tensorflow 深度学习框架对 CNN、RNN、CNN-RNN、RNN-CNN、RNN-attention 等模型在问答系统中的答案选择问题上进行了研究，并尝试了将问题和答案间的一些结构信息引入分类前的向量表示中。目前取得的结果相对可观。

工作实践经历

2014 学年 9 月 2015 年 6 月 地点：中央民族大学少数民族语言检测中心 内容：项目实习
参与实验室和清华合作的新华社项目，主要负责通过 Crawler4j 的数据抓取及 oracle 数据库的任务。熟悉了项目的开发流程，强化了自己的科研能力，对自然语言处理有了较深入的理解。
2015 学年 7 月至 2015 年 9 月 地点：浙江省杭州市阿里巴巴集团 内容：Java 研发工程师实习
参与公司“村淘”项目，主要通过公司内部框架 WebX3.0 从事服务端的开发工作，并借助公司内部框架 BUI 进行部分网页设计。此次经历使我熟悉了部门内的协同工作和部门间的交流沟通，了解了团队合作。
2016 学年 7 月至 2016 年 9 月 地点：江苏省苏州市金童技术公司 内容：自然语言处理工程师
主要从事财经数据方面的“实体关系”提取的研究工作，在实体识别的基础上，通过选取实体内、外、间的特征进行分析并构建 SVM 进行关系的初步提取（分类）。此次经历使我熟悉了如何将所学用于实际的市场需求之中，提升了知识运用能力。

专业技能

- 1、熟悉自然语言处理的智能问答领域；
- 2、了解自然语言处理的相关领域及机器学习算法；

- 3、了解 python 语言及 tensorflow 深度学习框架;
- 4、了解 linux 操作系统基本知识与基本命令。

自我评价

- 1、工作态度严谨，责任心强，不轻言放弃，善于沟通，理解能力较强，具有较强的团队协作精神;
- 2、学习能力强，喜欢钻研技术问题;
- 3、有较强的专业基础知识，能独立并开创性工作;

小论文

- 2013-2014 年，参加国家级**创新性科研项目**（GCCX），项目名称：基于视觉技术的四旋翼飞行器的目标定位系统的设计与实现，基于该项目以第一作者身份在 2015 应用科学与工程国际学术会议上发表论文“Design of Four-rotary-wing Aircraft Targeting System”，被 EI 检索
- 2014-2015 年，参加国家级**创新性科研项目**（GCCX），项目名称：面向藏文网络流通语料的语义分类研究，基于该项目以第二作者身份在 2015 计算机科学与通信工程国际会议上发表论文“Semantic Classification Method for Network Tibetan Corpus”，被 EI 检索
- 2014-2015 年，在老师的指导下进行藏文语义相似度的研究，以第二作者的身份在 2015 第六届 IEEE 软件工程与服务科学国际会议上发表论文 “Research on Text Similarity Computing Based on Word Vector Model of Neural Networks”
- 2016 年 4 月，在导师指导下参加华为在日本举办的 NTCIR-12 比赛，参赛系统位列所有系统的前三分之一，赛会的论文被邀请做 poster。