**南京大学软件学院研究生学位论文中期检查报告格式**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **南京大学软件工程硕士学位论文中期检查报告** | | | | | |
| 导师1姓名 | 伏晓 | 研究生姓名  （学号） | 安磊  MF1732001 | 方向 | 软件工程 |
| 导师2姓名 |  |
| 论文题目 | 基于深度学习的信息抽取技术的设计与实现 | | | | |
| 论文选题来源及研究的目的和意义（500字左右）：  随着人工智能技术的快速发展，越来越多的算法得到应用。自然语言处理作为人工智能领域中的一个重要分支一直是许多人关注的焦点，而知识图谱是自然语言处理中一个重要的研究方向。构建知识图谱的数据来源包括结构化文本和非结构化文本，作为构建知识图谱的重要数据来源之一的非结构化文本，如何从中提取高质量的信息成为一个研究热点。  面对上文所述的问题，一个能够提供高质量数据的信息抽取系统就显得十分重要。本文基于从非结构化文本中进行信息抽取的背景，以构建知识图谱时所需的三元组作为目标，实现从非结构化数据到知识图谱三元组的数据处理与转换。  本文基于针对非结构化文本的信息抽取系统的具体项目，针对金融上市公司领域进行命名实体识别、实体间关系抽取。由于基于深度学习的命名实体识别和关系抽取在不需要很多额外特征的基础上相对于传统机器学习就能得到更好的效果，所以本文从模型设计和实现的角度提出了基于深度学习框架Tensorflow实现模型、搭建信息抽取系统这一解决方案 | | | | | |
| 该方向的研究现状或技术进展综述（2000字左右） | | | | | |
| 论文的主要技术路线、研究思路和实现方法；相关项目应用前景：（重点说明变更部分）： | | | | | |
| 本人在相关项目中的扮演的角色和承担的工作（重点说明变更部分）：  项目中的两个算法模型都是由我一个人主导并进行开发的，后期的测试与调优也是我一个人完成。 | | | | | |
| 论文的主要工作（500字左右）： | | | | | |
| 实现论文三级大纲： | | | | | |
| 论文和相关项目的当前进度： | | | | | |
| 论文和相关项目进展过程中遇到的困难和问题，以及解决的措施： | | | | | |
| 主要参考文献： | | | | | |
| 导师意见： | | | | | |
| 学院备案意见：        年 月 日 | | | | | |