

工科专业毕业设计（论文）撰写指南

工科专业的毕业论文主要包括两类：应用开发型的毕业设计和理论研究类型的毕业论文。前者是应用所学知识开发系统或应用程序；后者是对本专业理论研究成果的展示。

1.应用开发型毕业设计：侧重于软件（系统）开发，要求能够运用所学专业知识和技能解决实际问题，锻炼实践动手能力。毕业设计突出原创性，即硬件系统的开发完全由自己完成，而不能抄袭。需要自主动手完成整个设计过程，最终的毕业论文应该把项目的系统分析、系统设计和开发的步骤等都详细地描述出来，提供可以检验的软件成果。

2.理论研究型毕业论文：研究思路清晰；方法适当，逻辑性强，对问题有较深入的分析；以及详细的新的算法或模型的设计开发过程，并有实验分析。论文注释和写作体例符合规范；参考文献具有适合性、准确性、专业性和权威性。

应用开发型毕业设计正文写作结构如下：

正文是毕业论文的主体和核心部分建议的正文结构如下：

第1章 绪论

1.1 项目背景与意义

1.2 设计目标

1.3 论文的章节安排

第2章 相关技术和开发环境

2.1 相关技术

2.2 开发环境

第3章 系统分析

3.1 可行性研究

3.2 系统需求分析

第4章 系统设计

第5章 系统实现

第6章 系统测试

6.1 测试策略和方法

6.2 测试内容和步骤

6.3 测试结果及分析

第7章 总结与展望

注：正文一般还包括附加的文献、引证、脚注。

第1章 绪论（或引言）

毕业论文的绪论通常包括项目背景、项目意义、项目内容、目标成果等部分。也可以包括论文结构介绍

第2章 相关技术

对应用系统开发中的相关技术做简明扼要的介绍。相关技术应聚焦在与本论文应用开发密切相关或关键技术上，对相关技术的发展状况有系统的综述与评价。文献引用规范，与文后参考文献相对应。

第3章 系统分析

系统分析应该包括需求分析和可行性分析。需求分析识别问题，应该明确以下几个方面的需求

- （1）功能需求：明确开发的软件必须具备什么样的功能；
- （2）性能需求：明确开发的软件的技术性能指标；
- （3）环境需求：明确软件运行时所需要的软、硬件要求

可行性分析清晰阐释软件（系统）开发是可行的。且应用开发环境、需要的条件、软件（系统）的运行环境和已经具备的条件均满足。

第 4 章 系统设计

系统设计建议包括：系统设计目标、系统功能设计、数据库设计。设计目标明确，阐释清晰。数据库设计绘制 ER 图，以及绘制由 ER 图导出的数据库表。用流程图来描述系统功能的总体流程，并介绍系统总体流程的。或者可以或绘制系统功能图，介绍功能之间的关联，并介绍各功能。

第 5 章 系统实现

系统实现主要包括各功能的实现。可以用伪代码或流程图来描述功能完成。

第 6 章 系统测试

系统测试主要包括测试方法，测试内容和步骤，测试结果等内容。

第 7 章 总结与展望

总结与展望一章对所做的工作进行总结。应该客观指出开发的应用所达到的水平，是否解决实际问题，存在的问题，以及未来改进的方向。

6. 参考文献

为了反映论文的科学依据，以及表明作者尊重他人研究成果的严肃态度，并向读者提供有关信息的出处，应在论文之后列出参考文献。

参考文献是论文的最后一部分，指作者在毕业论文工作中所参考的文献，是毕业论文不可缺少的组成部分。毕业论文的引用文献必须在文中引用处体现出来，按作者、文献名、卷册序号、出版地、出版单位、出版时间、章节或起止页码排列。如列出多位作者的参考文献，应按作者姓名的汉语拼音依次排列，外国作者按其姓汉语翻译的拼音排序；同一作者的不同著作按照出版时间先后顺序排列，其著作一律排在论文前面。引用在网络媒体首次发表研究成果的，必须注明网址；对于非首次发表的网络资料，必须注明其首次发表的媒体名称。

参考文献中列出的一般应限于作者直接阅读过的、最主要的、发表在正式出版物上的文献。所参考的私人通信和未公开发表的资料，一般不宜列入参考文献表，可紧跟在引用的内容之后在文内注释或标注在页下。参考文献一律放在论文结论后，不得放在各章之后。每条文献的项目必须完整，诸项缺一不可。

7. 附录

有些材料编入文章主体会有损于编排的条理性和逻辑性，或有碍于文章结构的紧凑和突出主题思想等，但是又有其特定价值的，可以编入毕业论文的附录中，位于全文的末尾。附录是作为论文主体的补充项目，但是，不是必需的组成部分。

理论研究型毕业论文正文写作结构

正文是毕业论文的主体和核心部分建议的正文结构如下：

1 引言（绪论）
1.1 研究背景
1.2 国内外研究现状
1.3 研究内容
1.4 论文组织结构
2 相关工作与理论基础
2.1 相关工作
2.2 理论基础
3 ***算法（模型）的设计
3.1 问题描述
3.2 ***算法
4 实验（仿真）分析
4.1 实验环境
4.2 实验数据
4.3 实验结果
5 总结与展望

注：正文一般还包括附加的文献、引证、脚注。

第1章 绪论（或引言）

一篇学位论文的绪论，大致包含如下几个部分：

研究背景及研究意义，国内外研究现状，研究内容和论文组织结构。

（1）研究背景中需要阐述清楚 2 个大问题：研究的问题“是什么”，为什么选择这个题目来研究，即阐述该研究的现实意义，比如说明该研究对学科发展有贡献，该研究对能够解决什么现实问题，该研究具有一定的前沿性等。

（2）国内外研究现状对本研究主题范围内的文献进行详尽的综合述评，“述”的同时一定要有“评”，指出现有研究成果的不足，讲出自己研究的动机。应简述本

课题在国内外的研究和发展状况；针对课题的实际背景和要解决的问题，对比国内外学者的相关工作，阐述清楚国内外学者对同类问题所采用的研究和解决方法，同时对比这些研究和解决方法的优缺点。当然也可适当简要地介绍一些与本课题有关的预备知识。

注：国内研究现状与第二章的相关工作是有区别的。国内外研究现状描述课题研究的大背景大方向，侧重讲述课题研究的先进性和重要意义，不用涉及具体技术。

(3) 研究内容包括需要解决什么问题，解决该问题采用的理论依据、研究方法和实验基础，预期的结果及其地位、作用和意义。在研究内容的最后需写清楚本研究的创新点或理论与(或)实践意义。如果研究的项目是别人从未开展过的，这时创新性是显而易见的，要说明研究的创新点。但大部分情况下，研究的项目是前人开展过的，这时一定要说明此研究与被研究的不同之处和本质上的区别，而不是单纯的重复前人的工作。

(4) 论文的组织结构是对整篇论文的概述，阐述清楚论文的章节，每一章的研究内容或者介绍章和章间的关系。

第2章 相关工作与理论基础 ●

论文的相关工作指的是与论文所研究内容直接相关的同行研究进展，在行文过程中需要有准确的参考文献支撑。

针对研究的问题，目前他人已经做了哪些方面的研究，提出了哪些算法，简单总结其解决的问题与相关算法，并指出现有模型或算法的不足之处。

注意这里的相关工作不仅仅包含现有工作的简单描述，还要在现有的研究工作中进行对比，指出其中的不同之处，并且说明现有算法的不足之处，本论文要在哪些方面做改进等。

理论基础指的是解决该问题使用的一些基本理论，但不是必需内容。如果论文中所涉及的相关理论，对后续论文的理解需要该知识，则有必要在这一章节进行简要介绍。否则，没有必要进行介绍。

这里的基本理论指的是与所研究问题直接相关的内容，而不应是该领域内所有知识的全部概括。

第3章 **算法（模型）的设计

算法（模型）设计这一章是论文的核心内容，包括所研究问题的描述与定义，解决该问题采用的算法，算法的详细描述，以及算法的复杂度分析等。可以包括问题描述和算法（模型）的设计两部分。

第4章 实验（仿真）分析

实验部分建议包括实验环境、实验数据的介绍、模型评价指标的描述和实验结果的分析。实验结果分析中

（1）包括给出结果，并对结果进行定量或定性的分析。写作要点是：以绘图和（或）列表（必要时）等手段整理实验结果。

（2）除了给出实验结果，另一个重点内容是将论文中提出的算法与已经存在的算法做对比，说明在哪些方面有改进，并说明性能提高的原因。

第5章 总结与展望

总结与展望包括结论部分和展望部分。结论不是研究结果的简单重复，而是对研究结果更深入一步的认识，其内容要点如下：

（1）本研究结果说明了什么问题，得出了什么规律性的东西，解决了什么理论或实际问题；

（2）对前人有关本问题的看法作了哪些检验，哪些与本研究结果一致，哪些不一致，作者做了哪些修正、补充、发展或否定；

展望部分阐述的不足之处或遗留问题。

6. 参考文献

为了反映论文的科学依据，以及表明作者尊重他人研究成果的严肃态度，并向读者提供有关信息的出处，应在论文之后列出参考文献。

参考文献是论文的最后部分，指作者在毕业论文工作中所参考的文献，是毕业论文不可缺少的组成部分。毕业论文的引用文献必须在文中引用处体现出来，按作者、文献名、卷册序号、出版地、出版单位、出版时间、章节或起止页码排列。如列出多位作者的参考文献，应按作者姓名的汉语拼音依次排列，外国作者按其姓汉语翻译的拼音排序；同一作者的不同著作按照出版时间先后顺序排列，其著作一律排在论文前面。引用在网络媒体首次发表研究成果的，必须注明网址；对于非首次发表的网络资料，必须注明其首次发表的媒体名称。

参考文献中列出的一般应限于作者直接阅读过的、最主要的、发表在正式出版物上的文献。所参考的私人通信和未公开发表的资料，一般不宜列入参考文献表，可紧跟在引用的内容之后在文内注释或标注在页下。参考文献一律放在论文结论后，不得放在各章之后。每条文献的项目必须完整，诸项缺一不可。

注意：未在正文脚注中出现的文献，一般不列入参考文献之中。

7. 附录

有些算法、程序等编入文章主体会有损于编排的条理性和逻辑性，或有碍于文章结构的紧凑和突出主题思想等，但是又有其特定价值的，可以编入毕业论文的附录中，位于全文的末尾。附录是作为论文主体的补充项目，但是，不是必需的组成部分。

毕业论文撰写遵照《西南财经大学本科生毕业论文（设计）撰写与印制规范》执行。具体情况由指导教师根据专业特点和论文体裁具体掌握。总体上，毕业设计撰写要求可参考如下标准：

研究问题清晰、明确且有价值；论点鲜明，论证充分，论据确凿；方法适当，逻辑性强，对问题有较深入的分析，实证过程严谨且工作量饱满；论文注释、引用和写作体例符合通行规范；文献综述完整，参考文献具有适合性、准确性、专业性和权威性；结构严谨，层次分明，语言表达准确。

实证学术论文评价标准参照学校统一的评价指标及评议要素（见表）。具体地评价细则见下表 2。

二、计算机专业毕业设计撰写要求及评价细则

毕业论文撰写遵照《西南财经大学本科生毕业论文（设计）撰写与印制规范》执行。具体情况由指导教师根据专业特点和论文体裁具体掌握。应用开发型毕业设计要求选题有理论或者实践意义，符合本学科专业的要求；系统分析和设计应有明确思路，相关技术选择恰当，工作量饱满；理论研究毕业论文要求选题要体现学生专业知识与专业技能的综合运用和科研创新能力的培养；学生在规定时间内工作量饱满，能对选题进行充分研究论证。具体评价标准见下表：

表 毕业设计评分标准

评议要素	优秀	良好	一般	不及格
选题意义	选题合理,属于该专业领域研究范畴;有明确的现实背景;研究目的及研究问题明确,研究具有显著的实践价值或学术意义。	选题合理,属于该专业领域研究范畴;研究目的较为明确;具有较大的实践价值或学术意义。	选题尚可,属于该专业学位领域研究范畴,具有一定的实践价值或学术意义。	选题存在一定问题,研究目的不清,不具备实践价值或学术意义。
写作安排	全面、综合分析国内外相关文献,追踪本领域最新研究现状,文献质量很高;相关技术介绍清晰、完整;时间进度安排合理;研究工作量饱满,有一定的研究或技术开发难度。	较为全面地分析国内外相关文献,追踪本领域研究基本现状,文献质量较高;相关技术介绍较为清晰、完整;时间进度安排合理;研究工作量较为饱满。	基本了解国内外相关文献和本领域研究基本现状,文献质量一般;时间进度安排较为合理;工作量一般。	不能全面、综合分析国内外文献,未追踪本领域研究现状,文献质量较低,时间进度安排不合理;工作量不足。
逻辑建构	从提出、分析到解决问题过程完整;研究思路清晰,有一定的提炼、分析和归纳能力,具有独立见解;开发技术使用和流程规范;研究方法规范。	从提出、分析到解决问题过程较为完整;研究思路比较清晰;开发技术使用和流程规范;研究方法比较规范。	从提出、分析到解决问题过程较为完整;技术使用和流程比较规范;研究方法比较规范。	提出、分析、解决问题的论证过程不完整;技术使用和流程不规范;研究方法缺乏规范。
专业能力	较好应用专业原理或理论,体现专业视角;较好运用专业研究方法或工具进行研究;有足够的编程的工作量,数据丰富、可靠。	应用了一定的专业原理或理论;运用了专业研究方法或工具进行研究;有一定的编程的工作量,数据比较丰富、可靠。	运用了专业研究方法或工具进行研究;编程工作量一般,数据不够丰富、欠缺可靠性。	欠缺对专业研究方法或工具的运用;工作量欠缺,欠缺数据资料。
学术规范	文字表达、书写格式、图表、公式符号、缩略词等方面严格符合通行学术规范;引用非常符合通行学术规范和知识产权相关规定。	文字表达、书写格式、图表、公式符号、缩略词等方面较好地符合通行学术规范;引用较好地符合通行学术规范和知识产权相关规定。	文字表达、书写格式、图表、公式符号、缩略词等方面基本符合通行学术规范;引用基本符合通行学术规范和知识产权相关规定。	文字表达、书写格式、图表、公式符号、缩略词等方面不符合通行学术规范;引用不符合通行学术规范和知识产权相关规定。

三、范文讲解

因两种类型写作范式不同，下面分别介绍两种类型的毕业论文的范本。以下范文讲解取自多篇优秀毕业论文部分内容。

（一）应用开发类型毕业设计

论文题目：校园社团管理 **app** 的设计与实验

1.选题：论文选题契合专业领域范畴。

2.题目：毕业设计要解决的问题和目标清晰。

摘要

大学生在刚步入校园的时候,常常会苦恼于无法找到志同道合的人一起交流,而社团正是将一群志趣相同的人聚集在一起的组织。但是大学校园中的社团组织的宣传渠道有限,传统的方式仍为通过发传单宣传社团。很多学生因为上课时间冲突错过了社团的宣讲活动,或是不了解社团的日常活动,因此错过自己心仪的社团。基于这些,本人研发了一个校园社团管理 APP,该 APP 主要实现:学生端通过注册登录,完善自己的基本信息,浏览各个社团的详情介绍后可选择关注或填写申请入社书,在广场中也可浏览各个社团发布的活动;社团管理人员可以进行招新审批、社员管理和发布活动。对于学生和社团双方来说都极大地节省了招新入社的时间,学生能更方便地了解社团的活动内容,社团也能免去复杂的管理程序,进行高效的社团管理。

本人在分析了校园社团管理 APP 用到的开发技术和开发理论之后,使用 Android Studio 平台,依赖 Bmob 云后台数据库。首先确定了系统的需求和功能模块划分,在进行可行性分析之后,设计页面 UI 和数据库,然后再实现逻辑功能的完善,进行系统的优化和测试,确保 APP 能够正常地运行。

关键词: 学生社团管理; Android studio; 数据库

3.摘要及关键词:

(1) 摘要语言简练,准确、流畅,具有独立性和自含性,能概括论文的研究目的、方法、结果、结论等。不与“绪论”、“结论”等部分内容雷同;摘要的时态以一般现在时为主,表示一种存在于自然界的客观规律。

(2) 关键词: 3-5 个,准确反映论文主要内容

(3) 摘要部分应该包括下面几部分的内容:

第一部分:系统所依赖的背景;

第二部分:描述开发了什么系统,解决当前背景下什么问题

第三部分:系统实现了 XXX, YYY, ZZZ 等功能,有何亮点特色

第四部分:软件测试结果怎样?是否符合预期?

摘要总长度约 600-700 字(大约控制 1 页内)

Abstract

When college students just enter the campus, they are often worried that they can't find people with the same interests to communicate with, and the club is an organization that gathers a group of people with the same interests. However, the publicity channels of community organizations in university campus are limited, and the traditional way is to publicize the community by distributing leaflets. Because of the conflict of class time, many students miss the propaganda activities of the club, or do not understand the daily activities of the club, so they miss their favorite club. Based on these, I have developed a campus community management app, which mainly realizes: students can register and log in, improve their basic information, browse the details of each community, then choose to pay attention to or fill in the application form, and browse the activities released by each community in the square; community managers can conduct recruitment approval, member management and release activities. For both the students and the club, it greatly saves the time of recruiting new members. Students can more easily understand the activities of the club, and the club can also manage the club without complicated management procedures.

After analyzing the development technology and theory of campus community management app, I use Android studio platform and rely on bmob cloud background database. Firstly, the requirements and functional modules of the system are determined. After the feasibility analysis, the page UI and database are designed, and then the logical functions are improved, and the system is optimized and tested to ensure the normal operation of the app.

Key words: student community management; Android studio; database

(4) 英文摘要及关键词: 英文翻译准确、流畅, 不存在语法、单词错误。

目录

1.绪论.....	1
1.1 研究背景	1
1.2 设计目标	1
1.3 本文结构	2
2.系统建构的相关方法和技术	3
2.1 系统开发环境	3
2.2 Bmob 后端云	3
2.2.1 Bmob 后端云简介	3
2.2.2 Bmob 后端云特点	3
2.2.3 Bmob 后端云存储数据类型	3
2.2.4 Bmob 后端云的缺点	4
2.3 Android	4
2.3.1 Android 操作系统简介	4
2.3.2 Android 存储访问机制	5
3.系统分析	6
3.1 系统需求分析	6
3.2 功能模块分析	6
3.3 系统可行性分析	7
3.3.1 技术可行性分析	7
3.3.2 系统可用性分析	7
4.系统设计	8
4.1 系统设计目标	8
4.1.1 系统技术性能目标	8
4.1.2 UI 设计目标	8
4.2 数据库设计	9
4.2.1 E-R 模型	9
4.2.2 具体数据表设计	10
4.3 设计难点	12
4.4 解决方法	12
4.5 系统总体流程	13
4.6 系统业务流程图	14
5.系统实现	16
5.1 开屏引导	16
5.2 登录界面	17
5.3 注册界面	18
5.4 个人信息界面	20
5.5 社团总览页面	21
5.6 社团详情页面	22
5.7 申请书页面	24
5.8 管理员界面	25
5.9 招新生	26
5.10 办活动	28
5.11 添加管理员	29
5.12 广场	30
5.13 活动详情	32
5.14 本章小结	33
参考文献	34
致谢	35

4.论文结构及目录

(1) 论文结构：论文总体结构合理，重点突出，逻辑清晰。满足应用开发型毕业设计的论文结构要求。

(2) 目录：格式规范、美观，各级标题准确；目录是论文的导读图，要求具有完整性。

建议的论文结构如下。

第1章 绪论

1.1 项目背景与意义

1.2 设计目标

1.3 论文的章节安排

第2章 相关技术和开发环境

2.1 相关技术

2.2 开发环境

第3章 系统分析

3.1 可行性研究

3.2 系统需求分析

第4章 系统设计

第5章 系统实现

第6章 系统测试

6.1 测试策略和方法

6.2 测试内容和步骤

6.3 测试结果及分析

7章 总结与展望

参考文献

致谢

1.绪论

1.1 研究背景

在学校的日常生活中，除了学习，社交生活也是大学生必不可少的。西财素有“西南文体活动大学”之称，每年的“百团大战”也会吸引众多新生参与，因此参加社团是西财学生扩展自己的社交圈的重要途径之一。但是在参与社团活动的同时，总会面临很多问题，或者在百团大战中被众多社团所吸引但又不知道哪一个更适合自己的，或者在加入心仪的社团后发现无法及时获取社团开展的活动信息，又或者在加入社团后发现缺少交流的机会仍然无法扩展社交圈等等。

与此同时，很多社团举办的活动因为场地原因会出现人数限制，若不及时与同学们沟通交流会导致参加活动的人数超过上限，使活动开展难度加大。

因此，开发一个具有个人信息完善、招新、关注、活动发布等功能的APP，能够有效地解决大学校园中的社团管理问题，提高招新的效率，增强学生与社团之间的互动，使得学生在校中能够更方便快捷地参与社团活动。

1.2 设计目标

针对这些大学生所苦恼的社团问题，我分析并设计了一款基于Android手机的简洁的校园社团管理应用，主要任务是简化社团管理人员和普通学生在社团互动中的工作，使社团的招新和管理更加高效，轻松发布活动让社团管理人员无需再使用大量人力进行纸质传单宣传。本APP能够让学生完善自己的个人信息，浏览所有的社团详情，选择自己喜欢的再填写入社申请书，同时也可以查看社团发布的活动，为活动留言点赞或报名参加，而社团管理人员需要通过管理员权限进入管理员界面，查看申请人的基本信息和申请理由进行新成员入社审批，另外也可以发布活动吸引更多学生入社，若工作量大，还可从普通社员中添加管理员共同

对社团进行管理。

APP满足的要求大致有以下几个方面：

- (1) APP设计清新简洁，简单易操作。
- (2) 实现了学生和社团之间的信息平等交流。
- (3) 运行基本稳定，给用户带来良好的体验。
- (4) 可操作空间很多，便于后续版本的更新和维护。

1.3 本文结构

第一章 绪论的主要内容为本APP应用的研究意义和研究背景，还有所需要解决的具体问题和目标。

第二章 本APP构建的开发环境与技术介绍。

第三章 系统分析，介绍了对于社团管理所需的具体信息分析、功能模块分析，以及系统可行性分析。

第四章 系统设计，介绍了系统的设计目标、数据库设计、设计难点和对应的解决方法、系统总体流程以及业务流程图。

第五章 系统实现，详细介绍了本APP的实现过程和主要代码以及技术难点，给出了每个界面的实现效果和截图。

第六章 系统测试，介绍了测试流程和结果。

5.绪论（或引言）

(1) 开门见山，直奔主题：不过多赘述历史渊源和历程，介绍文章的现实背景和学术背景。

(2) 言简意赅，突出重点：清晰阐明研究问题

(3) 尊重科学、实事求是：评价论文的价值要恰如其分、实事求是，用词要科学，对论文的创新性最好不要使用“本研究国内首创、首次报道”、“填补了国内空白”、“有很高的学术价值”、“本研究内容国内未见报道”或“本研究处于国内外领先水平”等不适当的自我评语。

(4) 符合规范，结构完整：毕业论文的绪论通常包括项目背景、项目意义、项目内容、目标成果等部分。也可以包括论文结构介绍

2.系统建构的相关方法和技术

2.1 系统开发环境

本系统软件环境配置为 JDK1.8.0_152，数据库依赖于 Bmob 后端云，开发工具是 Android Studio 操作系统，服务器操作系统为 Windows 10。

2.2 Bmob 后端云

2.2.1 Bmob 后端云简介

Bmob 云平台能为移动应用提供一个完整的后端解决方案，利用官方提供的 SDK，在 App 中进行调用，便可以很方便地进行网络数据的存储和下载。

2.2.2 Bmob 后端云特点

随着 Android 系统在手机和平板市场的普及以及 AndroidApp 越来越多地出现，AndroidApp 也吸引了很多开发者的加入。传统的网络数据存储、用户管理、消息推送等功能都需要数据库以及服务器的支持，个人开发者要是想在 App 中加入这些功能，就必须学习 Java、xml 语言之外的其他语言，增加了开发 App 的难度，并且搭建服务器也额外增加了开发成本。Bmob 云平台简化了这一过程，使开发者能够将更多的精力投向软件功能的改善和界面的美化设计，使应用的开发变得更加简单。

2.2.3 Bmob 后端云存储数据类型

6.相关技术

对应用系统开发中的相关技术做简明扼要的介绍。应该满足以下条件

(1) 问题导向性：相关技术应聚焦在与本论文应用开发密切相关或关键技术上；

(2) 表达清晰：能够清晰、合理介绍某项技术的原理，复杂的内容可以用图来展示，并对图上的内容做介绍。

(3) 逻辑性：对相关技术的发展状况有系统的综述与评价。

(4) 饱满性：文献数量饱满，没有重要文献缺失。

(5) 规范性：文献引用规范，与文后参考文献相对应。

3.系统分析

3.1 系统需求分析

大学社团信息一般包括的有社团介绍、活动发布、成员管理等,而对于普通用户而言,个人基本信息管理、社团选择以及对活动的看法发布等更为重要,因此本社团管理 APP 主要针对的就是这些目标来进行设计和实现。

(1) 个人基本信息:用户注册时需要完善自己的基本信息,包括姓名,学号,性别,专业,出生年月日等,然后才可以进行其他的功能的实现。

(2) 社团信息:每个社团都有自己的基本信息介绍,除此之外还有普通社员和管理员的区别,另外社团还需要发布活动来吸引更多的学生加入社团,组织社员进行活动,发布活动的同时由于场地等原因要限制参与活动的人数,当报名人数满后其他未报名的同学则无法参加该活动。

(3) 广场信息:成立社团的目的就是为了有更多的相同兴趣的人聚在一起交流,学生们可以及时地看到这些活动的内容并报名参与。

3.2 功能模块分析

(1) 开屏引导模块:三个介绍引导页面,浏览或跳过后进入登录页面。

(2) 登录注册模块:功能包括用户名、密码以及个人基本信息的填写注册、登录进入主页面。

(3) 个人信息模块:查看已加入社团等个人基本信息。

(4) 社团模块:社团详情浏览,普通用户关注社团、申请加入社团、填写申请书以及查看申请状态;社团管理员可进入管理员界面,进行申请书的审批、活动的发布以及其他管理员的添加功能。

(5) 广场模块:浏览社团发布的活动,可进行点赞和评论以及根据活动人数限制报名参与活动。

3.3 系统可行性分析

3.3.1 技术可行性分析

本系统在开发中用到了 Android Studio 集成开发环境,依赖于 Bmob 云后台数据库,能够适应中小型系统的开发。具有稳定的数据存取和更新能力,在逻辑数据处理和应用界面显示方面都十分优秀。分析结果说明技术方面是可行的。

3.3.2 系统可用性分析

大学生在进入校园后,缺少与其他志同道合的人交流的渠道,并且常常因为上课时间冲突而无法获得了解社团、参加社团的机会。本系统应用于安卓平台,学生可轻松应用于自己的手机,查看各个社团的详情,提高了社团的宣传和招新效率,减少了学生加入社团的难度,因此本系统的设计开发十分适用于大学校园中的社团使用。

本系统的开发基于 Java 语言,便于后续的版本更新和维护,应用在安卓手机上,方便于社团管理人员和学生可随时随地在手机上进行操作。同时,界面设计可爱简洁,易于操作,为使用本系统的人在进行社团活动的时候带来了极大的便利。

7.系统分析应该包括需求分析和可行性分析

需求分析识别问题,应该明确以下几个方面的需求

(1) 功能需求:明确开发的软件必须具备什么样的功能;

(2) 性能需求:明确开发的软件的技术性能指标;

(3) 环境需求:明确软件运行时所需要的软、硬件要求

可行性分析应满足以下条件

(1) 可行性:清晰阐释软件(系统)开发流程,以说明毕业设计是可行的。

(2) 合理性:清晰阐释应用开发环境、需要的条件、软件(系统)的运行环境和已经具备的条件,以说明软件(系统)是可以开发出来的,且开发出的软件(系统)具备运行条件。

4.系统设计

4.1 系统设计目标

4.1.1 系统技术性能目标

因为社团管理 APP 是为了服务在校学生与社团互动，实现招新和互动的方便性和简洁性，提高学生参与社团活动、管理社团的效率，所以需要具有稳定快速的的数据处理能力的数据库作为基础，因此本 APP 在技术方面上的设计目标如下：

(1) 数据库稳定：该系统依赖的是 Bmob 后端云作为后台来进行开发。开发时将数据存储在 Bmob 后端，需要时再获取，提高了系统性能。Bmob 后台相比于本地数据库在功能扩展和性能优化方面都具有优势，不仅可以存文字信息，并且支持文件上传，下载。Bmob 云后台很适合移动开发者做一些轻量级的应用，非常适合本项目的完成。

(2) 可扩展性保证：Android studio 的基本语言为 JAVA，面向对象开发的特点保障了后期的维护便利性和可操作性，可轻松进行其他功能的完善和继续开发。

(3) 用户体验保证：页面设计简洁，让用户看着心情舒适没有其他冗余的复杂信息，操作易懂使得学生在进行与社团之间的互动时方便了许多。

(4) 操作准确性保证：在用户注册输入密码时进行二次确认，防止用户偶然输错导致后续无法正常登陆的情况出现。

4.1.2 UI 设计目标

UI 界面就是 APP 的显示界面，采用 xml 布局设计。UI 设计将直接影响用户的使用体验，因此从以下两方面设计 UI 实现目标：

(1) 功能方面：社团管理 APP 主要分为三个模块：个人信息、社团和广场，因此在底部导航栏设计这三个板块可随时切换。对于管理员而言，在对应的社团

详情中区别普通学生会显示管理员图标，为管理员提供招新、发布活动、社团管理的权限。UI 界面的按钮都使用了对应功能的图标，使没有查看过相关文档的人也能够方便地使用该 APP。

(2) 风格方面：因为适用人群为大学生，因此设计风格整体灵动跳跃，引导界面和头像的设计则较为可爱，而对于管理界面则较为简洁易懂。既能够吸引众多学生，又能够方便用户的使用。

4.2 数据库设计

4.2.1 E-R 模型

数据库逻辑关系如图 4-1 所示：

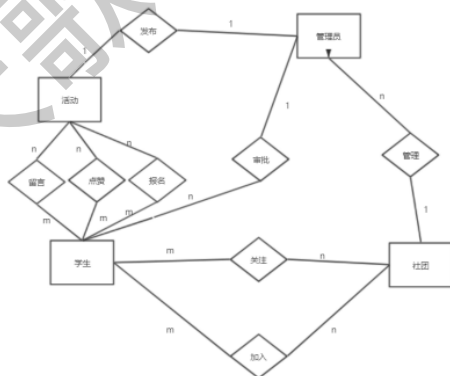


图 4-1 E-R 图

8.系统设计应包括：系统设计目标、系统功能设计、数据库设计。

(1) 设计目标明确，阐释清晰。

(2) 数据库设计应该绘制 ER 图，以及绘制由 ER 图导出的数据库表。对表的内容做必要的介绍。

参照上图：管理员可以发布活动；管理员也可以对学生进行审批，一位管理员可审批多个学生；管理员可对社团进行多对一的管理；学生可对活动进行点赞或者留言或者报名，一个学生可对多个活动进行点赞留言或报名，并且可以对一个活动多次留言；学生也可以对社团进行关注或加入，一个学生可关注或加入多个社团。

4.2.2 具体数据表设计

本系统设计的主要数据表如下所示：

(1) 用户信息表如表 4-1 所示：

表 4-1 用户信息表

字段	名	类型
学号	UserID	Number
密码	Password	String
专业	Major	String
出生日期	Birthday	Date
性别	Sex	String
真名	RealName	String

用户信息表中储存了用户的基本信息。

(2) 社团成员申请信息表如表 4-2 所示：

表 4-2 社团成员申请信息表

字段	名	类型
是否为管理员	Admin	Boolean
申请理由	Reason	String
审批状态	Apply_condition	String
学号	UserID	Number

当用户提交了入社申请书后便会将个人基本信息和申请理由存在社团成员信息表中，同时将审批状态置为“待审批”，是否为管理员置为“false”。当管理员进行新社员入社审批后状态会更改为“不通过”或者“已通过”，当审批状态为“已通过”时，用户将在社团详情中查看到入社状态变成“已加入”，若审批状态为“不通过”，用户将在社团详情中查看到入社状态变成“未通过，可再次提

表 4-5 留言信息表

学号	UserID	Number
内容	Content	String
留言编号	MessageID	Number
活动编号	ActivityID	Number

(6) 点赞信息表如表 4-6 所示：

表 4-6 点赞信息表

学号	UserID	Number
点赞编号	FollowID	Number
活动编号	ActivityID	Number

4.3 设计难点

- (1) 打开应用若直接进入登陆界面略显突兀，难以给新用户留下对本 App 的特殊印象，若强制每个用户都需经过引导页面才可进行登录注册，则会浪费部分老用户的时间，使引导界面变成累赘。
- (2) 用户首次登陆时未进行个人信息的完善，若在此基础上进行社团申请或对活动留言则会出错。
- (3) 当用户提交入社申请书后，无法得知自己的审批状态，容易重复提交申请。
- (4) 一个管理员无法有足够的时间和精力管理一个社团。
- (5) 广场中的活动若只有标题和简要内容则无法体现活动的热度。
- (6) 很多社团活动由于场地等原因要求限制人数，而报名者的数量可能会超过限制人数上限。

4.4 解决方法

- (1) 打开应用即进入引导界面，用户需要对三个引导页面左右滑动查看后再

必要的话,可以阐述设计难点以及解决方法。

进入登陆界面，若用户不想浏览引导界面，也可点击跳过直接进入登陆界面。既可以让新用户对此应用留下印象，准备登录注册的缓冲时间，也解决了想直接登陆进行社团活动的用户的需求，跳过引导界面直接登陆注册，节省时间。

(2) 在注册时就要求用户填写相关基本信息，防止在登录后进行活动留言或报名活动时出现错误。

(3) 用户提交申请书后，在该社团详情界面可看到自己的申请状态，当管理员还未审批的时候，状态会变更为“待审批”，当管理员进行审批后，若通过，则该申请人的社团详情界面则显示“已加入”，对应图标也被点亮，若不通过，则该申请人的状态显示为“审核未通过，可再次提交申请”，则该申请人可再次填写入社申请书提交。

(4) 设置管理员可以从普通社员中挑选自己信赖的人作为新的管理员协助自己管理。

(5) 广场中的活动列表除了标题和简要内容外同时也显示浏览量和点赞量，用户在浏览广场时可直接看出活动的热度和受欢迎程度。

(6) 管理员发布活动的时候需要设置限制人数，用户在广场浏览活动时可看到该活动的已报名人数和限制人数，若该活动报名人数已超过人数上限则无法报名。

4.5 系统总体流程

打开本 APP，进入引导界面，可左右滑动三个引导界面也可直接点击跳过按钮直接进入登陆界面。

新用户可在注册界面输入学号即用户名并且确认两次密码一致，填写个人基本信息包括真实姓名，专业，性别和出生日期进行账号注册。

普通学生端登录后可在个人信息界面查看已加入的社团和个人基本信息；在社团总览界面可点击社团 logo 进行详情查看，可关注社团或填写入社申请书；在广场界面可浏览活动的点击量和点赞量以及已报名人数和限制人数，点击活动

可查看详情，留言点赞或报名参加该活动。

管理员登录后可在社团详情界面点击管理员按钮进入管理界面，在招新生界面中可查看待审批与已审批的申请书，点击申请书查看申请人基本信息和申请理由，对其进行审批；在办活动界面可发布新活动于广场中；在添加管理员界面可查看所有普通社员和管理员，点击普通社员将其设置为新管理员协同管理。

4.6 系统业务流程图

系统业务流程图如图 4-2、4-3 所示。

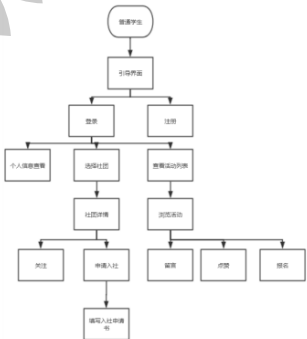


图 4-2 普通学生系统业务流程图

用流程图来描述系统功能的总体流程，并介绍系统总体流程的。或者可以绘制系统功能图，介绍功能之间的关联，并介绍各功能。

5.系统实现

5.1 开屏引导

界面介绍：可通过向左或向右滑动至不同的引导界面，共有三个引导界面，如图 5-1 所示，也可以点击右上角的跳过直接进入登录界面。

实现难点：开屏引导左右滑动的效果实现借助了 ViewPager 视图翻页工具，实现了多页面切换效果，每个页面为一个 fragment，并监听页面的滑动事件，实现视差动画。



图 5-1 开屏引导界面

服务端在进行文章内容的传递时，分为两种情况：当日文章信息和历史文章信息。考虑到结果不止一条，单个的 JSONObject 不能满足，需要对各条信息进行重组，使其成为一个完整的 Json 数据。获取文章信息的流程图参见图 3-6。

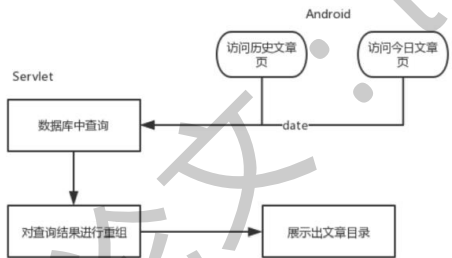


图 3-6 获取文章目录流程图

9.系统实现主要包括各功能的实现。建议不要直接使用源代码，可以用伪代码来描述功能完成，或者用流程图来描述功能的完成。并辅以对伪代码或流程图工作的解释。如下图所示。应该提供软件运行界面的截图。

6. 系统测试

本章将介绍测试的主要方法，并给出针对本软件的具体测试实例。

6.1 测试方法

测试贯穿在整个开发过程中，但为了验证开发的软件是否满足事先预期的功能，还需要制定测试计划，对软件进行单元测试、集体测试和系统测试。

黑盒测试：不考虑系统内部的情况，只根据输入输出进行判断，且不会破坏外部信息。

白盒测试：直接在程序各模块内部测试，在代码层次判断其正确性。

6.2 测试用例

(1) 登陆测试

用户登陆功能测试用例表参见表 6-1。

表 6-1 用户登录功能的测试

用例编号	操作内容	正确结果
1	使用正确的用户 id 和密码后点击登陆	验证通过，进入今日文章目录页
2	填写错误的密码	弹出密码错误提示，停留在登陆页
3	点击取消	退出 APP
4	返回	退出 APP

6.3 测试结论

进行上述黑盒测试，可将客户端和服务端所有实现的功能都覆盖到。通过测试结果，可以得出本软件能够正常登陆，能够准确获取文章目录，能够在线浏览文章内容，能够在线播放音频，运行过程中无闪退等异常情况出现。同时通过输出到客户端的结果，也能得出客户端与服务端数据交互的正确性。

10.系统测试主要包括测试方法，测试内容和步骤，测试结果等内容。

7. 总结与展望

本文选择基于 Android 系统的在线英语学习软件开发作为研究内容, 综合应用了 Android 开发的各工 具, 以及本地服务器的搭建。最终实现了登陆, 在线 浏览文章内容, 在线播放音频文件的基本功能, 完成了一个可以正常使用的 APP。

软件现阶段还未完全达到最初的设想, 它并未实现学习进度记录这一功能。所以在下一阶段, 打算对此功能板块进行具体的设计和开发。将会涉及从客户端到服务端的数据传递以及数据库的更新操作等。

另外, 本地服务器中的文章资源均为手动添加, 并未实现通过服务端定时爬取网站上每日更新的数据。为了软件的科学性, 应继续对这一块进行开发。将会涉及到网页的解析, 在线下载 mp3 文件, 对文件名的管理及数据库的写入操作。

11. 总结与展望一章对所做的工作进行总结。应该客观指出开发的应用所达到的水平, 是否解决实际问题, 存在的问题, 以及未来改进的方向。

参考文献

- [1] 周冉, 高玉竹. Bmob 云平台在 AndroidApp 开发中的应用[J]. 微型机与应用, 2015.
- [2] 杜巍. Android App 开发中 Bmob 云平台的运用[J]. 通讯世界, 2016.
- [3] 黄鑫. 基于 Android 的大学生个人课程助理系统的设计与实现[D]. 厦门大学, 2014.
- [4] 黄莹. 基于 Android 平台智能手机多方通话软件测试系统的研究与实现[D]. 华中师范大学, 2013.
- [5] 祝忠方. 基于 Android 的移动互联终端的设计和实现[D]. 北方工业大学, 2014.
- [6] 倪红军. 基于 Android 系统的数据存储访问机制研究[J]. 计算机技术与发展, 2013.

12. 参考文献:

(1) 规范性: 严格按照学校学位论文要求编排参考文献以及每条参考文献的样式。

(2) 对应性: 凡是前文中引用了别人成果

(文字、数字、事实以及转述他人的观点)之处, 均应加标注列于参考文献中。

(3) 前沿性: 参考文献需能看出论文较好地跟踪了近期的该领域研究动态。

(二) 理论研究类型毕业论文

论文题目：基于可解释机器学习的谣言检测研究

1.选题：论文选题契合专业领域范畴，瞄准学科发展前沿和关键领域。

2.题目：研究对象非常明确，研究问题非常清晰，研究意义显著。

摘要

现代社会互联网技术的发展愈发迅猛，各式各样的社交媒体成为了信息产生与交换的一种应用最为广泛的渠道，同时也成为了虚假不实信息产生与传播的理想场所。网络谣言具有一定的煽动性和危害性，破坏着网络环境和秩序。因此，对网络谣言进行有效鉴别，对改善网络环境，维护网络中的安全与稳定，具有一定的积极意义。现有的谣言检测算法经过前人的研究，得到了很大的发展，同时取得了一些成果，但由于存在各种条件的限制，得到的这些成果会有些不足，还可以进行进一步的提升。比如，现有的谣言检测算法只是判断一条信息是否为谣言，并不能对其成为谣言的原因；还有，在实际情况中，谣言数据集属于一种不平衡的数据集，但多数的谣言检测研究并没有考虑到这个问题，这样会造成检测算法准确率的降低。本文提出的基于可解释机器学习的谣言检测模型正是为了解决以上这些问题的。

本次研究模型采用循环生成对抗网络（Cycle-consistent Generative Adversarial Networks, CycleGAN）这一新的框架，循环生成对抗网络由两个生成器和两个判别器组成，生成器可以生成少数类别的样本数据，达到数据增强的效果，提升训练数据集的容量，为解决数据不平衡性的问题提出了相应的方案。使用两个生成器和一个判别器实现谣言数据到非谣言数据的转换，学习谣言数据的指标性特征，不断提高模型的学习能力，实现模型的可解释性。本次研究通过该模型与其他的谣言检测算法进行了对比验证，确认了本次研究设计的谣言检测模型具有较好的效果。

关键词：谣言检测；循环生成式对抗网络；可解释性；数据不平衡

3.摘要及关键词：

（1）摘要语言简练，准确、流畅，具有独立性和自含性，能概括论文的研究目的、方法、结果、结论等。不与“绪论”、“结论”等部分内容雷同；摘要的时态以一般现在时为主，表示一种存在于自然界的客观规律。

（2）关键词：3-5个，准确反映论文主要内容

（3）建议摘要的写法：摘要一般分为2-3段，字数在300~500之间。不要出现第一人称我或我们的字样，要从客观的角度来阐述。

第一段：简述论文背景，引出为什么要研究该项目（意义）。

第二段：是摘要的主要内容，对全文进行总概。一般按照论文的顺序进行阐述。

第三段：可从论文的特点和贡献上进行一下总结

Abstract

The development of Internet technology in modern society is becoming more and more rapid, all kinds of social media have become the most widely used channels for information generation and exchange, and at the same time, it has become an ideal place for the generation and dissemination of false and false information. Network rumors have certain incendiary and harmful, destroying the network environment and order. Therefore, the effective identification of network rumors is of positive significance to improve the network environment and maintain the security and stability of the network. The existing rumor detection algorithm has been greatly developed by previous research, and some achievements have been made, but due to the limitation of various conditions, these results will be somewhat inadequate, and can be further improved. For example, the existing rumor detection algorithm only determines whether a message is a rumor, and cannot be the cause of the rumor; Also, in practice, the rumor data set belongs to an unbalanced data set, but most rumor detection studies do not take this problem into account, which will lead to a decrease in the accuracy of the detection algorithm. The rumor detection model based on explanatory machine learning is proposed in this paper to solve these problems.

This research model uses the new framework of circular generation adversarial networks (CycleGAN), which consists of two generators and two discriminators, which can generate a small number of categories of sample data to achieve data enhancement and increase the capacity of training data sets. In order to solve the problem of data imbalance, the corresponding scheme is put forward. Using two generators and a judge to realize the conversion of rumor data to non-rumor data, learn the index characteristics of rumor data, continuously improve the learning

ability of the model, and realize the explanatory nature of the model. This study verifies the rumor detection model with other rumor detection algorithms by comparing the model, and confirms that the rumor detection model designed by this study has good effect.

Keywords: Rumor detection; Cycle-generated against the network; Explanatory; The data is unbalanced

(4) 英文摘要及关键词: 英文翻译准确、流畅, 不存在语法、单词错误。

目录

1. 绪论.....	1
1.1 研究背景及研究意义.....	1
1.2 研究现状.....	2
1.2.1 人工检测方法.....	3
1.2.2 基于机器学习的检测方法.....	4
1.2.3 基于深度学习的检测方法.....	5
1.3 研究内容与结构安排.....	7
1.3.1 研究内容.....	7
1.3.2 结构安排.....	8
2. 相关理论概述.....	9
2.1 不平衡数据概述.....	9
2.2 机器学习的可解释性.....	10
2.3 生成式对抗网络.....	11
2.4 循环生成式对抗网络.....	13
3. 基于可解释机器学习的谣言检测算法.....	15
3.1 问题描述.....	15
3.2 网络结构的设计与实现.....	16
3.2.1 模型的整体架构.....	16
3.2.2 生成器模型.....	17
3.2.3 判别器模型.....	18
3.2.4 损失函数.....	19
3.2.5 谣言检测算法实现流程.....	20
3.3 实验设计与分析.....	22
3.3.1 数据集.....	22
3.3.2 实验设置.....	24
3.3.3 实验结果分析.....	27
4. 总结与展望.....	32
4.1 研究总结.....	32
4.2 研究展望.....	33
参考文献.....	35
致谢.....	39

4.论文结构及目录

(1) 论文结构：论文总体结构合理，重点突出，逻辑清晰。满足应用开发型毕业设计的论文结构要求。

(2) 目录：格式规范、美观，各级标题准确；目录是论文的导读图，要求具有完整性。建议的论文结构如下。

目录

1 引言（绪论）

1.1 研究背景

1.2 国内外研究现状

1.3 研究内容

1.4 论文组织结构

2相关工作与理论基础

2.1 相关工作

2.2 理论基础

3 ***算法（模型）的设计

3.1 问题描述

3.2 ***算法

4 实验（仿真）分析

4.1 实验环境

4.2 实验数据

4.3 实验结果

5 总结与展望

参考文献

1. 绪论

1.1 研究背景及研究意义

互联网的蓬勃发展使得以新浪微博、Twitter、Facebook 等为代表的社会网络已成为当今人们消费和分享新闻的一个重要平台，深刻地影响了人们的生活方式。社交媒体已成为全世界人们日常生活中不可或缺的一部分，根据 We Are Social 与 Hootsuite 合作发布的最新《2020 全球数字报告》^[28]显示，全球使用互联网的人数已增长至 45.4 亿，与 2019 年 1 月相比增长了 7%（2.98 亿新用户），到 2020 年 1 月，全球社交媒体用户达到 38 亿，自去年同期以来，这一数字增长了 9% 以上（3.21 亿新用户）。实际上，互联网的普及与应用，社交媒体平台的爆发式增长在其发挥了关键作用。与传统媒体（如报纸、广播和电视）不同，社交媒体作为新媒体的成员，打破了生产、消费、传播的界限。用户既可以是内容生产者，又可以是消费者，同时还可以是传播者，用户可以随意发表自己的观点和看法，或者浏览、转载、点赞别人的观点或文章，大大提高了用户的参与度。社交媒体平台每天都产生海量的信息，这些信息具有传播速度快、范围广、即时性强等优点，社交媒体平台给人们带来信息交流和分享便捷的同时，由于其缺乏对信息发布的有效监管机制和手段，使得在社交媒体中传播的信息的质量参差不齐，其中就不少掺杂着一些故意捏造事实的虚假新闻或谣言信息。根据《微博辟谣 2020 年度报告》^[5]显示，2020 年度，微博站方共有效处理 76107 条不实信息，新增 782 条谣言案例，平均每一条不实信息微博被 34 个网友举报，标记不实信息 2642 条。一般来说，假新闻^[18]的定义为：假借新闻报道形式传播的错误虚假、耸人听闻的信息。“假新闻”的出现率大大增加的现象是从 2016 年开始的，主要是与 2016 年美国总统大选和之后的传播有一定的关系。而谣言^[26]则是指被广泛传播的且并未得到官方证实的信息或者量已够被证实为不实的虚假信息。这几年，研究界对假新闻和谣言，已经有大量威胁国家或地区的安全。

互联网的普及与应用助长了假新闻和谣言的泛滥，一些个体或机构为了制造话题度，博取关注度，任意发布一些虚假新闻或谣言，完全不理睬其带来的不良影响。客观来讲，新闻是一种对客观事实再现的社会活动，为社会中每一个人的生活带来了深刻的影响，提供一个自己与外界连接的媒介，从中认识自己，认识社会，认识世界，人们可以非常方便地享受其带来的经济、政治、文化以及各种专业领域等信息服务。而假新闻的泛滥则破坏了新闻所具有的公信力和权威性。例如曾经流行的“塑料大米”^[3]引起消费者恐慌，“日本福岛核污染扩散”^[4]引发的加碘盐被疯狂抢购的事件，还有针对 2020 年新冠肺炎所报道的一系列假新闻，如 2020 年 1 月 30 日，塞尔维亚 pink 电视台、B92 网站等多家塞尔维亚媒体转载外媒关于新冠肺炎疫情的虚假视频报道，并以《恐怖市场 (pijaca strave)》等为标题吸引关注，舆论炒作，让其观众对中国造成很大的误解^[6]。再如外媒报道“武汉疑似火化大量新冠肺炎患者尸体导致二氧化硫破千”^[7]，经过中国监测总站的对比核实，外媒所用数据跟实际监测数据相差 200 倍，报道严重失真，一来影响公众对武汉的理解，二来对其气象服务网站能力也提出了很大的挑战。假新闻和谣言的所带来的影响力一次次突破大众的想法。为此，进行谣言检测的研究显得很有必要，从而能够更加有效和准确地检测社交媒体平台中的谣言，遏制假新闻在社交媒体上的传播，对于营造积极健康的网络环境和促进整个新闻生态系统的信任建设发挥着重要作用，而且，社交媒体平台可以从及时控制舆情和引导舆论中发挥积极作用。

1.2 研究现状

针对社会网络中的谣言泛滥问题，国内外研究学者进行了大量的研究和探索，对社会媒体的谣言进行自动检测也逐渐发展成为一个十分热门的研究领域。目前，大多数的谣言检测方法被看作是一种分类问题，可以对一些检测指标来进行谣言

5.一篇学位论文的绪论（引言），大致包含如下几个部分：

研究背景及研究意义；

国内外研究现状；

研究内容；

论文组织结构。

（1）研究背景中需要阐述清楚 2 个大问题：研究的问题“是什么”，为什么选择这个题目来研究，即阐述该研究的现实意义，比如说明该研究对学科发展有贡献，该研究对能够解决什么现实问题，该研究具有一定的前沿性等。

（2）国内外研究现状对本研究主题范围内的文献进行详尽的综合述评，“述”的同时一定要有“评”，指出现有研究成果的不足，讲出自己研究的动机。应简述本课题在国内外的研究和发展状况；针对课题的实际背景和要解决的问题，对比国内外学者的相关工作，阐述清楚国内外学者对同类问题所采用的研究和解决方法，同时对比这些研究和解决方法的优缺点。当然也可适当简要地介绍一些与本课题有关的预备知识。

检测,例如基于时间、空间、用户等特征以及它们的融合特征。现有的谣言检测方法主要有三种,分别是人工检测方法,基于机器学习的检测方法和基于深度学习的检测方法。以下则是对这三种谣言检测方法展开的具体介绍。

1.2.1 人工检测方法

人工谣言检测方法^[9],它是当前社会网络平台上主流的一种谣言检测方法,其工作机制是:平台将网络中的可疑信息交给具有丰富经验的编辑或相关行业专家,通过使用他们的经验和行业知识对可疑信息进行甄别,判断其真伪。新浪微博、Twitter、Facebook 等主流的社交网络平台都是采用人工谣言检测方法。新浪微博平台使用了两种对不实信息检测举措,一是“微博辟谣”¹平台,该平台旨在针对微博内可能出现的不实信息或重大争议内容,向网友传递更加全面、真实、客观的信息,其中,对于重大辟谣信息,微博辟谣官方会以粉丝头条的形式进行扩大传播。二是“微博举报处理中心”²,采用众包方法,微博用户向该平台举报可疑的信息,微博平台的专家们对举报的信息进行甄别,之后在通过平台发布结果。Twitter 同样采用众包方法,对平台上的信息的真伪进行甄别,通过设计一种信息真实性判别算法,这种算法能够根据 Twitter 上的用户信息评价来计算每一条信息的真实度。Facebook 则是采用人工标注与权威媒体证实相结合的一种方法来对其平台上的信息进行真伪性判别。Facebook 的用户可将 Facebook 平台上的可疑信息通过该平台的接口进行提交,之后,一些权威媒体将可疑的信息通过 API 提交给该媒体的编辑,然后,权威媒体和领域专家对上报的可疑信息的真伪进行甄别。由以上发现,Facebook 与新浪微博这两者的检测方法有些相似。人工检测方法准确率高,但是,通过以上分析,发现会存在以下问题:

- (1) 人工谣言检测方法将会耗费大量的人力来对上报的可疑信息进行逐一甄别,同时,还须依靠主观进行判断的滞后性。

1.3 研究内容与结构安排

1.3.1 研究内容

近年来,谣言检测研究已经取得了一些优秀成果,但大多数研究都侧重于有效检测出具有潜在特征的谣言,对于研究模型为什么可以检测新闻或信息为假的并没有给出确切的判定依据。可解释的谣言检测,就是解决这个问题,即让人们理解新闻或信息被认定为谣言的原因,提高模型的检测性能。在这个基础上,结合分析现有的谣言检测算法,考虑到循环生成网络 CycleGAN 可以实现两个相近数据集相互转换的特点,使用 CycleGAN 网络架构来开展此次研究,主要研究的内容包括以下几点:

- (1) 本文对谣言检测的背景以及研究的意义进行了相关和具体的介绍,对谣言检测的研究现状的也进行了具体的介绍与分析,在这个介绍与分析中,根据现有谣言检测缺乏可解释性等问题来阐明本次研究所要研究的内容和方向。
- (2) 本文对可解释机器学习谣言检测模型涉及到的相关理论进行了介绍,主要包括不平衡数据概述、机器学习的可解释性、对抗生成网络概述和循环生成网络概述。
- (3) 基于谣言数据集中谣言样本数量与非谣言样本数量不平衡问题,实际情况中,谣言数据的数量是远远少于非谣言数据的数量的。本次研究通过使用循环生成对抗网络的生成器网络生成谣言样本数据克服了数据类别分布不平衡的问题。
- (4) 基于本次研究提出的可解释机器学习的谣言检测模型,利用循环生成对抗网络将不同域数据相互转换的特点,实现谣言数据与非谣言数据的相互转换,在这个过程中学习谣言数据具有的关键特征,为模型的判定结果提供一个可靠的依据,即可解释性,同时提高模型的鉴别能力。
- (5) 对于本次研究进行了具体的实证分析,同时与常用的机器学习分类技术进行实验对比,通过常用的评估指标,证明了基于可解释机器学习的谣言检测模型的有效性。

注:国内研究现状与第二章的相关工作是有区别的。国内外研究现状描述课题研究的大背景大方向,侧重讲述课题研究的先进性和重要意义,不用涉及具体技术。

(3) 研究内容包括需要解决什么问题,解决该问题采用的理论依据、研究方法和实验基础,预期的结果及其地位、作用和意义。在研究内容的最后需写清楚本研究的创新点或理论与(或)实践意义。如果研究的项目是别人从未开展过的,这时创新性是显而易见的,要说明研究的创新点。但大部分情况下,研究的项目是前人开展过的,这时一定要说明此研究与被研究的不同之处和本质上的区别,而不是单纯的重复前人的工作。

1.3.2 结构安排

本文的正文部分主要由四个部分组成，一个章节代表一个部分，如下是对每个章节的具体内容进行的详细陈述：

第一章：绪论。绪论部分有研究背景及意义，研究现状和研究内容与结构安排。首先对谣言检测研究的背景及研究的意义进行了具体的描述与介绍，接着对谣言检测研究的现状进行论述和分析，最后就是对本文的研究内容和文章的结构安排进行具体详细的阐述。

第二章：相关理论概述。一个是对于谣言数据集类别不平衡的问题，介绍了不平衡数据的相关概述。一个是对机器学习的可解释性进行论述。其余两个是对生成

式对抗网络以及循环生成式对抗网络的相关内容介绍，为进一步的分析和研究奠定了理论基础。

第三章：基于可解释机器学习的谣言检测算法。首先描述了可解释机器学习的谣言检测研究涉及的相关问题。其次介绍了网络结构的设计与实现，包括模型的架构、生成器模型、判别器模型、损失函数以及谣言检测算法的实现流程等内容。最后论述了实验设计与分析，介绍了实验相关的谣言数据集，实验设置的内容以及对实验结果进行分析等。

第四章：总结与展望。这一部分就是根据本文的相关内容，总结了本文的进行的一些工作内容，并且对研究中的不足的地方进行了具体的分析，而且针对这些不足之处分别进行对未来的研究方向的一些展望。

(4) 论文的组织结构是对整篇论文的概述，阐述清楚论文的章节，每一章的研究内容或者介绍章和章间的关系。

2. 相关理论概述

2.1 不平衡数据概述

不平衡数据是指数据集中各个类别的样本量极不均衡,即某些类的数据量远远高于其他类。比如,二分类问题,正类的样本量远远大于负类的样本量,在这样的情况下,可能会为一个无意义的度量指标进行优化,最终取不到良好的效果,这对机器学习以及深度学习的结果造成极为严重的后果。

在谣言检测研究中,所收集到的谣言数据集中的谣言样本量远远小于非谣言样本量,是一个不平衡的数据集,如果不解决数据不平衡的问题,直接对其进行建立模型,那么由该模型得到的预测结果将会严重偏向非谣言而忽视谣言,导致该模型的分类准确度偏低。为了使本次谣言检测研究能够取得良好的效果,本次研究采用 CycleGAN 的生成网络生成谣言数据,来解决数据集不平衡的问题。

6. 论文的相关工作指的是与论文所研究内容直接相关的同行研究进展,在行文过程中需要有准确的参考文献支撑。一般来说,相关工作与第一章绪论篇幅加起来不能超过全文的 1/3。

针对研究的问题,目前他人已经做了哪些方面的研究,提出了哪些算法,简单总结其解决的问题与相关算法,并指出现有算法的不足之处。

注意这里的相关工作不仅仅包含现有工作的简单描述,还要在现有的研究工作中进行对比,指出其中的不同之处,并且说明现有算法的不足之处,本论文要在哪些方面做改进等。

理论基础指的是解决该问题使用的一些基本理论,但不是必需内容。如果论文中所涉及的相关理论,对后续论文的理解需要该知识,则有必要在这一章节进行简要介绍。否则,没有必要进行介绍。

这里的基本理论指的是与所研究问题直接相关的内容,而不应是该领域内所有知识的全部概括。

3. 基于可解释机器学习的谣言检测算法

3.1 问题描述

通常来讲, 谣言检测的主要任务在于对一些没有标记的信息的真假情况进行预测: 信息是真实的信息则预测为非谣言, 信息是虚假的信息则预测为谣言。这实质上也就属于一种二分类的任务, 此研究也是立足于这一点出发, 在此基础上进行了进一步的探索。也就是说, 此研究除了模型在已经被标记的谣言或非谣言的训练数据集上进行训练后, 根据训练好的模型能够预测出没有被标记的测试数据的真伪性之外, 还可以通过该模型学习谣言的关键特征, 以此来解释信息被预测为谣言的原因, 提高模型对谣言的鉴别能力, 得到的预测结果也更加具有可靠性和说服力。

首先, 训练数据集是一个有标签的数据集合: $S = \{(x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_n, y_n)\}$, 其中 n 是表示数据集的大小, 每条记录为 (x_i, y_i) ($i=1, 2, \dots, n$), 其中 x_i 表示第 i 条数据, 是一个文本向量, $y_i \in \{0, 1\}$ 表示 x_i 的标签, 标签值分为“谣言”和“非谣言”, 其中, 用 1 表示谣言, 0 表示非谣言。模型的输入数据是待判断的文本向量, 输出是对应输入数据的类别标签, 即“谣言”或“非谣言”。其次, 模型把一个谣言数据转换为非谣言数据, 再把非谣言数据转换为谣言数据, 在这个过程中, 模型会从中学习到谣言数据所具有的关键特征因素, 成为判断信息成为谣言的重要依据, 达到可解释性的效果。

3.2 网络结构的设计与实现

3.2.1 模型的整体架构

针对现有谣言检测算法的不足和 CycleGAN 的特点, 本文提出一个基于可解释机器学习的谣言检测算法, 使用 CycleGAN 框架来搭建模型, 输入数据输入模型中, 能够判断该数据是谣言还是非谣言, 除此之外, 谣言数据输入到模型中, 能够得到非谣言数据的输出。本模型设计了两个生成器和一个判别器, 不同于上述 2.4 节中描述的 CycleGAN, 研究所设计的模型中的判别器的作用在于: 除了可以判断输入数据是真实的数据还是生成器生成的数据之外, 还有一个非常重要的作用, 就是可以判断输入数据是谣言数据还是非谣言数据。对于判别器而言, 假如其输出结果为 1, 则表示输入数据是真实的数据; 假如输出结果为 0, 则表示输入数据就是生成器生成的数据; 类似的, 假如判别器的输出结果为 1, 代表输入数据是谣言数据, 假如判别器的输出结果为 0, 则代表输入数据是非谣言数据。总的来说, 判别器有

7.算法(模型)设计这一章是论文的核心内容, 包括所研究问题的描述与定义, 解决该问题采用的算法, 算法的详细描述, 以及算法的复杂度分析等。可以包括问题描述和算法(模型)的设计两部分。

4. 实验评测与结果分析

4.1 实验设置

4.1.1 实验环境

本文进行实验时使用的编程语言为python 3.6.9，使用的深度学习框架为Tensorflow 1.13.1，将代码与数据集存储于 Google Drive，通过实验平台 Google Colab 挂载 Google Drive 使用 GPU 运行程序，实验平台相关信息如表 4-1 所示：

表 4-1 实验环境

显卡	Tesla P100
驱动	NVIDIA-SMI 418.67
CUDA 版本	10.1
显存	16GB
磁盘空间	68.40GB

虽然 Colab 平台提供的算力可观，但是在运行程序的过程中，系统会根据资源

4.1.2 实验数据集

本实验采用 NLPCC – ICCPOL 2016 KBQA 任务发布的数据集，该数据集包含三个文件，数据规模如 4-2 所示：

表 4-2 实验数据集信息统计

文件名称	数据量
训练集	问答对 14,609 个
测试集	问答对 9,870 个
知识库	实体 6,502,738 个，属性 587,875 个，三元组 43,063,796 个

训练集与测试集原数据只包括问答对，本文使用的数据为该任务评测第一名预处理后的数据^[17]，增加对于每个问题标注的三元组，三元组的构造方法为从知识库中反向查找答案，根据问题过滤实体，最后进行筛选数据样例如表 4-3 所示：

4.1.3 评测指标

NLPCC-ICCPOL 2016 KBQA 任务官方评测最终结果采用的评价指标为平均 F1 值，此处将准确率 accuracy 值看作 F1 值，可得到：

8.实验部分建议包括实验环境、实验数据的介绍、模型评价指标的描述和实验结果的分析。

实验结果分析中

(1) 包括给出结果，并对结果进行定量或定性的分析。写作要点是：以绘图和（或）列表（必要时）等手段整理实验结果。

(2) 除了给出实验结果，另一个重点内容是将论文中提出的算法与已经存在的算法做对比，说明在哪些方面有所改进，并说明性能提高的原因。

4.2 实验与分析

4.2.1 数据预处理

首先，本文从训练集与测试集中的问答对文本数据中，提取出了每个问答对所对应的问题、三元组、答案，存储到结构化表格中，便于实体链接、属性映射任务的训练集构建与系统测试等操作。数据样例如表 4-5 所示：

表 4-5 表格内存储的数据样例

q_str	t_str	a_str
<question id=1> 《机械设计基础》这本书的作者是 谁？	<triple id=1> 机械设计基础 作者 杨可桢，程光蕴，李仲生	<answer id=1> 杨可桢，程光蕴，李仲生

数据集划分：从训练集中划分出 2609 个问答对作为验证集，得到包含 12000 个问答对的训练集，包含 2609 个问答对的验证集合包含 9870 个问答对的测试集。

将知识库中的三元组存入云数据库中，如图 4-6 所示：

entity	attribute	answer
机械设计基础	作者	杨可桢，程光蕴，李仲生
高等数学	出版社	武汉大学出版社
线性代数	出版时间	2013-12-30
安德烈	国籍	摩纳哥
线性代数	isbn	978-7-111-36843-4
高等数学	书名	高等数学一（微积分）
万达广场	外文名	amoywandaplaza
李明	出生日期	1963.1
霸陵守护力	小说进度	连载

4.2.2 主题实体识别

(1) 数据集构建

根据每个问题对应三元组的实体，对于问题序列使用 BIO 标签进行反向标注，其中，本文命名实体识别的主要任务是识别出问题的实体提及，因此对于实体的具体类别不做区分，均以“B-LOC”、“I-LOC”的形式进行标注。将每个问句以空行分隔，标注结果如图 4-7 所示：

```
《 O
机 B-LOC
械 I-LOC
设 I-LOC
计 I-LOC
基 I-LOC
础 I-LOC
》 O
这 O
本 O
书 O
的 O
作 O
者 O
是 O
谁 O
? O
```

图 4-7 反向标注得到的训练数据样例

(2) 关键参数（相关代码详见附录）

表 4-9 命名实体识别任务相关参数设置

参数	参数值
max_seq_length	128
batch_size	128
learning_rate	0.0001
droupout_rate	0.6
Gradient clip	5
hidden_size (bert base)	768
num_hidden_layers (bert base)	12
embedding_size (bert)	128
hidden_unit (lstm_size, 使用了 bert-bilstm-crf 对比实验)	128
number of layers (lstm)	1

此外，需要将数据在输入前处理为图 3-5 的形式，如表 4-10 所示：

表 4-10 模型输入样例

INFO:tensorflow:*** Example ***
INFO:tensorflow:tokens: 《 机 械 设 计 基 础 》 这 本 书 的 作 者 是 谁 ？
INFO:tensorflow:input_ids: 101 517 3322 3462 6302 6360 1825 4704 518 6821 3315

(3) 结果分析

实体识别任务结果示例如下，其中，predict 属性对应的值为实体识别得到的结果，可见，大部分问题的实体提及都能够被准确识别。

表 4-11 实体识别结果示例

question	entity	predict
你知道计算机应用基础这本书的作者是谁吗？	计算机应用基础	计算机应用基础
计算机应用基础这本书的出版社是哪个？	计算机应用基础	计算机应用基础
告诉我高等数学的出版时间是什么时候？	高等数学	高等数学
我想知道戴维斯是什么国家的人？	戴维斯	戴维斯

此外，本文使用 BERT-BiLSTM-CRF、BERT-CRF 等方法进行了对比实验，测试结果如表 4-12 所示：

表 4-12 命名实体识别任务实验结果对比

模型方法	准确率 (%)	召回率 (%)	F1 值 (%)
CRF	86.22	86.49	86.35
BiLSTM	89.13	88.25	88.69
BiLSTM-CRF	90.87	89.42	90.13
BERT-BiLSTM-CRF	94.92	95.93	95.42
BERT-CRF	95.74	96.21	95.97

可见，预训练模型的应用大大提升了命名实体识别的效果，其中，将 BERT 与 BiLSTM 叠加反而降低了模型效果，分析原因是，由于 BERT 模型内部已存在双向编码，transformer 已足够强大，在模型中加入 BiLSTM 没有必要性，直接将 BERT

5. 总结与展望

5.1 总结

本文主要研究基于中文知识图谱的问答系统，对于传统的语义解析方法使用深度学习进行了改进，解决了传统语义解析存在的句法分析需要大量语言学先验知识、人工模板构建复杂且模型泛化能力差等问题。

针对单关系型问答任务，本文构建了包含主题实体识别、属性映射、答案选择环节的系统问答流程，提出了 BB-KBQA 模型。该模型由命名实体识别模型 BERT-CRF 和属性映射模型 BERT-Softmax 构成，采用 BERT 类预训练语言模型，解决了 Word2Vec 难以解决的一词多义等问题，获得了问题的深层语义表示。实验证明，BERT-CRF 融合了 BiLSTM、CRF 等模型的优点，提升了命名实体识别的准确率；BERT-Softmax 能够实现高达 98.43% 的属性映射准确率，使得系统能够更好地实现问题与知识库的语义匹配。

最后，本文将 BB-KBQA 模型应用于 KBQA 系统，实现了基于中文知识图谱问答的完整流程，在 NLPCC-ICCPOL 2016 KBQA 任务上进行评测，回答通用领域的单关系型事实问题，达到了 81.43% 的平均 F1 值，接近该评测任务的最好水平。

5.2 局限与展望

本文仍存在许多不足，同时也是未来研究的努力方向：

(1) 问题提及实体可能存在歧义：一些问题所提及的实体存在错误或者歧义，导致答案错误。例如“博士来拜是什么年代的作品？”问题中所提及的实体为“博士来拜”，在知识库中可查询到“博士来拜（维登画作）”、“博士来拜（曼特尼亚画作）”等候选实体，且均具有“年代”属性，导致虽然命名实体识别和属性抽取均

9. 总结与展望包括结论部分和展望部分。结论不是研究结果的简单重复，而是对研究结果更深入一步的认识，其内容要点如下：

(1) 本研究结果说明了什么问题，得出了什么规律性的东西，解决了什么理论或实际问题；

(2) 对前人有关本问题的看法作了哪些检验，哪些与本研究结果一致，哪些不一致，作者做了哪些修正、补充、发展或否定；

(3) 展望部分阐述的不足之处或遗留问题。

参考文献

- [1] 崔万云. 博士学位论文: 基于知识图谱的问答系统关键技术研究[D]. 复旦大学. 2017.
- [2] 刘康. 基于深度学习的知识库问答[R]. 知识图谱与问答系统前沿技术研讨会. 2016.
- [3] Kang L, Yuanzhe Z, Guolaing J, et al. 基于表示学习的知识库问答研究进展与展望[J]. 2016.
- [4] 周博通, 孙承杰, 林磊, 等. 基于 LSTM 的大规模知识库自动问答[J]. 北京大学学报: 自然科学版, 2018, 54(2): 286-292.
- [5] 曹明宇, 李帅驰, 王鑫磊, 等. DUTIR 中文开放域知识库问答评测报告[R]. CCKS. 2019.
- [6] qrlhl. 词法分析之 Bi-LSTM-CRF 框架.
<https://blog.csdn.net/qrlhl/article/details/78561342>. 2017.
- [7] 周正平, 张钰晖. 基于知识库的问答系统实现[R]. 清华大学. 2017.
- [8] Liang P. Lambda dependency-based compositional semantics[J]. arXiv preprint arXiv:1309.4408, 2013.
- [9] Berant J, Liang P. Semantic parsing via paraphrasing[C]//Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2014: 1415-1425.
- [10] Yao X, Van Durme B. Information extraction over structured data: Question answering with freebase[C]//Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2014: 956-966.

10. 参考文献:

(1) 规范性: 严格按照学校学位论文要求编排参考文献以及每条参考文献的样式。

(2) 对应性: 凡是前文中引用了别人成果(文字、数字、事实以及转述他人的观点)之处, 均应加标注列于参考文献中。

(3) 前沿性: 参考文献需能看出论文较好地跟踪了近期的该领域研究动态。