

2019 届硕士专业学位论文

分类号: \_\_\_\_\_

学校代码: 10269

密 级: \_\_\_\_\_

学 号: 63154501405



華東師範大學

East China Normal University

硕士专业学位论文

MASTER' S DISSERTATION

# 论文题目：基于挖掘算法的学生成绩分 析与预警系统设计与实现

院 系: 软件工程学院

专 业: 工程硕士

研 究 方 向: 软件工程

指 导 教 师: 刘小平 讲师

学位申请人: 吴烨

2019 年 11 月 16 日

Dissertation for master degree in 2019

Student ID: 63154501405

University code:10269

# East China Normal University

**Title: Design and implementation of student achievement  
analysis and early warning system based on  
mining algorithm**

Department:	<u>Software Engineering Institute</u>
Major:	<u>Master of Engineering</u>
Research direction:	<u>Software Engineering</u>
Supervisor:	<u>Liu Xiaoping Lecturer</u>
Candidate:	<u>Wu Ye</u>

November , 2019

## 华东师范大学学位论文原创性声明

郑重声明：本人呈交的学位论文《基于挖掘算法的学生成绩分析与预警系统的设计与实现》，是在华东师范大学攻读硕士/博士（请勾选）学位期间，在导师的指导下进行的研究工作及取得的研究成果。除文中已经注明引用的内容外，本论文不包含其他个人已经发表或撰写过的研究成果。对本文的研究做出重要贡献的个人和集体，均已在文中作了明确说明并表示谢意。

作者签名： 吴辉

日期： 2019 年 11 月 16 日

## 华东师范大学学位论文著作权使用声明

《基于挖掘算法的学生成绩分析与预警系统的设计与实现》系本人在华东师范大学攻读学位期间在导师指导下完成的硕士/博士（请勾选）学位论文，本论文的研究成果归华东师范大学所有。本人同意华东师范大学根据相关规定保留和使用此学位论文，并向主管部门和相关机构如国家图书馆、中信所和“知网”送交学位论文的印刷版和电子版；允许学位论文进入华东师范大学图书馆及数据库被查阅、借阅；同意学校将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于（请勾选）

（ ） 1. 经华东师范大学相关部门审查核定的“内部”或“涉密”学位论文\*，于 年 月 日解密，解密后适用上述授权。

（☒） 2. 不保密，适用上述授权。

导师签名 刘坤

本人签名 吴辉

2019 年 11 月 16 日

\* “涉密”学位论文应是已经华东师范大学学位评定委员会办公室或保密委员会审定过的学位论文（需附获批的《华东师范大学研究生申请学位论文“涉密”审批表》方为有效），未经上述部门审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权）。

吴烨 硕士学位论文答辩委员会成员名单

姓名	职称	单位	备注
安平	教授	上海大学	主席
张民	副教授	华东师范大学	
李成举	副教授	华东师范大学	

## 摘要

针对传统学生成绩分析管理的不足，如分析效率不高、准确性差、成绩无法预测等问题，本文设计和实现适合普通高中的学生成绩分析与预警系统。学校可以有效地提升学情管理和成绩分析能力，从而提高工作效率。因此，学生成绩分析与预警系统的开发与设计具有重要的价值和意义。

本文利用 UML 统一建模技术，解决了用户业务流程分析、用例分析、非功能性需求分析等问题。同时，采取工作流技术，解决了传统成绩管理中记载、汇总和统计出现的效果较差、效率较低等问题。利用数据挖掘技术分析处理数据的优势，实现预测分析，满足用户成绩分析与预测等需求。在功能业务设计上，本文通过 UML 工具对主要业务处理学籍管理、考试分析、成绩管理、预测分析等建模设计，实现成绩预警机制去激励学生学习积极性，最后是学生成绩分析与预警系统的实现与测试，按照前面的设计搭建开发环境，并在这个基础上实现模块界面和关键代码的实现。系统测试实验结果表明系统的功能基本符合要求，测试数据与结果是达标的。

学生成绩分析与预警系统的研发采取信息化的手段，在提升成绩分析管理水平的同时，也促进成绩管理的科学化，为提高中学学生的升学率和改进教学管理提供数据支撑。

**【关键词】** 成绩预警；考试分析；决策树算法；预测分析

**【论文类型】** 应用基础

## **Abstract**

In view of the shortcomings of traditional student achievement analysis management, such as inefficiency, poor accuracy and unpredictability, this paper designs and implements a performance analysis and early warning system suitable for senior middle school students. It can effectively enhance the ability of student status management and achievement analysis, thereby improving work efficiency in schools. Therefore, the development and design of student performance analysis and early warning system is of great value and significance.

Firstly, UML unified modeling technology is used to solve the problems of user business process analysis, case analysis and non-functional requirement analysis. Workflow technology is adopted to solve the problems of poor effect and low efficiency of record, summary and statistics in traditional performance management. The algorithm in data mining has good analysis processing number. According to its advantages, it can better realize the prediction analysis and meet the needs of user performance analysis and prediction. In the functional business design, through UML tools for the main business processing student status management, examination analysis, performance management, prediction analysis and other modeling design, through the performance early warning mechanism to stimulate students' learning enthusiasm. According to the system's data requirement modeling and processing, the data object entity analysis and modeling, get the data table structure. Finally, the implementation and testing of student performance analysis and early warning system, according to the previous design to build a development environment, on this basis, the realization of module interface and key code. After completing the system development, it needs to be tested. The results show that the functions of the system basically meet the requirements, and the test data and results are up to the standard.

The research and development of student achievement analysis and early

warning system adopts the means of informationization, which not only improves the level of performance analysis and management, but also promotes the scientific management of achievement, and provides data support for improving the rate of high school students' entrance and teaching management.

**【Key Words】** Performance early warning; Examination analysis; Decision tree algorithm; Predictive analysis.

**【Thesis Type】** Application of basic

# 目录

摘要 .....	I
ABSTRACT .....	II
目录 .....	IV
第 1 章 绪论 .....	1
1.1 研究背景与意义 .....	1
1.2 国内外应用现状 .....	2
1.2.1 国外现状 .....	2
1.2.2 国内现状 .....	3
1.3 主要研究内容 .....	4
1.4 论文组织结构 .....	5
第 2 章 相关技术简介 .....	6
2.1 软件开发技术 .....	6
2.1.1 JavaEE 介绍 .....	6
2.1.2 B/S 架构 .....	7
2.1.3 数据库技术 .....	8
2.2 数据挖掘技术 .....	9
2.2.1 数据挖掘算法 .....	9
2.2.2 不同算法比较 .....	10
2.3 本章小结 .....	11
第 3 章 学生成绩分析与预警系统需求分析 .....	12
3.1 系统概述 .....	12
3.2 系统功能需求分析 .....	13
3.2.1 系统管理 .....	13
3.2.2 学籍管理 .....	14
3.2.3 成绩管理 .....	15
3.2.4 考试分析 .....	16
3.2.5 预测分析 .....	17
3.3 系统性能需求分析 .....	18
3.4 本章小结 .....	18
第 4 章 学生成绩分析与预警系统的详细设计 .....	19
4.1 系统总体设计 .....	19
4.1.1 网络结构设计 .....	19
4.1.2 体系结构设计 .....	20
4.1.3 功能架构划分 .....	21
4.2 功能模块设计 .....	21
4.2.1 系统管理模块设计 .....	21
4.2.2 学籍管理模块设计 .....	23
4.2.3 成绩管理模块设计 .....	25
4.2.4 考试分析模块设计 .....	27



4.2.5 预测分析模块设计 .....	29
4.3 数据库设计 .....	31
4.3.1 概念结构设计 .....	32
4.3.2 逻辑结构设计 .....	32
4.4 本章小结 .....	36
<b>第 5 章 学生成绩分析与预警系统主要功能实现 .....</b>	<b>37</b>
5.1 开发环境搭建 .....	37
5.2 系统功能模块的实现 .....	39
5.2.1 系统管理模块实现 .....	39
5.2.2 学籍管理模块实现 .....	41
5.2.2 成绩管理模块实现 .....	43
5.2.4 考试分析模块实现 .....	44
5.2.5 预测分析模块实现 .....	46
5.3 数据访问的实现 .....	47
5.3.1 数据库连接 .....	47
5.3.2 访问性能优化 .....	49
5.4 本章小结 .....	51
<b>第 6 章 学生成绩分析与预警系统的测试 .....</b>	<b>52</b>
6.1 测试环境 .....	52
6.2 功能测试 .....	52
6.3 性能测试 .....	56
6.4 本章小结 .....	58
<b>第 7 章 总结与展望 .....</b>	<b>59</b>
7.1 总结 .....	59
7.2 展望 .....	59
<b>参考文献 .....</b>	<b>61</b>
<b>致谢 .....</b>	<b>65</b>

## 第1章 绪论

### 1.1 研究背景与意义

随着信息技术发展,信息技术越来越融入到教育教学活动中。先进的信息技术可以为教学活动和教学改革服务,解决教学管理中的实际问题<sup>[1]</sup>。对于学校来说,日常教学活动中会产生相当多的基础数据,如学生成绩、学生作业、学生考勤等数据。随着大数据技术的发展,我们可以利用大数据算法对这些数据进行前期的基础采集、统计分析、分类处理等,有效的分析学生的学习生活等行为<sup>[2]</sup>。利用数据挖掘技术,将学生在校期间的表现通过数据分析的方式展现出来,从多个角度展现学生的学习活动,我们甚至可以为以后的成长发展提供预测分析服务,也可以为教育教学的管理者提供参考<sup>[3]</sup>。对于学生来说,在学校期间,都会产生海量具有意义价值的数据,如果这些数据能够得到有效的收集、处理和分析,它将可以为学校的相关人员提供良好的指导。通过分析学生的学习数据,可以对在校期间表现不好的科目或成绩进行预警<sup>[4]</sup>。因此,学生的管理出现在前端,我们可以有效的防止学生出现辍学等现象。

针对学生在校期间的各类表现进行分析处理,有利于对学生的教育方式的转变,改变传统依靠人工管理的方式。将大数据的应用价值延伸到学生管理教育领域,对现有的教育进行深入分析和挖掘,实现教育领域数据化管理,为用户提供更加良好的服务<sup>[5]</sup>。目前,在教育领域,特别是在普通高中,学生在校期间积累了大量的具有结构化的数据。对这些结构化的数据,如学生信息、学生成绩等进行分析;与此同时,对一些非结构化的数据,如学生的考勤数据等进行充分的分析,就可以有效的帮助我们建立学生在校期间的准确画像。通过建模,我们就可以对学生进行学业预警和学业指导<sup>[6]</sup>。数据挖掘技术可以帮助我们对数据进行准确采集和处理,将所需要的数据进行标准化和规范化的转换,最终得出我们所需要的数据。

综上所述,通过对教育领域特别是学生成绩数据的采集、处理、分析,可以对在校期间学习表现进行全方位的分析,形成具有学生个性的学习表现发展趋势,为存在学业风险的学生提供预测服务。对于在教师来说,通过学生学业的分析,可以改善教学方式手段,提高教学质量。对于学校的管理者来说,通过系统改变管理模式,提升教学管理质量和水平,具有重要的价值和意义。

## 1.2 国内外应用现状

### 1.2.1 国外现状

在国外,学者们对数据挖掘技术的研究和运用开始得比较早,学生考试成绩的数据挖掘得到极大的应用和推广<sup>[7]</sup>。斯坦福大学,麻省理工大学以及卡内基梅隆大学等都对教育领域的数据挖掘技术的分析和应用进行较为深入的研究,取得很好的研究成果<sup>[8]</sup>。目前,数据挖掘技术的应用呈现出不断递增的趋势。在2012年,美国就对全国范围内的教育数据包括学生考试成绩等数据进行有计划的开发,将现今的数据挖掘分析技术应用其中<sup>[9]</sup>。通过数据挖掘教育教学数据,建立学生个性化的学习模型,在一定的训练和学习数据下,完成个性化学习的模式框架创建<sup>[10]</sup>。通过这个模式,学校有效地帮助教师来改变教学方法,提升教育教学效果。有了学生全方位的教育数据,教师就可以通过系统发现学生学习的缺陷,提供更好的学习的建议。

国外的学者对学生成绩预测的模型和算法开展相关研究取得了一定的成果。例如,德国学者汉斯教授在2015年通过研究提出数据挖掘技术、机器学习在教育行为具有良好的应用前景和价值的观点<sup>[11]</sup>。2016年美国学者吉安慕斯等人通过研究,提出在学生学习的冬季和提高学生成绩之间的规律数据集分析,采取粗糙集理论和方法具有很好的准确率的论断<sup>[12]</sup>。日本学者井村一等人对某学习平台的日志分析,对学习的行为进行数据的采集分析,构建出具有预测功能的模型<sup>[13]</sup>。2018年,法国学者泽曼等人对在巴黎大学的本科生和研究生中的成绩数据进行多种手段的采集分析。对其中的成绩在分析基础上进行预测,预测学生在研究生阶段的学习表现,建立了成绩预测的模型,取得良好的应用效果<sup>[14]</sup>。

很多大型的商业公司,也在积极投入到教育大数据的研究应用。例如美国的IBM公司,通过机器学习技术,对学生的在校期间的数据进行采集处理。利用公司强大的技术研发实力,提出了针对学生学习的“Civitas Learning”项目,对学生成绩进行研究和做出预测分析<sup>[15]</sup>。对海量的学生在校期间各类有价值的数据进行重新的梳理和分析,提炼出影响学生学习的重要因素。对学生在校期间成绩起到较大影响的原因进行系统的分析处理,包括了原始的百万的学生数据、记录信息进行分析,通过预测分析,对可能存在重修风险的学生挖掘和发出预警,促进学生在此期间调整学习方式<sup>[16]</sup>。不仅仅是美国,加拿大的软件公司Desire to Learn也对教育数据积极研究。他们主要是针对在大学学习期间的学生课程历史成绩进行分析,并且将其延伸到课堂外,对学生的社会活动等非结构化数据采集处理,从而建立学生课程预测模型,预测学生学习成绩情况<sup>[17]</sup>。在2018年的教

育数据分析国际会议上,很多学者对教育挖掘进行深入的探讨,并且在大数据的浪潮下,逐渐形成教育数据挖掘的新兴交叉学科,将其应用在教育数据的分析研究中<sup>[18]</sup>。

综上所述,通过对国外研究现状的分析,可以看到教育数据是重要的数据资产。通过对其采集分析利用可以很好的分析学生成绩的问题所在,可以为用户提供更为精准的数据服务,以此不断地优化学校的教学活动,为教育教学实践提供良好的依据和支撑。

### 1.2.2 国内现状

国内对学生成绩管理分析系统的研究和应用相对国外来说起步得比较晚。最早对系统的研究和建设是上个世纪九十年代,后续对系统的研发逐渐成熟,在多个学校得到应用,取得较好的价值<sup>[19]</sup>。总的来说,学生成绩管理分析系统在国内的研发可以划分为二个阶段,具体如下:

第一阶段,从上个世纪的九十年代到本世纪初。在这个阶段中,随着互联网技术迅速发展,很多学校逐渐使用网络技术,建立起属于自己的网络。在这个良好的网络环境下,对传统初级 EXCEL 软件方式进行有效改造,建立起了电子化的系统<sup>[20]</sup>。对传统成绩管理中的数据和资料进行电子化处理,有一定的数据汇总和分析的功能。但是在成绩管理中,这些系统仅仅是用于考试成绩的记录,没有对数据做出任何分析<sup>[21]</sup>。究其原因,主要是缺乏技术与方法去收集更详细的信息,也没有相关的系统可以做出分析。

第二阶段,从本世纪初到至今。这个阶段是自动化信息化新技术阶段,主要是对传统的自动化系统进行改造,引入新技术。数据挖掘技术也加入到成绩管理中,使得新技术得到广泛应用,促进成绩分析和预测工作<sup>[22]</sup>。与此同时,多媒体的积极引入,便于获取学生个体更加细微的表现数据。与此同时,很多学校在不断改变,不断应用新技术和新方案,对原来的系统进行积极改造,加入一些移动互联网的元素,使得业务场景可以得到很多的延伸,数据贯穿于学生的整个学习和生活<sup>[23]</sup>。

我国的学者在成绩数据分析的算法研究中,也取得一定的成果。例如,魏顺平等介绍了数据挖掘技术在教育方面的发展历程和相关的概念,五类数据挖掘方法被列举其中<sup>[24]</sup>。叶福兰等人利用数据挖掘技术分析已有的高校教育教学数据,探究不及格课程的内在联系和关键因素,从而得出了学生成绩的预警信息<sup>[25]</sup>。刘斌等人则是利用数据挖掘中的 K-Means 算法,对目标数据进行聚类分析,最终得出优化后的学生评价方式。学生成绩预测预警研究方面,有的学者通过研究提出了高考回归预测模型。针对教学环节和相关数据的统计计算,建立回归计算

方程,处理预测区间中的误差处理后,促进教师可以在日常的教学活动中对内容作出调整,提升了教学水平<sup>[26]</sup>。在普通高中,我们可以针对高三的月考成绩与高考成绩的研究,建立相关性,为高考的复习和决策提供依据和支撑。目前大部分的研究采取的方式是机器学习中的线性回归,对成绩等教育数据进行分析和预测<sup>[25--26]</sup>。但是,在预测的准确性方面还有待于提高,需要提升预测的能力水平。如果引入神经网络、决策树、支持向量机、贝叶斯学习等相关算法,应用到教育课程数据中,实际上更加有利于计算,预测的效果更好<sup>[47-48, 52, 58]</sup>。在这方面,已经有学者已经对高考中的录取业务活动进行预测,采取的方式是贝叶斯网络,经过实际校验,已经取得较好的应用效果<sup>[58]</sup>。

综上所述,随着数据挖掘技术的不断发展和教育数据的不断完善,对在校期间学生课程成绩的采集分析和处理,建立学生学业成绩的预测预警模型,目前已经有一定的研究成果。本文希望通过利用数据挖掘算法设计,实现学生成绩分析与预警系统,改变落后的教育教学方式,从而提升教学质量。

### 1.3 主要研究内容

随着信息技术迅速发展,大数据、人工智能技术在教育领域积极应用。人们可以通过对教育数据的分析利用,提升教育教学质量,帮助用户改进教学方法,因此,针对教育数据中的课程成绩数据,本文采取 Java EE 软件开发技术、成绩数据挖掘技术等对学生成绩进行系统分析和设计,完成学生成绩分析与预警系统。希望本系统的开发与应用能够对学生成绩进行分析预测,为学校管理者提供数据支撑。论文主要研究内容为:

(1) 针对面向学生成绩分析的数据挖掘等流程和特点,进行系统业务的调研,取得系统总体的需求分析,为研发设计提供基础,从而建立起学生成绩分析与预警系统的业务框架和需求模型。

(2) 学生成绩分析与预警系统功能模块化设计和成绩挖掘模型构建。利用模块化设计技术,本文的总体设计为成绩管理、学籍管理、考试分析、成绩预测等业务功能模块,针对挖掘过程、算法和应用,设计了学生成绩分析数据挖掘模型。在模型中对挖掘算法及其应用场景进行了针对性的设计和优化,尤其是决策树算法在成绩预测具体应用,有效提升成绩分析工作效能。

(3) 学生成绩分析与预警系统实现和测试。针对学生成绩预测预警模型,对系统分层设计,采用 Java EE 技术有效降低层级的耦合性,使得系统具有良好的可扩展性,能够适应不同业务场景需求,有效提升系统的业务扩展性能。在开发阶段,利用 Struts、Spring 和 Hibernate 等开发框架,设计系统开发接口,

支持业务拓展的开发设计。根据前面的设计，对主要业务模块代码的实现和界面开发，完成系统研发。在测试阶段，搭建相应测试环境，设计测试用例，完成系统的功能测试和性能测试，验证系统是否符合开发的要求。

## 1.4 论文组织结构

论文的组织结构根据安排，整个论文划分为七个章节，具体如下：

第1章 绪论 介绍研究的背景和意义。对国外和国内目前学生成绩分析与预警系统研发情况做对比介绍，分析现有系统的不足。此外，还介绍了论文主要研究内容和论文的组织结构的安排。

第2章 相关技术分析 针对系统开发所要用的一些技术，如 Java EE 软件开发规范、B/S 架构模式、数据库存储技术等，在成绩数据分析中主要对一些常见的挖掘算法进行比较，分析其优点和缺点。

第3章 学生成绩分析与预警系统实现需求分析 介绍学生成绩分析与预警系统开发可行性，主要从技术可靠性等三个方面介绍系统研发是可行和必要的。接着是系统功能需求分析，通过 UML 工具对获取到的成绩管理、学籍管理、考试分析、成绩预测等四个方面业务进行描述。最后是系统性能需求分析，调研用户对系统运行在指标、速度、稳定性等方面的需求。

第4章 学生成绩分析与预警系统实现的设计 首先是制定系统总体的设计方案，包括了系统运行的网络环境、系统划分的层次、系统各个模块。其次是业务功能模块的详细设计，根据前面划分的业务功能进行需求分析，针对成绩管理、学籍管理、考试分析、成绩预测等四个方面业务采取 UML 工具详细设计，为后面的开发打下基础。接着是工作流模型设计，对于智能网络办公系统来说，最重要的是工作流引擎，完成对模型设计、流传处理设计等。最后是数据库设计，包括了概念结构设计和逻辑结构设计。

第5章 学生成绩分析与预警系统实现的实现 搭建相应的开发环境，对前面所确定的技术路线和选型进行各项核心功能相应的代码与页面开发，从而得到系统的功能。

第6章 学生成绩分析与预警系统实现的测试 主要阐述了本系统的业务功能测试和性能测试。利用测试的结果和数据，对系统进行改进和优化，验证系统是符合要求的，是达到目的。

第7章 总结和展望 对前面工作进行工作的小结，主要是对研发成果和内容进行总结，针对系统在本次研发过程中不足的地方进行展望，并且结合新技术和新业务需求部署下一阶段的工作。

## 第2章 相关技术简介

本章主要是进行技术分析,对学生成绩分析与预警系统开发设计所需要的JavaEE 软件开发技术、数据存储管理技术、B/S 架构模式进行比较和分析,最终确定系统开发的技术路线和选型。对学生课程成绩数据采集分析处理需要用到的数据挖掘技术、挖掘工具、挖掘算法等进行分析比较,确定学生成绩分析与预警系统开发需要的相关技术。

### 2.1 软件开发技术

#### 2.1.1JavaEE 介绍

学生成绩分析与预警系统在经过开发后,用户主要采取浏览器的方式访问系统。在这个模式下,对于用户来说,不需要安装部署客户端,可以带来更加方便快捷的用户体验。同时,学生成绩分析与预警系统中的数据都是在后台完成的,是典型三层架构,可以有效地提升用户访问的速度和运行效率,并且在后期的维护会变得非常便捷<sup>[30]</sup>。

在基于JavaEE 规范进行软件开发技术中,需要遵循它的规范标准,利用其所提供的相关服务框架组件,在不同的平台具有较好的兼容性,能够在后台运行的时候对数据对象进行处理<sup>[31]</sup>。JavaEE 软件开发技术与标准的Java 类有一些不同的区别。对于某一个应用来说,对服务器的管理,主要是通过提供组件来完成,包括了Java Servle、JSP、EJB 等组件在服务端的运行,完成对业务处理<sup>[32]</sup>。通过上述方式可以有效地提升运行处理的效率,可以很好地完成目标系统的开发,使复杂的软件系统的开发变得更为简单和便利。

通过JavaEE 所提供的软件开发方案,包括服务框架与组件能够实现在多种模式下的开发编码工作,具有良好的中间件开发功能<sup>[33]</sup>。在开发方案中,目前多个版本的系统方案能够实现较为灵活、开放和兼容的需求,实现目标系统的开发工作。通过基于JavaEE 规范所开发的软件系统具有如下的特点和优势<sup>[34]</sup>:

(1) 基于JavaEE 的开发具有较好的层次。对不同的层次进行自定义,系统的逻辑架构较为清晰。在不同的分层中,可以设计不同的业务模块以及支撑业务模块的各类插件。在完成开发设计之后,对模块和组件进行设置,系统就可以快速部署。

(2) 基于JavaEE 的进行软件开发具有较好的兼容性和规范性,能够适应不同的服务框架要求,可以在不同的平台中运行<sup>[36]</sup>。

(3) 基于 JavaEE 的进行软件开发可以运行在不同的用户终端上，无论是基于 PC 的用户还是基于移动端的用户都可以登录、访问、使用系统。

综上所述，基于 JavaEE 的进行软件开发遵循 Java 的规范，具有开发设计的优势。对于本系统的研发来说，基于 JavaEE 技术能够完成开发工作。因此，系统开发平台选用 JavaEE 技术路线。JavaEE 如图 2-1 所示<sup>[37]</sup>。

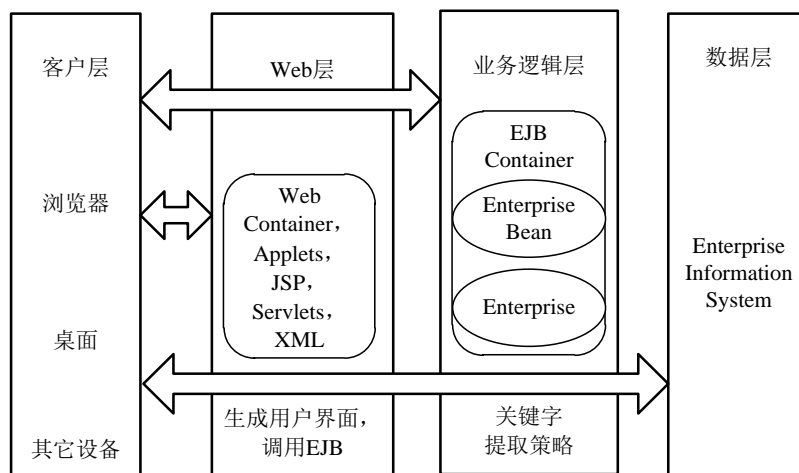


图 2-1 JavaEE 结构图

### 2.1.2 B/S 架构

B/S 架构是目前较为流行的用户访问系统的软件开发模式。因为这个架构是基于浏览器的，软件研发主要关注后台的业务处理逻辑，通过三层架构可以将数据与表现分离出来。前端是客户端，主要是用户的浏览器<sup>[38]</sup>。用户只需使用自己设备中的浏览器进行访问，不用安装部署任何软件，后期的使用更没有升级维护软件的麻烦。中间是服务器端，是进行业务处理的。后台是数据库，详细的架构具体如图 2-2 所示<sup>[39]</sup>。

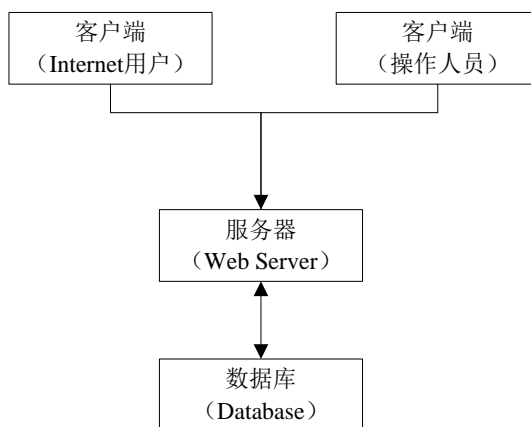


图 2-2 B/S 架构图

在上述的 B/S 架构中主要包括了三个部分，分别是客户端、服务器、数据库：



(1) 客户端：主要的工作是接收用户的数据信息。由用户端的浏览器工具来完成，显示系统所处理结果。用户端在浏览器中通过 HTTP 请求将数据发送给服务器端，对于符合要求的请求在服务器端进行处理，将结果显示在用户端中。

(2) 服务器端：主要是进行业务的逻辑处理，在三层架构中起到上下承启作用。

(3) 数据库：主要是进行中间业务的请求，对各类请求的响应和处理，并将处理结果的数据完成存储<sup>[40]</sup>。

综上所述，在本系统的开发中，主要工作是中间业务的处理和数据库设计，充分利用三层架构的优势，建立业务与数据之间的连接。

### 2.1.3 数据库技术

经过选型和比较，学生成绩分析与预警系统在数据库存储上选定的是微软公司的 Microsoft SQL Server 2014，在这个数据产品中可以满足用户的业务和性能需求，可以按照用户的需求进行数据对象的存储，在结构化数据中处于相对领先的地位<sup>[41]</sup>。对于数据库的产品来说，不仅仅是考虑数据容量的问题，还需要考虑的是数据存储的成熟度问题。通过对各类的数据库产品的比较，本文选定的 Microsoft SQL Server 2014 这个数据库可以满足需要<sup>[42]</sup>。

随着信息技术发展，目前的数据类型变化多端，对于数据库产品来说，也需要跟进，使之能够适应这样的变化，可以为用户提供良好的数据存储的服务，并对海量的数据对象进行安全稳定的存储<sup>[43]</sup>。对于数据存储来说，一方面是存储容量，另一方面是需要有较好的管理便利性。目前，数据库产品在经过几十年发展之后，能够适应现代社会对数据存管理的需要，能够完成良好的数据传送的要求<sup>[44]</sup>。

此外，数据库管理需要有统一标准规范的语言，才能够完成数据对象的设计和存储，完成各类的数据对象的操作，通过标准规范的数据管理语言和操作，完成数据对象的各类需求。目前这个数据产品管理系统可以很好地满足用户的需求，可以对数据对象提供增加、删除、修改和更新的服务<sup>[45]</sup>。

在 Microsoft SQL Server 2014 数据库管理系统中，具有如下的优势与特性，有如下几个方面<sup>[46]</sup>：

首先，它可以很好地满足数据的移植性能，可以在不同的平台中运行，代码经过良好的测试，具有安全稳定的运行效率<sup>[47]</sup>。

其次，它具有较好的适应性。在不同的操作系统中可以很好地运行，可以支持不同的用户语言的习惯，满足多语言的接口。可以在源代码中实现数据对象的高效查询，满足极短时间范围内的数据对象的查询<sup>[48]</sup>。

最后，数据访问方式便捷，数据连接方式提供多种工具实现。可以有很好地提高数据管理效率，能够支持数据存储引擎，满足海量数据存储需要<sup>[49]</sup>。

综上所述，经过分析比较，Microsoft SQL Server 2014 与其他的数据库产品和系统具有一定的特性和优势，可以满足用户的数据存储和管理的要求。

## 2.2.数据挖掘技术

### 2.2.1 数据挖掘算法

数据挖掘是通过一定的工具对目标具有海量数据对象进行采集、处理，经过一定的去除噪声等数据处理之后，可以对其中具有价值的数据进行有效提取，完成数据价值的分析与发现过程<sup>[50]</sup>。在数据挖掘中，最重要的是数据的发掘算法。通过数据的挖掘，使得数据中内在的有用数据价值信息得到提取，实现数据知识的发现。数据挖掘主要包括了如下的几个步骤：数据预处理、数据挖掘、评估验证等，具体的处理的流程如图 2-3 所示<sup>[51]</sup>。

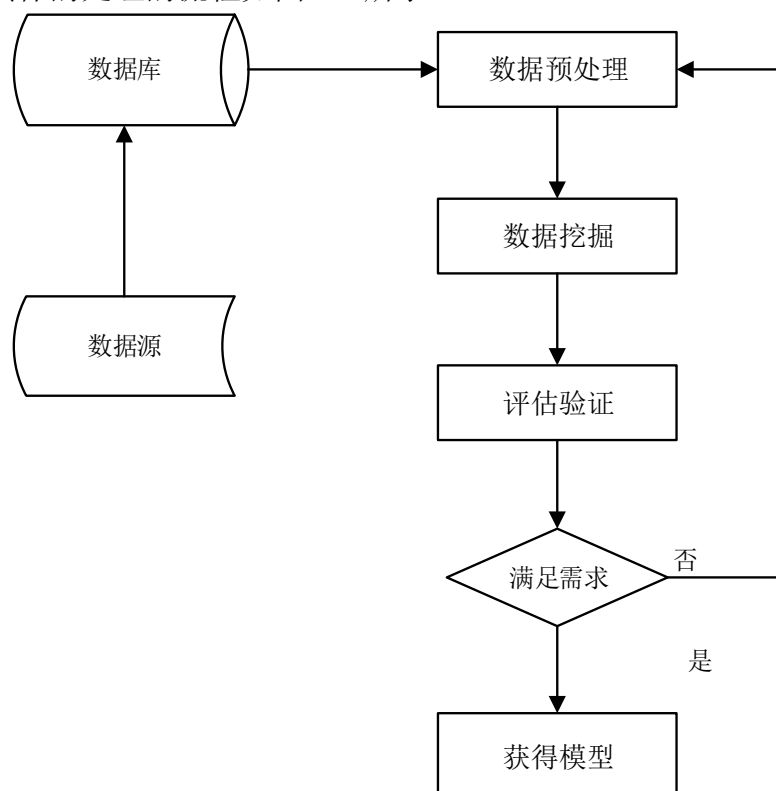


图 2-3 处理流程图

在数据挖掘中，需要对数据对象进行一定预处理，为后续的数据潜在的价值提供基础服务。通过预处理实现数据的准备，进行数据源的获取和管理，在准备好相关的数据之后，就可以进行挖掘的设计<sup>[52]</sup>。利用现有的数据挖掘的算法进行数据分析和知识提取，最后对提取的数据信息进行有针对性的分析，获取得到

数据挖掘的深度结果,可以为相关的人员提供良好的决策支持服务。在进行数据挖掘时,对于原始的数据对象,可以是结构化的数据,也可以是非结构化的数据类型。对数据对象的分析,获得数据之间的关联和知识,利用现有的数据挖掘的技术和算法,实现数据的有价值的体现。目前对数据对象处理涵盖了人工智能、机器学习、数理统计等相关领域的知识,也包括了计算机科学数据信息的内容。因此,数据挖掘是具有交叉学科的特性<sup>[53]</sup>。

综上所述,目前的数据分析的工具和技术主要是数据挖掘的算法,主要包括了模糊理论、决策树、遗传算法等等。对本系统开发来说,采取决策树算法。

### 2.2.2 不同算法比较

对于学生成绩这种数据对象来说,主要是具有结构化的数据库存储数据。在进行数据分析和挖掘的时候,可以采取如下几种算法进行处理:K-Means 算法、Support Vector Machine 算法、决策树算法等,不同的算法对数据处理时候有不同的环境和适应要求,通过对比其性能,选定适合学生成绩数据对象的算法<sup>[54]</sup>。

#### (1) K-Means 算法

数据挖掘中的 K-Means 算法属于对目标数据对象的聚类分析。这个算法是通过数据实体对象中的属性进行一定的划分,并且将其聚类的特性进行抽取获得得到正太分布求最大期望结果<sup>[55]</sup>。在这个算法中可以对数据对象的平均数据值进行计算,具有比较好的收敛的性能,能够在海量数据情况进行数据聚类的处理。缺陷是在进行收敛过程中需要聚类中心的处理,如果数据对象处于非结构且不平均情况下,聚类的计算就会出现精度过小的问题<sup>[56]</sup>。

#### (2) Support Vector Machine 算法

Support Vector Machine 算法在对数据的挖掘分析时,从支持向量机的监督学习中,可以有效的减少数据计算的依赖经验,能够实现几何边缘最大化处理,计算出目标参数的最大似然估计。因此,在这个算法中,能够实现对目标数据中小样本数据的处理。但是,如果数据规模较大的时候,算法具有一定的缺陷,计算效率就会受到制约。

#### (3) 决策树算法

决策树算法能够完成对目标数据的聚类处理,从海量的具有无关联规则的数据经过处理完成有用价值的提取,实现对数据的影响因素的计算和分析。目前实现决策树的算法包括了两种,分别是 ID3 算法和 C4.5 算法。上述的两种算法实际上在实现对原始数据的关联处理中具有自身的优势和特性<sup>[57]</sup>。总的来说,在 ID3 算法中,主要实现的是信息增益计算,需要利用这个重要的因素建立相关的树,在这个决策树中实现数据信息抽取。C4.5 算法实际上是对连续属性的处理,

对决策树进行剪枝处理，利用信息增益率消除无关因素，最后形成具有价值的结果树，如图 2-4 所示。

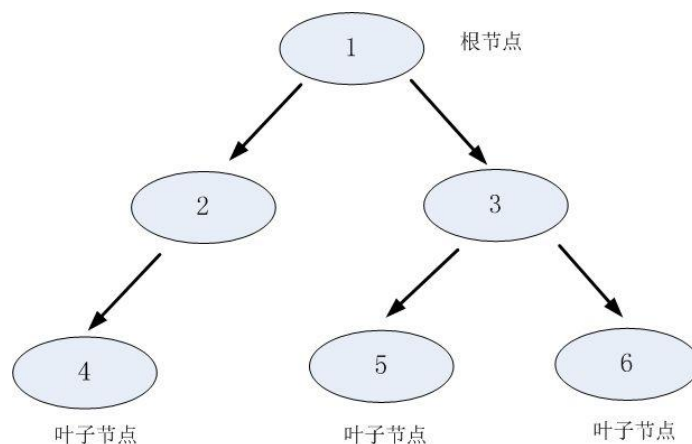


图 2-4 决策树算法图

综上所述，经过对 K-Means 算法、Support Vector Machine 算法和决策树算法等不同的数据挖掘算法的比较，对本系统开发中涉及到的成绩数据来说，成绩原始数据的无关性、数据规模等几个因素的考虑分析，认为 C4.5 算法是比较适合的。与此同时，C4.5 算法具有一定的缺陷，需要对此加以改进。

## 2.3 本章小结

本章主要是进行技术分析，对学生成绩分析与预警系统开发设计所需要的 JavaEE 软件开发技术、数据存储管理技术、B/S 架构模式进行比较和分析，最终确定系统开发的技术路线和选型。

## 第3章 学生成绩分析与预警系统需求分析

本章主要是学生成绩分析与预警系统的需求分析。首先是介绍学生成绩分析与预警系统开发可行性，从技术、经济、操作等三个层面分析系统在研发阶段是否可行。接着是对学生成绩分析与预警系统功能的需求分析，介绍了功能方面的要求，实现用户需求的建模，分析系统的设计需求。最后是系统性能的需求。

### 3.1 系统概述

学生成绩分析与预警系统的设计主要是针对传统系统设计中成绩数据仅仅只是做数据记录和统计的问题，对现有系统进行分析和改进，以及对基本学生成绩录入和统计系统中的各个业务进行分析，解决目前系统中出现的业务功能不全、无数据挖掘分析功能、数据缺少业务流程、数据统计不准确等问题，完成系统的信息化数字化建设。随着互联网+、大数据、人工智能技术的迅速发展，对学校的管理业务来说，需要通过新信息技术对业务处理流程进行重构和优化，改变传统的人工汇总分析的模式，加强数据分析挖掘的功能，实现更为有效的信息化功能。

在本系统的设计中，需要对学生的成绩数据进行原始的分析，对其中的数据进行预处理。在采集完学生成绩数据之后对其进行一定的去噪等预处理工作，实现数据的科学化和规范化。在传统的学生成绩的管理中，成绩的管理业务主要是依赖于人工方式，通过利用口头等人工方式来告知学生成绩的数据结果。在这种情况下，由于教师对学生的了解不够，加之带有一定的主观性，使得学生成绩的分析与预警没有办法实现，甚至是学生处于成绩的低谷是期也无法做出具有预见性的判断。此外，依靠传统的人工方式也会在一定程度上增加教师的工作。对于学校管理人员来说，没有办法及时有效地获取到学生的学业预警数据和情况，也没有可能在有数据情况下进行预警和决策。本系统的设计目的就是为了解决上述的问题。一方面是对学生成绩有效分析，减少人为因素，提升工作的效率。另外一方面，也可以科学分析成绩数据，做出预警等具有前瞻性的判断。本系统基本满足目前数据分析的趋势，能够很好的改进现有的教学管理模式，提升成绩数据的使用价值，对学生的成绩进行多维度、多形式的分析，从而获取到具有较大价值和意义的数据关联性。

综上所述，学生成绩分析与预警系统就是解决原始数据统计分析中的弊端，提高教学特别是成绩管理的有效性和针对性，完成课程成绩的关联性分析。

## 3.2 系统功能需求分析

### 3.2.1 系统管理

在系统管理中主要完成的业务功能是进行数据参数配置、用户账号管理、用户权限配置处理等，详细的业务功能，通过统一建模技术将其描述为系统管理用例，如图3-1所示。

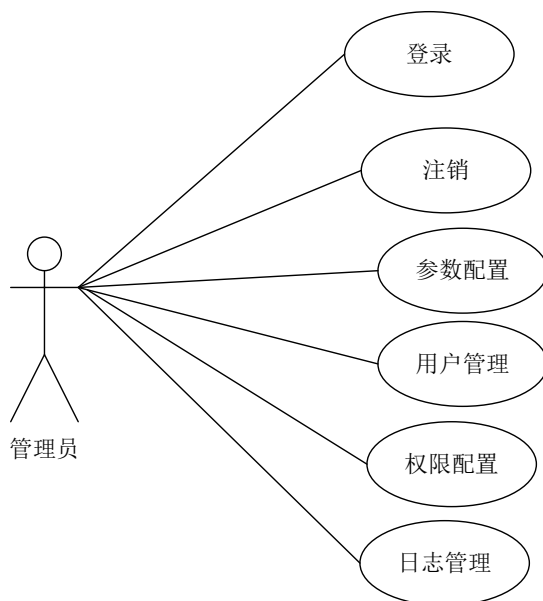


图 3-1 系统管理用例图

在上述系统管理用例中，系统管理员用户完成六个用例，下面对上述的不同的用例作简要介绍。

（1）登录用例。在这个用例中主要完成的是用户身份数据信息的验证，使得进入到系统中的用户是符合要求的，是存在的。

（2）注销用例。在这个用例中主要完成的是用户在完成系统的各项操作之后需要退出系统将 session 信息销毁，使得其他用户能利用此数据进入。

（3）参数配置用例。在这个用例中主要完成的是对系统运行中的各个参数的管理，保证系统稳定高效运行。

（4）用户管理用例。在这个用例中主要完成的对系统中的各个用户账号信息处理，对用户登录或操作系统中信息维护管理。

（5）权限配置用例。在这个用例中主要完成的是用户在系统中操作各个界面或模块所需要的权限内容的管理。

（6）日志管理用例。在这个用例中主要完成的是对系统中的运行的数据进行记录的存储与管理。

### 3.2.2 学籍管理

学籍管理模块主要是对学生信息的管理维护,对学生基本情况维护,如学生的年级、班主任、科任老师等数据对象处理,查询目标的学生详细的学籍管理的用例如图3-2所示。

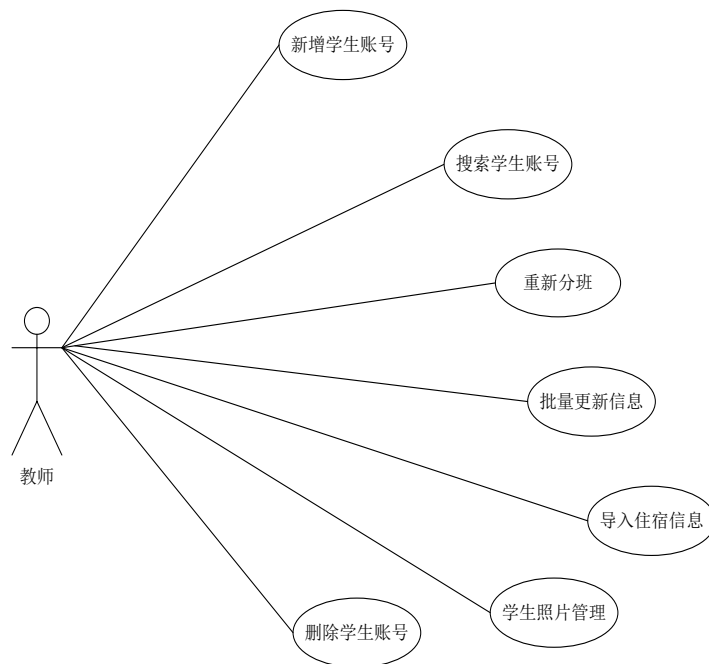


图 3-2 学籍管理用例图

在上述学生学籍管理用例中,包括了七个用例,下面对上述用例简要介绍。

(1) 新增学生账号用例。在这个用例中主要完成的对学生基本信息处理,维护管理学生的学籍数据。

(2) 搜索学生账号用例。在这个用例中主要完成的是按照一定的信息检索目标学生信息,并将其结果显示在用户端。

(3) 重新分班用例。在这个用例中主要完成的是对学生的学籍的移动,通过此功能改变某学生的班级信息。

(4) 批量更新信息用例。在这个用例中主要完成的是对学生的信息出现规模的变动时候,可以通过此导入数据信息到系统中。

(5) 导入住宿信息用例。在这个用例中主要完成的是对学生住宿信息的处理,批量改变更新目标数据。

(6) 学生照片管理用例。在这个用例中主要完成的是维护学生的照片数据信息。

(7) 删除学生账号用例。在这个用例中主要完成的是去除掉学生的学生账号与内容。

### 3.2.3 成绩管理

在成绩管理中主要实现的是对学生课程考试成绩数据的存储,通过导入学生的课程成绩、班级成绩等内容,完成学生成绩录入和管理,在管理过程中班主任老师主要负责本班级的成绩的管理、查询和维护工作,学生可以查询自己的考试成绩,用例如图 3-3 所示。

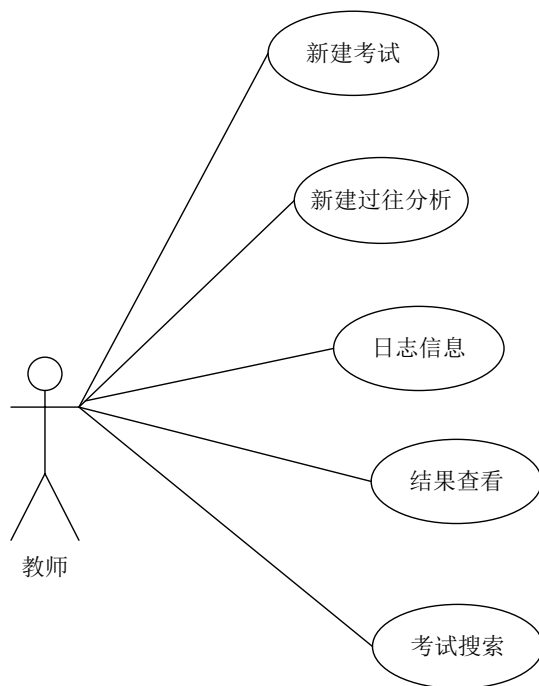


图 3-3 成绩管理用例图

在上述成绩管理用例中,教师用户完成五个用例,下面对上述的不同的用例简要介绍。

(1) 新建考试用例。在这个用例中主要完成的是在系统中建立考试批次,如新建 2019 年第一次月考等考试内容等。

(2) 新建过往分析用例。在这个用例中主要完成的是对之前的考试数据对象建立考试分析的数据。

(3) 日志信息用例。在这个用例中主要完成的是对导入的成绩情况的日志记录的存储,对导入失败的数据对象记录和分析,为后续的处理提供良好的基础。

(4) 结果查看用例。在这个用例中主要完成的是对成绩管理结果的显示和查询,获取得到成绩导入内容和考试结果数据。

(5) 考试搜索用例。在这个用例中主要完成的是对以往的考试成绩查询,查询以往考试的情况。



### 3.2.4 考试分析

在考试分析中主要完成的是对各个分析对象的汇总统计,对具有分析权限的用户,可以查询和显示,并且可以获取得到分析结果的附件信息,具体如图 3-4 所示。

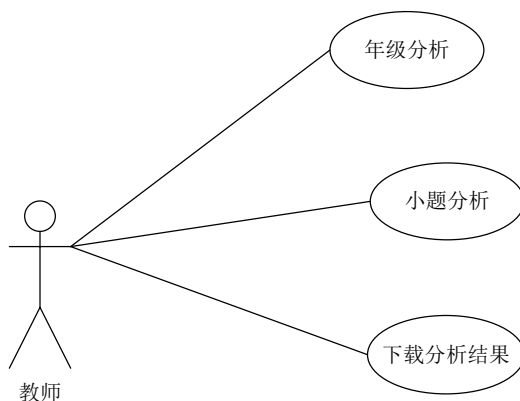


图 3-4 考试分析用例图

在上述考试分析用例中,教师用户完成三个用例,下面对上述的不同的用例简要介绍。

(1) 年级分析用例。在这个用例中,主要完成的是对不同的年级建立考试分析的数据视图。在完成成绩数据的导入之后,对成绩数据中进行一定的处理,如清理和汇总。构建分析的模型,并调用相关的分析模块对成绩数据进行评价分析,形成考试成绩的统计分析,为用户提供具有良好数据支撑的智能化分析结果。

(2) 小题分析用例。在这个用例中,主要完成的是对考试成绩中学生的答题的情况进行分析,如易错知识点的分析,为学生成绩分析提供数据统计的支撑,为教学改革提供有价值的决策。此外,还可以根据实际需要进行填空题、选择题、简答题、计算题等等的分析,进行各个小项目的处理,形成良好的分析报告,可以详细的了解到学生考试成绩的情况,对部分没有完成好的内容进行强化和重点处理。

(3) 下载结果分析用例。在这个用例中,主要完成的是对分析结果的运用,对各个分析的结果通过附件形式提供给用户。对考试成绩进行分析,从而得到有价值的结果,用户可以对此查询和下载。

### 3.2.5 预测分析

成绩预测是本系统的特色，根据前面的考试成绩的数据的导入、清洗，按照数据挖掘的算法对预处理之后的数据进行集成。对课程的成绩深入的分析，对其中具有关联性的数据进行处理，最终获取得到学生个体的成绩趋势报告。系统对学生的学业发展趋势做出一定的分析判断，对学业可能存在一定问题的同学发出预警信息等。因此。教师可以很好地提升教育教学的效果，详细的用例如图3-5所示。

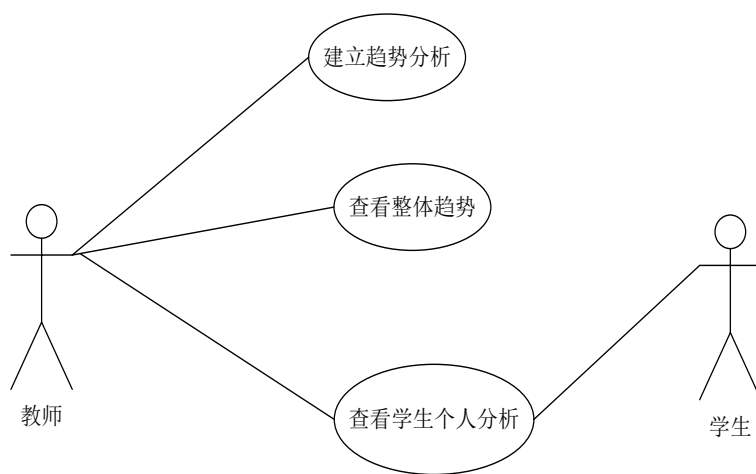


图 3-5 预测分析图

在上述预测分析管理用例中，教师和学生用户完成三个用例，下面对上述的用例简要介绍。

(1) 建立趋势分析用例。在这个用例中，主要完成的是对学生的成绩可能出现的变化和发展做出判断。在对以往的历史的数据进行挖掘和分析，得到学生的趋势发展变化，对学生的成绩做出判断。如果偏离一定的参数，那么就有可能存在学业的问题，系统就会对此学生发出预警的消息和善意的提醒。

(2) 查看整体趋势用例。在这个用例中，主要完成的是学校对整个成绩的数据进行基本的分析，做出课程成绩的发展趋势的判断，并且做出成绩预测功能。

(3) 查看个人分析用例。在这个用例中，主要完成的是对具体的某一个的学生进行个人成绩变化发展趋势分析和报告，查看各个课程之间的挖掘分析报告。

### 3.3 系统性能需求分析

前面已经从五个层面完成系统可行性分析，本系统的开发是可行和必要的。在这个基础上，下文将主要从五个方面分析系统性能的需求，分析了解系统在性能即非功能性方面的需求和要求。

(1) 安全性需求。系统在研发过程中重视系统环境和数据安全，通过权限和业务访问控制模型保证系统用户的隔离，使得系统中的不同的用户组具有一定的操作权限，按照科学合理的要求进行操作业务。系统还要具备有数据备份的能力，完成结构化数据的数据备份，保证数据的安全。

(2) 稳定性需求。系统在运行阶段不能经常出现故障，一旦出现故障可以在极短的时间内恢复。因此，系统的故障修复时间需要保障在 2 个小时之内，确保系统稳定可靠的运行。对于出现的故障，需要系统有相应的数据表进行日志记录，提供给相关的用户查询和分析。

(3) 健壮性需求。系统运行需要保障各个代码运行容错能力，保障系统中的千行代码只允许出现不多于 2 个错误，使得系统具有良好的健壮性能。

(4) 共享性需求。系统是开放的，对有需要的外部系统的集成可以提供二次开发的接口，使得系统可以与其他外部系统实现资源共享和软件集成。

(5) 易用性需求。系统有着比较简洁、非常友好的操作界面，提供可视化的操作界面和窗口，使得系统能够在不同的环境下方便展示，满足不同用户的操作要求和习惯。

### 3.4 本章小结

本章系统的需求分析。通过前面的分析，可以得到系统功能需求和性能需求，从三个层面分析系统开发是可行和必要的。在这个基础上完成系统性能需求分析，分析了解系统在性能即非功能性方面的需求和要求，从而可以建立系统的需求模型。

## 第4章 学生成绩分析与预警系统的详细设计

本章是系统的详细设计。根据前面的需求分析得到的系统业务要求，对系统进行总体设计，对网络架构进行部署和规划，将系统划分为三个逻辑结构层次完成体系架构设计；接着是系统功能架构划分，将系统分为系统管理、学籍管理、成绩管理、考试分析和预测分析等不同业务功能模块，针对上述业务功能进行详细设计；最后是数据库设计。

### 4.1 系统总体设计

学生成绩分析与预警系统将根据前面需求分析和设计要求，进行网络结构设计、体系结构设计和业务功能模块设计，下面是其详细设计过程。

#### 4.1.1 网络结构设计

系统运行需要一定的物理和网络环境，需要物理层面的支持。因此根据前面的需求分析，设计相应的区域网络，部署系统网络结构和设备，完成客户端、应用端和数据端的设备部署。在系统的网络结构设计中，因为办公区域的网络主要是非军事区网络部署。实际上用户端的设备，通过网络识别，穿过内部防火墙，从而可以访问到机房网络的服务器设备。在机房网络中，主要包括了 Web 服务器、数据库服务器、FTP 文件服务器等设备。对于外部用户来说，可以通过互联网连接到外部防火墙设备进入到机房网络，在获取授权后可以访问到系统。此外，为了加强网络区域安全，部署了入侵检测设备，网络部署和结构如图 4-1 所示。

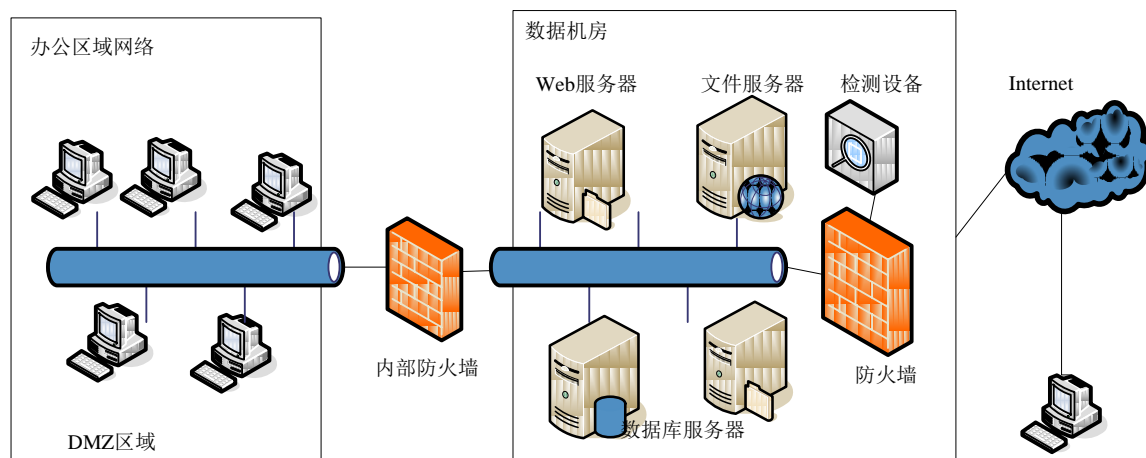


图 4-1 网络部署结构图

### 4.1.2 体系结构设计

系统逻辑结构设计中，主要包括了三个层次内容，不同层次之间具有逻辑关系，完成不同的功能。下面对各个层次的设计详细介绍其功能。对于系统的分层来说，三个层次是相互独立的，不同层次之间通过数据接口实现数据的传送，这样就可以使得系统业务模块之间是相对松耦合的，最终形成层次分明、架构清晰体系结构。详细的体系架构设计如图 4-2 所示。

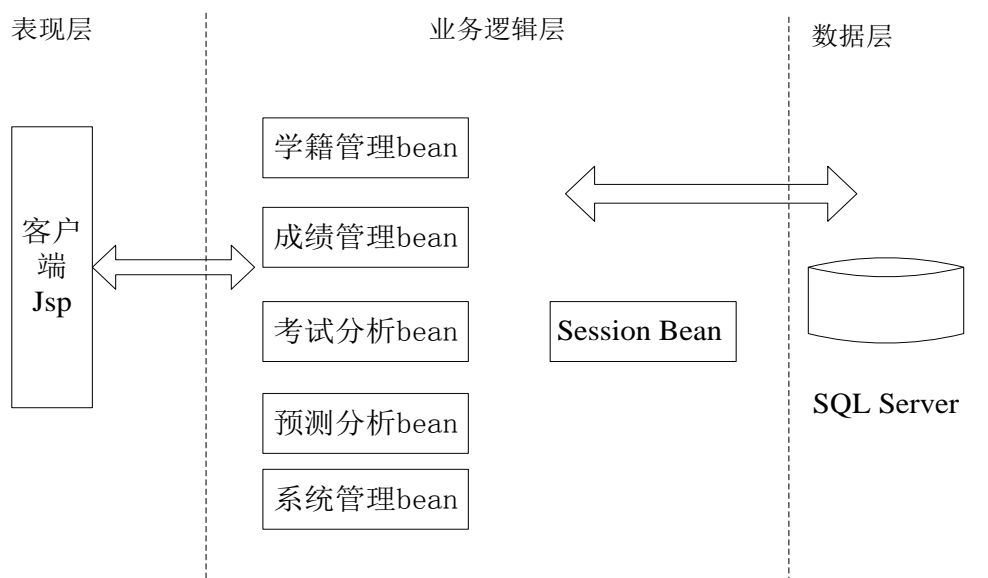


图 4-2 体系架构图

(1) 用户界面层。首先是视图层，在视图层的页面显示验证身份，将用户各类数据信息展示在用户端。对于用户登录的 session 值表单数据处理，按照一定的规则对各类用户端的数据进行处理，将不符合规则的数据过滤掉，并且将完成过滤的数据传送到业务逻辑层中。此外，控制层按照用户配置要求将相关的请求跳转到某个页面中。

(2) 中间业务逻辑层。在这个层次中主要完成的是对各类业务处理，通过对类和方法的处理，完成业务处理，按照算法的运行方式完成成绩数据的处理。在基于 SSH 框架技术的软件开发中采取的技术手段是 Spring 框架，通过 Spring 框架完成业务逻辑处理，并且利用 Spring 框架技术对数据访问实体的定义和开放相应接口完成业务的处理。为此，在业务逻辑层中需要将 Spring 框架与 Hibernate 框架相结合，实现面向接口的设计，这样也有利于系统的进一步扩展，开发完成业务处理接口。

(3) 数据库层。在这个层次中，主要完成的是数据对象的存储，对数据对象进行封装处理之后，为用户提供良好的访问、处理环境，如实现数据对象的添

加、修改、删除等。

### 4.1.3 功能架构划分

根据前面对学生成绩分析与预警系统功能需求分析，完成业务功能的建模，学生成绩分析与预警系统采取模块化设计理念，详细的业务功能模块设计具体如图 4-3 所示。

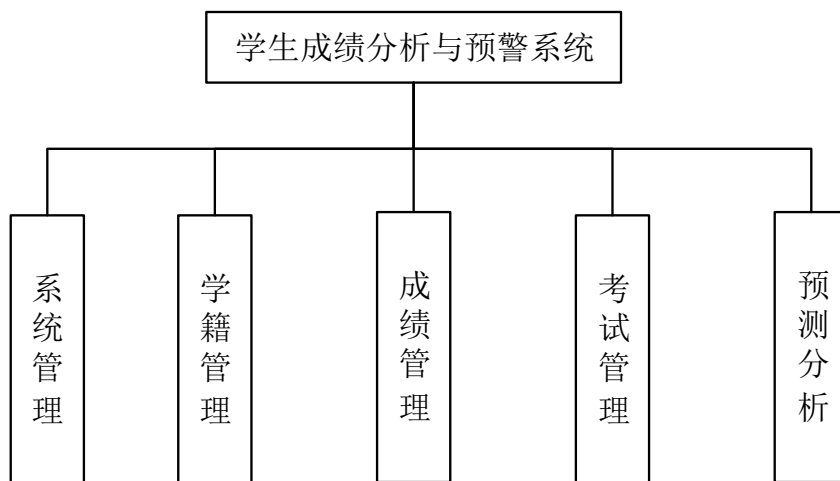


图 4-3 业务功能模块图

## 4.2 功能模块设计

在上述学生成绩分析与预警系统模块划分中将整个系统分为五个主要功能业务模块，分别是系统管理模块、学籍管理模块、成绩管理模块、考试管理模块和预测分析模块，下面将针对上述的业务功能模块详细设计进行介绍。

### 4.2.1 系统管理模块设计

在系统管理功能中，用户可以在模块中完成身份信息的验证，之后符合要求的用户才允许获得访问系统的权限，才能进入到系统模块中操作页面。在系统管理也包括了对系统正常运行中所需要的参数的配置，通过配置完成系统运行基本要求。在系统管理中，针对用户、权限和角色的管理，设计相应的实体类，不同实体类之间的内容和关系详细如图 4-4 所示。

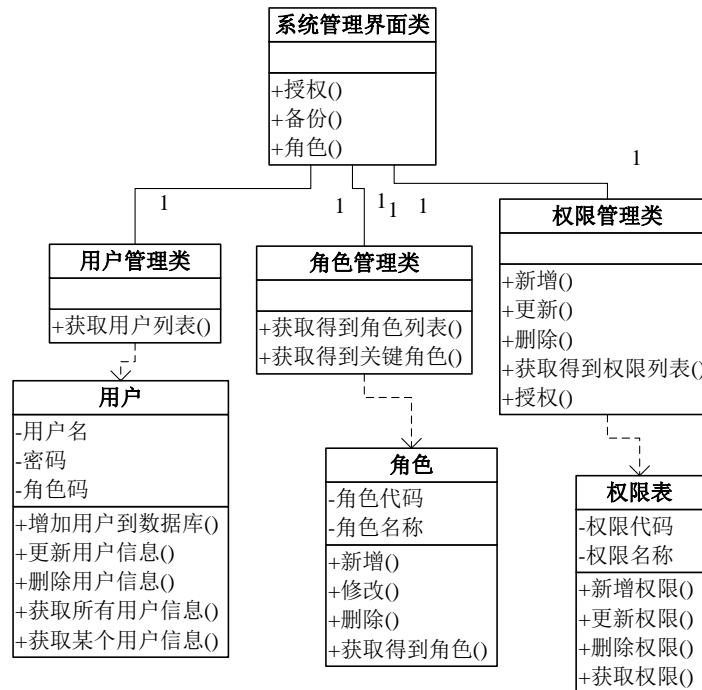


图 4-4 实体类关系图

在上述系统管理实体类之间的关系中，可以看到不同实体类之间存在着的一对一关系，如系统管理界面类与用户管理类之间。对于系统管理来说，日常工作中需要对用户的权限进行配置处理，详细处理的时序如图 4-5 所示。

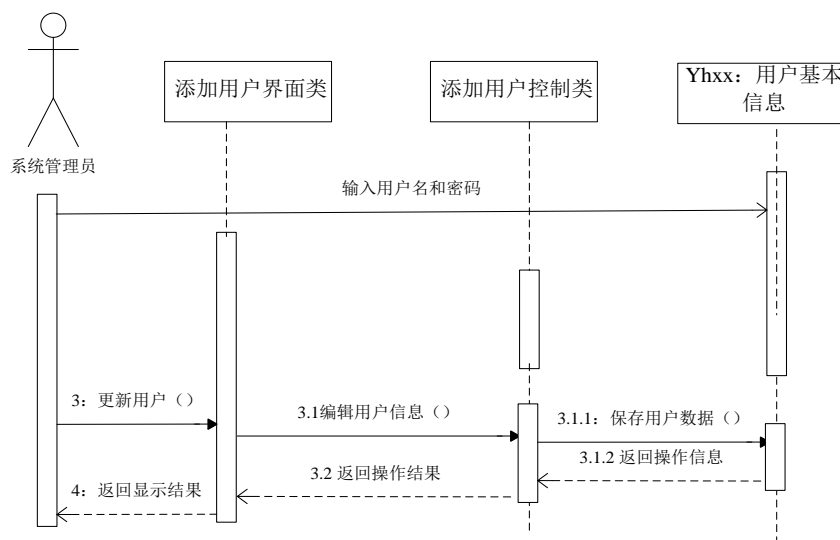


图 4-5 权限配置时序图

在上述关于用户权限配置的时序处理中，系统管理员在输入用户名和密码之后就可以进行用户身份信息的验证。完成验证之后，系统管理员就可以进入添加用户权限界面操作。在添加用户界面，系统管理员可以编辑用户的权限信息和类型，将用户的权限配置后提交到添加用户控制类中进行处理，并且将处理结果提

交到数据库中保存，最终存储操作结果信息。最后将结果返回到操作人员的界面中完成权限信息配置。

此外，成绩数据是用户的核心重要资产，在系统管理中需要对成绩数据库进行数据的备份，防止由于误操作或其他系统故障导致的数据丢失，需要采取一定的备份的策略和方式来提高其安全性能。详细数据备份的业务流程如图4-6所示。

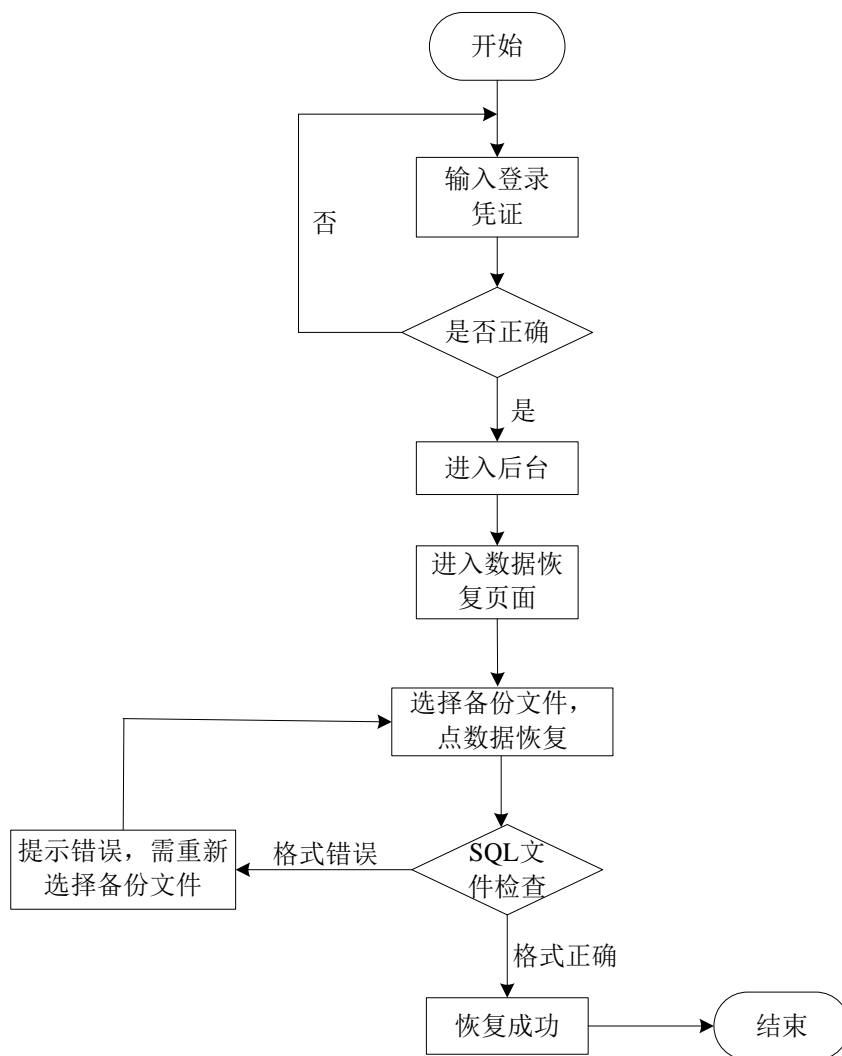


图 4-6 数据备份流程图

#### 4.2.2 学籍管理模块设计

学籍管理模块完成的是对学生基本信息的存储维护。系统会在学生的班级、年级等基本信息在出现异常变动时，及时更新和处理，完成学生任课教师的关系维护管理。对于学生的学籍管理来说，包括了查询目标的学生详细的学籍情况等业务。学生的学籍基本信息直接关系到成绩分析预测的结果，因此，对于学生基



本信息的维护非常重要。添加学生学籍信息的流程如图 4-7 所示。

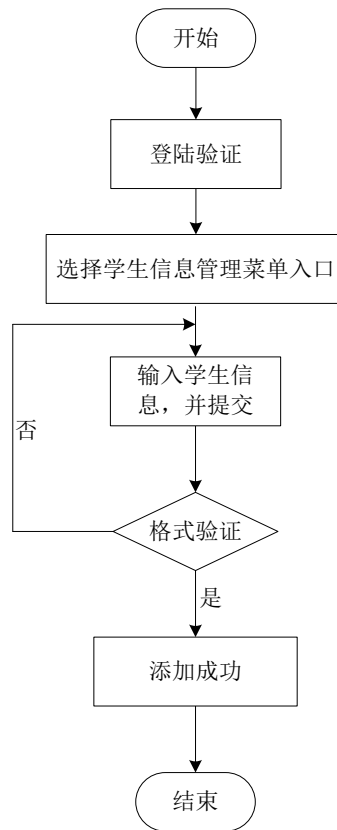


图 4-7 学生学籍信息添加流程图

在学生的学籍管理中, 对于学生需要重新分班的业务, 需要通过如下的实体类来完成。为此本文设计了学生基本类、任课教师基本类、班级信息类等, 不同的实体类设计及其关系详细如图 4-8 所示。

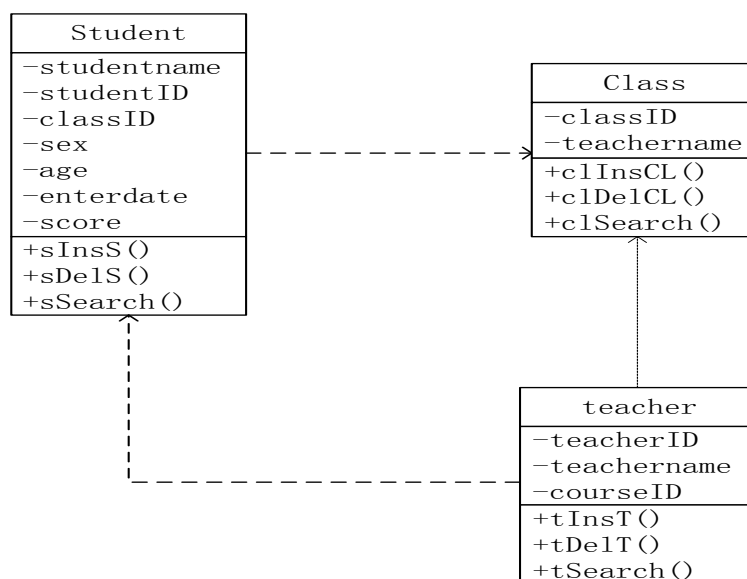


图 4-8 班级管理类图

在上述班级管理类中, 对于学生类来说, 一个学生只能在一个班级中, 因此

是一对一的关系。但是，一位教师可以带多个班级，也可以教授多个学生。因此，对教师类来说，与学生和班级类都是一对多的关系。通过对班级信息的编辑处理，可以完成对班级信息、任课科目信息的处理，也可以查询班级等数据信息。此外，在班级管理，也可以对学生基本信息进行维护，如学生个人、班级信息等。

### 4.2.3 成绩管理模块设计

成绩管理模块的主要功能是对学生课程考试成绩数据的存储和处理。通过导入学生的课程成绩、班级成绩等内容，完成学生成绩录入和管理，在管理过程中班主任老师主要负责本班级成绩的管理、查询和维护工作，学生可以查询自己的学生成绩。在进行学生成绩中的处理，还涉及到学生考试中的题型管理，通过对成绩中的题型的数据对象的处理，为后续的成绩小题分析提供良好的数据基础。

(1) 题型管理。对于题型的深入分析和处理，可以有效的分析考试情况，挖掘学生在题型中的不同的表现，在题型管理中主要的业务包括如图 4-9 所示。

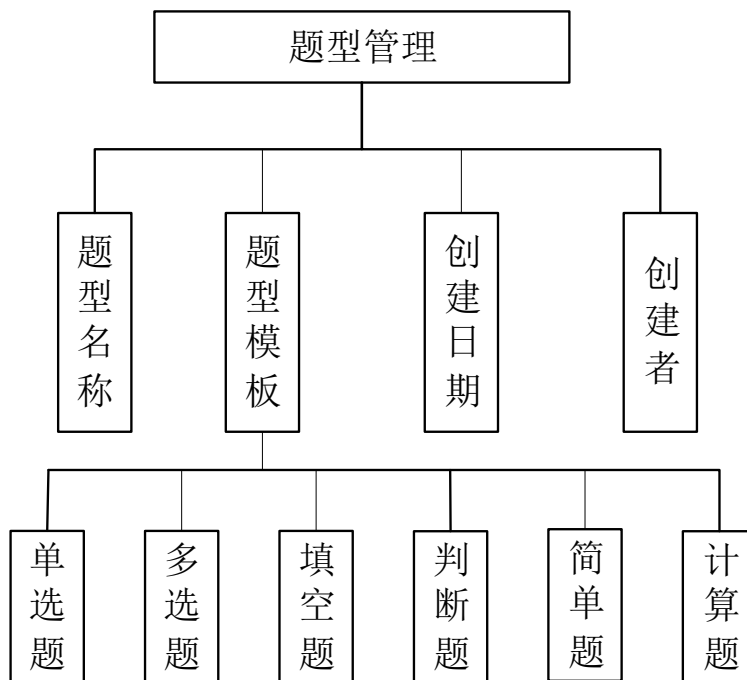


图 4-9 题型功能框图

在上述的题型管理中，包括了题型管理的名称、模板、创建者和创建日期等信息，对不同的题型如单选题、多选题等有效的分析，深入挖掘题型的得分和学生做题情况，直接关系到成绩挖掘的效果，因此，需要设立成绩分析数据的特征。

(2) 成绩数据查询。对于已经完成学生成绩的数据导入，对成绩的数据需要按照一定的范围使用和查询，教师用户在完成身份信息的验证处理之后，就可

以将获取权限的教师用户进行成绩的查询操作，成绩查询设计相应的实体类完成，具体如图 4-10 所示。

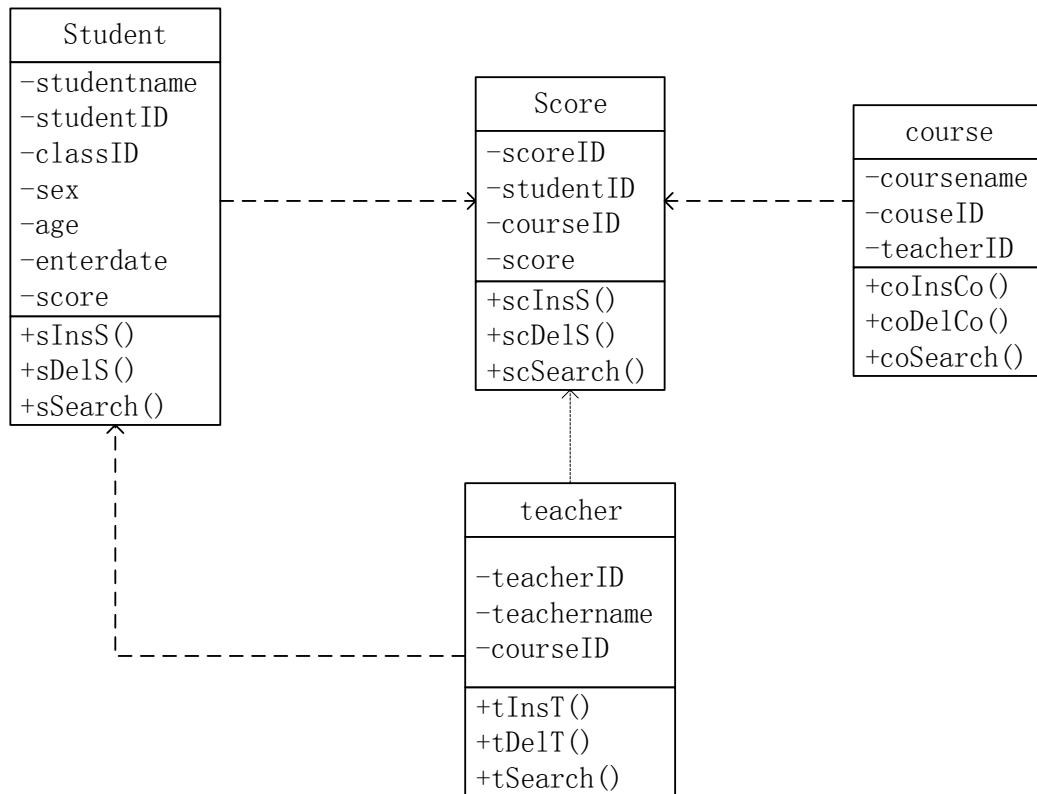


图 4-10 成绩查询管理类图

在上述的学生成绩数据查询业务中，对于学生用户来说，包括了多个科目的成绩数据信息，因此是一对多的关系。成绩实体类，包括了多个的课程，因此也是一对多的关系。对于教师实体类来说，可以担任多个学生的任课教师，也可以多个科目任教，因此与学生和成绩实体类是多对多的关系。

(3) 学生查询成绩。对于学生用户来说，可以查询属于个人的成绩结果，详细的业务处理的时序如图 4-11 所示。

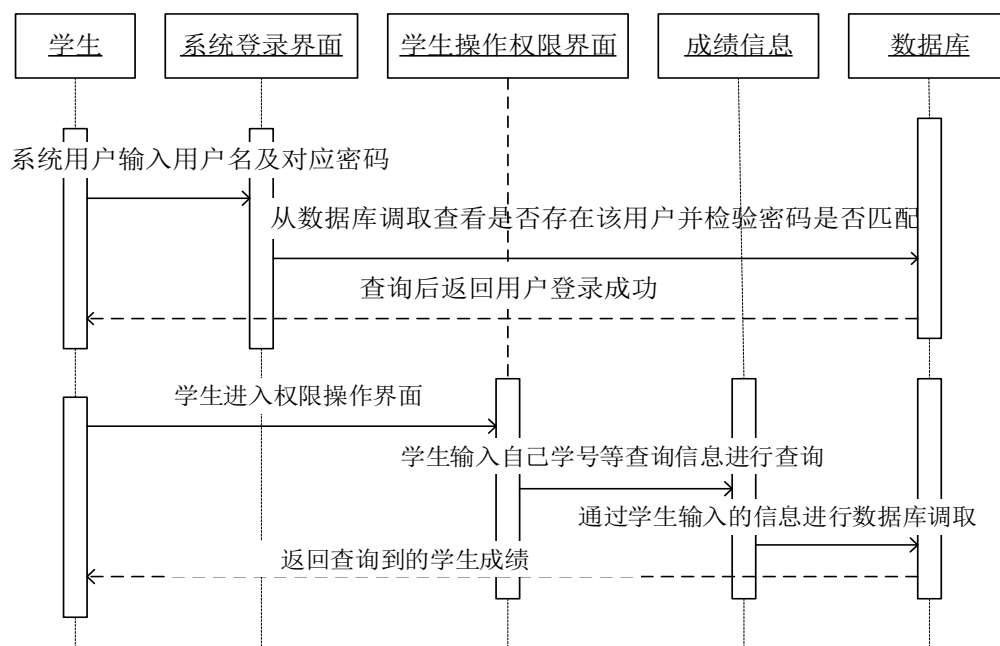


图 4-11 学生查询成绩时序图

在上述学生查询成绩的业务中，学生用户通过自身的账号密码进行身份的验证，在符合系统用户身份之后就可以进去到系统中相应的界面和模块中，按照课程或学号查询到成绩数据。

#### 4.2.4 考试分析模块设计

考试分析的业务功能主要对前面的成绩数据对象进行汇总统计，对小题、年级等按照不同的组对象的汇总统计成绩数据信息。在进行考试分析的时候对具有分析权限的用户，可以查询和显示，并且可以获取得到分析结果的附件信息。

在进行成绩分析时候对导入数据业务处理，详细如图 4-12 所示。

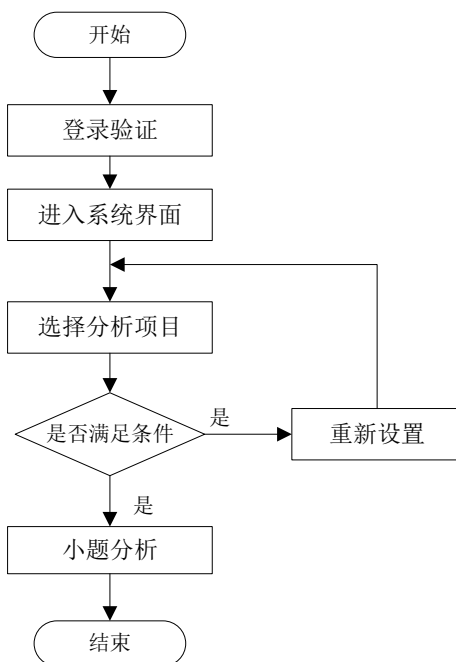


图 4-12 考试分析流程图

在考试分析中，按照用户要求设置分析的参数和内容，系统将根据用户提交的各类分析具体的对象进行配置，并且最终得到考试分析结果显示在用户端，在考试分析中包括了试卷信息实体类、成绩信息实体类、答题情况信息实体类等相关数据信息表，考试分析实体类图详细设计及其关系，具体如图 4-13 所示。

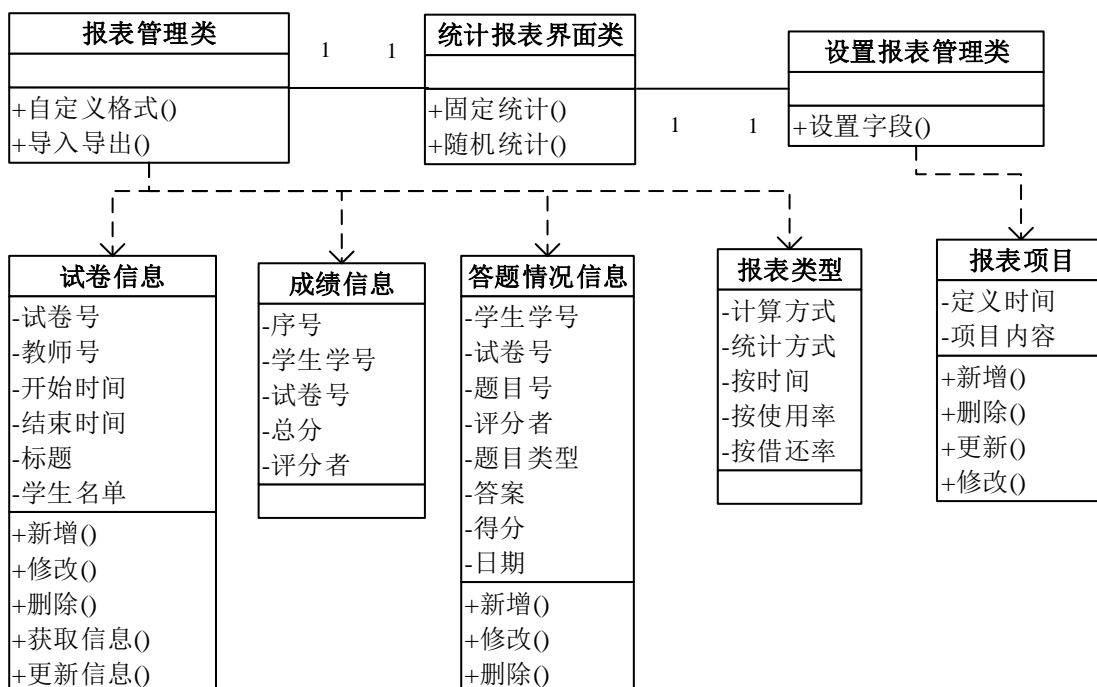


图 4-13 考试分析类图

在上述考试分析管理类中，统计报表界面类与报表管理类是一对一关系，与

设置报表管理类也是一对一关系，报表管理类中包括了多个实体类如试卷信息、成绩信息、答题情况信息等。下面是用户在考试分析模块中完成对成绩等数据的分析处理，详细的业务如图 4-14 所示。

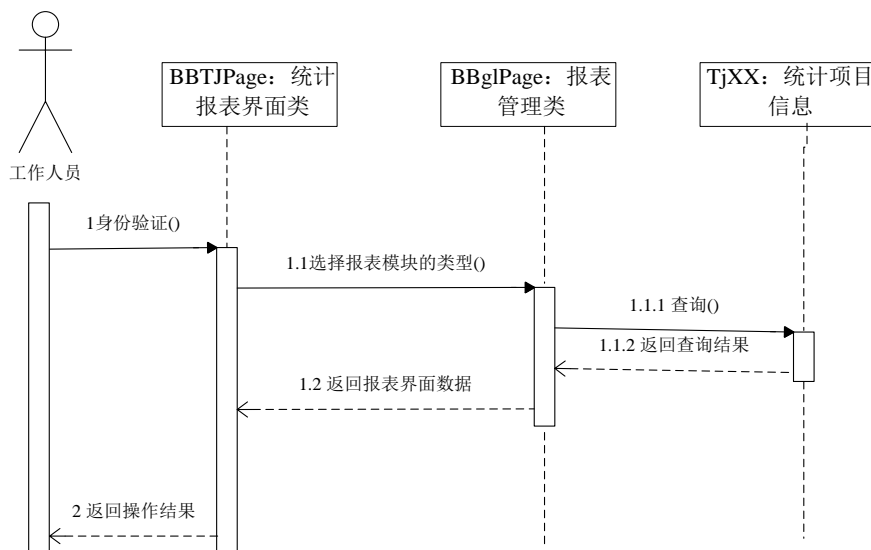


图 4-14 考试分析时序图

在上述考试分析业务处理中，用户在实现对身份的验证后就可以在统计报表界面中设置需要处理报表的类型，并且设置相应的参数和内容，选择之后就可以提交其参数到系统进行处理。在考试分析的业务逻辑收到用户提交的内容之后，如没有数据进行过滤那么就会调用统计项目处理函数，对成绩数据的查询，并将查询结果数据返回到用户端中，最后在报表的界面中获取得到操作的结果内容。

#### 4.2.5 预测分析模块设计

成绩预测是本系统的特色，根据前面的考试成绩的数据的导入、清洗，按照一定的数据分析挖掘的算法对预处理之后的数据进行集成，对课程的成绩深入的分析，对其中具有关联性的数据进行处理获取得到学生个体的成绩趋势的报告，对学生的趋势发展做出一定的分析判断，对学业可能存在一定问题的同学发出预警信息等，可以很好地提升教学管理的效果，具体的业务功能如图 4-15 所示。

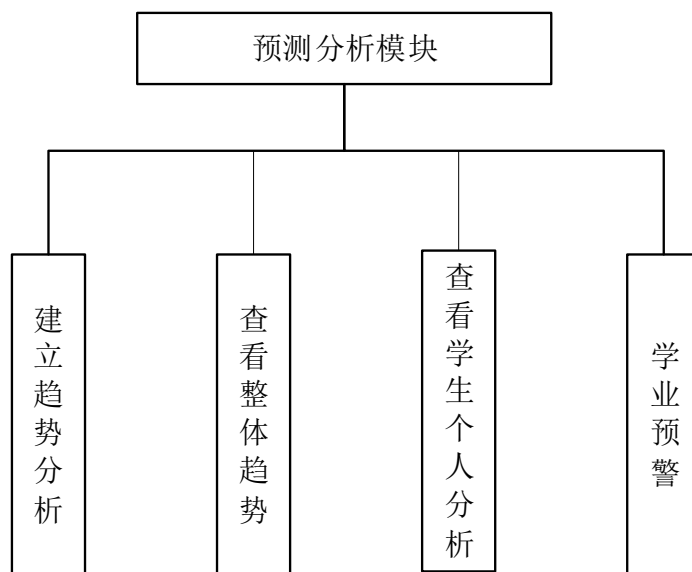


图 4-15 预测分析模块图

对学生的学习情况和成绩进行预测和预警需要依靠海量的数据信息，并且对其进行处理和深入挖掘，建立学生成绩预测模型，才能较好地挖掘在海量成绩数据后面的潜在的联系，挖掘可能存在的知识和价值信息，从而为教学提供决策和帮助。

在对各类的数据挖掘算法进行比较之后，分析不同算法其优缺点，分析其在成绩数据挖掘中的适用性等，在经过权衡后本系统的成绩预测选取的是基于决策树的 C4.5 算法，下面介绍基于 C4.5 算法的成绩数据预测模型的详细设计。

(1) 设置参数及选择分类属性。首先是需要进行数据参数的配置，通过配置完成算法中的最小信息增益率阈值，可以有效地对建立的树的分支进行控制，避免树的分支过多或过小。因此，在进行成绩数据的分析业务中需要有效地对其参数进行选取和配置，选用合适的参数。另外一个参数是选择分类属性。在这个参数中将使得分类的属性在界面中进行有效地分类处理，丰富属性的功能。

(2) 数据预处理。对于成绩的原始数据来说，可能会产生一些噪音的数据，对此数据需要加以一定的处理。对于在日常管理中的成绩数据，需要完成简单转换和离散化的处理，将其复杂的部分进行简单化直观化。例如学生的成绩数据，可以将其转换为更为简单的五个等级为非常优秀、优秀、一般、差和非常差等，这样就可以有效的降低数据类型，可以将其更为简便化。另外对于一些相对复杂的数据，如学生的户籍，如广州、深圳，可以简化为广东等，统一进行简便的转化处理。通过上述的两个的处理可以使得决策树不会过于繁冗，不会产生更多的分支等，这样就可以提高分析的速度和精度，有利于问题的分析和解决。

(3) 产生决策树。根据前面的处理之后就可以对分类属性的信息熵及预先定义的阈值进行计算处理，使得决策树中的分支和节点具有相应映射，建立相应

的映射关系之后就可以生成成绩数据预测的决策树。

(4) 显示处理过程及决策树。通过前面的模型建立，就可以对成绩数据进行处理，并且将建立好的决策树通过工具显示在用户端中。

学生成绩数据预测分析的业务如图 4-16 所示。

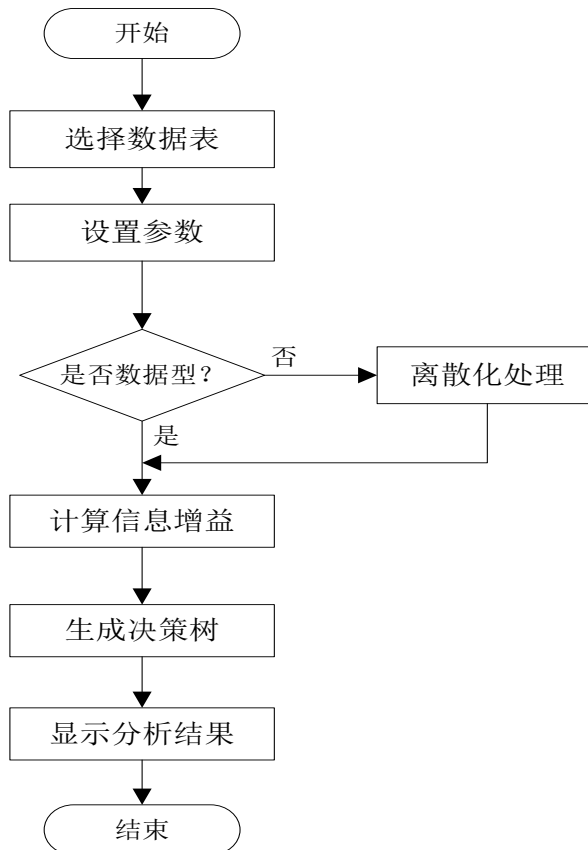


图 4-16 业务流程图

根据上述的业务流程就可以建立起对学生成绩预测分析的模型，根据这个模型进行决策树的生成和显示处理结果。对于所建立的模型可以按照一定的标准对其进行评估分析，验证模型是可行并且是有效的，可以适应学生成绩预测业务需求的，并且满足用户要求的。对所评估的结果进行后续的优化，包括对模型算法中的参数进行调整或对其中的剪枝进行重新配置，使得在目前的成绩预测的数据集下是达到最优效果的。

### 4.3 数据库设计

在系统的数据存储设计中，需要考虑多种因素，不仅仅是存储容量大小的问题，还需要考虑到数据存储的效率以及存储数据的安全等。在本系统的设计中，需要设计相应的各个实体属性，也需要设计相应的表结构、数据访问接口，通过对数据表重要参数设置，用户可以高效地执行查询和访问，对数据进行安全的存



储等。

### 4.3.1 概念结构设计

概念结构设计主要针对前面对系统数据分析,对企业管理中所产生的数据采取技术方式存储,通过科学合理的设计,使得数据存储有质量,提供良好的安全的数据支撑环境,是系统建设重要内容。对数据库设计的好坏,直接关系到系统能否安全稳定高效运行,在进行数据库设计时候需要关注数据存储需求,关注业务处理模式,对各类基础数据和业务数据设计标准规范的存储,为数据的查询、修改提供支撑。通过前期的调研,对真实数据进行抽象分析,建立数据概念模型,使得数据对象可以形成实体关系映射,也就是 ER 图模型,对于本系统的 ER 图模型。E-R 模型如图 4-17 所示:

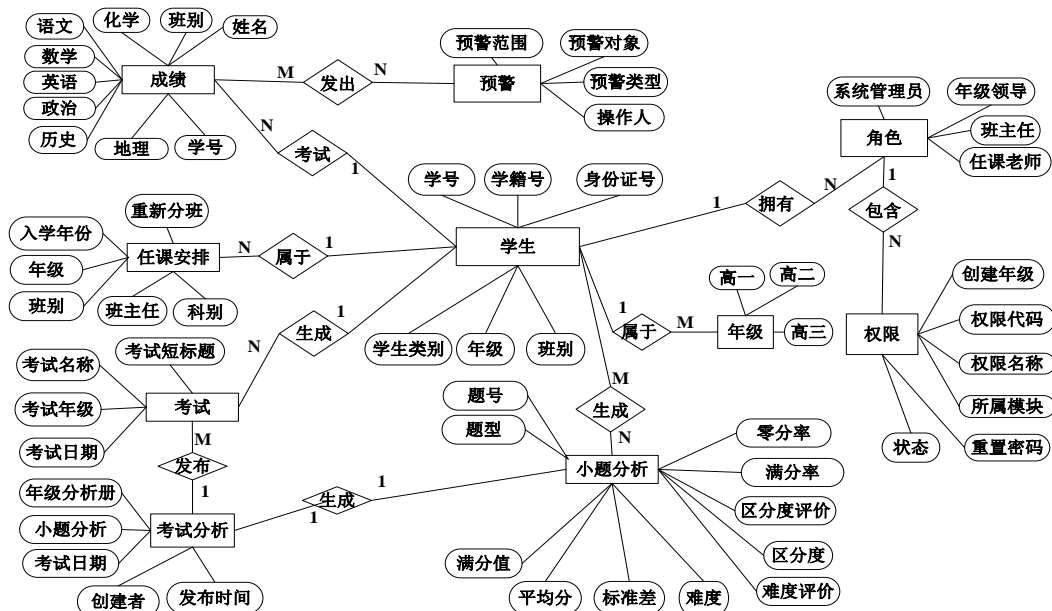


图 4-17 E-R 图

在上述的 ER 图模型中,可以看到,系统包括了用户、会议、公文、考勤等相关的实体对象,实际上是系统中各个业务功能的实体反映。不同实体对象之间存在一对一和一对多的关系,具体表现为 1: N 和 M: N, 由于篇幅所限,目前只是列举了部分的 ER 图模型。

### 4.3.2 逻辑结构设计

根据前面对系统中各个实体的概念结构设计,完成了建立系统数据模型也就是 ER 模型图,利用 ER 模型可以完成系统中后台数据中的数据表设计,从而获取数据抽象。得到逻辑数据表,由于篇幅限制,下面描述主要的重要的数据表以及表结构,具体如下。

表 4-18 学生信息表

字段	数据类型（长度）	字段描述	备注
xh	varchar(10)	学号	
xjh	varchar(10)	学籍号	
xslb	varchar(5)	学生类别	
sfzh	varchar(18)	身份证号	
nj	int(8)	年级	
bb	varchar(20)	班别	

表 4-19 成绩信息表

字段	数据类型（长度）	字段描述	备注
xh	varchar(10)	学号	
xm	varchar(10)	姓名	
bb	varchar(20)	班别	
yw	int(8)	语文	
sx	int(8)	数学	
yy	int(8)	英语	
zz	int(8)	政治	
ls	int(8)	历史	
hx	int(8)	化学	
wl	int(8)	物理	
dl	int(8)	地理	

表 4-20 预警信息表

字段	数据类型（长度）	字段描述	备注
yjfw	varchar(50)	预警范围	
yjdx	varchar(20)	预警对象	

yjlb	varchar(10)	预警类别	
ccr	varchar(10)	操作人	

表 4-21 任课安排信息表

字段	数据类型（长度）	字段描述	备注
km	varchar(10)	科别	
bzr	varchar(10)	班主任	
bb	varchar(10)	班别	
nj	int(8)	年级	
rxnf	int(8)	入学年份	
cxfb	varchar(10)	重新分班	
js	varchar(10)	教师	

表 4-22 考试信息表

字段	数据类型（长度）	字段描述	备注
ksmc	varchar(40)	考试名称	
ksnj	int(8)	考试年级	
ksrq	date(8)	考试日期	
ksdbt	varchar(20)	考试短标题	

表 4-23 考试分析信息表

字段	数据类型（长度）	字段描述	备注
cjz	varchar(10)	创建者	
kssj	date(8)	考试时间	
njfxc	varchar(10)	年级分析册	
fbsj	date(8)	发布时间	

表 4-24 小题分析信息表

字段	数据类型（长度）	字段描述	备注
th	int(8)	题号	
tx	varchar(10)	题型	

pjf	int(8)	平均分	
bzc	int(8)	标准差	
lfl	int(8)	零分率	
mfl	int(8)	满分率	
qfd	int(8)	区分度评价	
qfdr	int(8)	区分度	
ndpj	varchar(10)	难度评价	
nd	varchar(10)	难度	

表 4-25 角色信息表

字段	数据类型（长度）	字段描述	备注
xtgl	varchar(10)	系统管理员	
njld	varchar(10)	年级领导	
bzr	varchar(10)	班主任	
krls	varchar(10)	任课老师	

表 4-26 权限信息表

字段	数据类型（长度）	字段描述	备注
cjnj	int(8)	创建年级	
qxdm	int(8)	权限代码	
qxm	varchar(10)	权限名称	
ssmk	varchar(10)	所属模块	
czmm	varchar(10)	重置密码	
zt	varchar(10)	状态	

## 4.4 本章小结

本章是系统详细设计，根据前面的需求分析得到的系统业务要求，对系统进行总体设计、详细设计，完成网络结构设计包括系统的网络区域划分，系统中的数据服务器、web 服务器等的部署设置，提供良好网络运行环境。接着是逻辑机构设计，将系统分为数据层、业务逻辑层和表现等多个层次，层次比较清晰，按照不同层次分别设计。然后是功能模块设计，根据前面业务分析，进行相应的五个主要业务功能详细设计，最后是数据库设计。

## 第5章 学生成绩分析与预警系统主要功能实现

本章是系统的详细实现，根据前面的需求分析和系统详细设计，完成开发设计阶段的工作，已经确定系统所要达到的物理模型和逻辑模型，为后面的实现编码阶段提供好的基础，本章在前面工作基础上，继续完成系统的实现，实现系统的各项功能业务，具体包括了在搭建开发环境中代码实现和界面实现。

### 5.1 开发环境搭建

根据前面对目标系统分析和设计，完成系统详细设计为后续的开发打下基础，接下来需要搭建符合要求的软件和硬件开发环境。按照前面技术选型要求，系统开发采取的是基于 JavaEE 的技术框架，利用 JSP 页面开发技术、SSH 多层框架，在搭建好的集成开发工具 MyEclipse2014.12.0 下进行代码编写和设计。后台数据存储采取的是 Microsoft SQL Server 2014，通过前端开发和后台设计的工具的搭建和集成，构建出符合本系统开发的软件环境。具体软件的工具列表如表 5-1 所示。

表 5-1 软件开发工具表

软件工具	版本
开发的操作系统	Windows Server 2012 R2
集成开发工具	MyEclipse2014.12.0
数据库	Microsoft SQL Server 2014
编程语言	JSP
Web 运行	Apache-tomcat-8.0.53-windows-x64
数据建模	Microsoft Viso 2010

根据前面的开发要求，下面介绍 MyEclipse 安装部署，在 MyEclipse 中可以对 Java 9 有较好的支撑，能够有效提升开发效率，根据模块完成自定义类的路径处理，修复程序和快速添加 module-info.java 文件。在完成 IDE 集成环境下，可以支撑 JDK 10 的运行，可以说是基于 SSH 框架技术开发的良好环境，JDK 包如图 5-2 所示。



图 5-2 下载 JDK 包

MyEclipse 软件环境如表 5-3 所示。



图 5-3 MyEclipse 界面图

与此同时，良好的集成环境需要一定的硬件支撑，通过良好的硬件环境可以有效支持代码编写和业务功能开发设计，Web 的服务器的环境配置详细如表 5-4 所示。

表 5-4 硬件配置表

功能	品牌	性能参数
数据库服务器	华为 HUAWEI 5885Hv5	Intel 酷睿 i7 4770 内存类型：ECC 内存：8GB 存储硬盘大小：2TB 电源：冗余
Web 服务器	戴尔（DELL）型号 PowerEdge T30	Intel 酷睿 i7 4770 内存：8GB 存储硬盘大小：1TB

经过前面的软件和硬件环境的配置，可以形成良好的开发环境和运行环境，多组件、简使得开发环境部署。接着根据需求分析结果和软件设计进行代码编码实现。由于篇幅限制，在前面的设计阶段完成的每一个业务功能模块都包括了较多的功能，下面知识对部分的业务功能的实现进行介绍。

## 5.2 系统功能模块的实现

### 5.2.1 系统管理模块实现

在系统管理模块中可以对基础数据的配置管理，对用户信息维护管理，实现对学校、班主任、学校领导、教师、学生等基本权限的设置，不同用户其获取的权限是不同的，进入系统后获取到的业务模块也是不同的，这些数据的配置是通过系统管理模块来完成的。系统管理模块的页面如图 5-5 所示。

	学生	教师	考试
高三 2016	918	52	10
高二 2017	924	80	7
高一 2018	907	58	2

图 5-5 系统管理模块图

在上述系统管理页面中，可以看到，包括了学生、教师两个基本用户组的角色，通过其链接页面进入到相应的用户角色中对其配置相应参数信息，从而建立起用户角色与权限的映射关系，在完成配置之后用户就获取得到权限字符串，就可以对完成身份验证的用户分配具有权限的业务，其实现的功能函数 `Split()`。在系统管理模块中也可以添加用户基本信息，可以手工方式添加用户基本信息，也可以编辑用户基本信息，如重置用户密码等。

在系统管理中对于用户权限的处理，在基于三层架构的开发中，主要是通过中间业务逻辑层来完成的，其中对权限的处理的函数为 `ValidatePermission()`，对用户权限的业务流程进行处理，下面是详细的代码。



```
public string ValidatePermission(string varchar,string permissions)
{
    string[] ps = permissions.Split(',');
    switch (varchar)
    {
        case "A":
        {
            foreach (string obj in ps)
            {
                if (obj == "A") return "A";
            }
        }
        break;
        case "B":
        .....//省略部分代码
        case "E":
        {
            foreach (string obj in ps)
            {
                if (obj == "C") return "C";
            }
        }
        break;
    }
    return "";
}
```

通过前面的权限业务功能处理完成角色的配置，下面是对 test 用户进行密码基本信息的处理，在对当前密码完成数据校验之后就可以重新配置新的密码信息，具体页面如图 5-6 所示。

图 5-6 用户配置密码图

在系统管理中还有业务功能是对用户数据的身份验证，下面是系统的用户信息的登录页面，具体如图 5-7 所示。



图 5-7 系统登陆界面

在上述页面中，用户输入符合身份信息的用户名和密码信息，提交到后台数据的处理，之后符合用户数据表中的信息才能反馈用户的身份，才能完成身份的校验，获取进入到系统中的权限。

5.2.2 学籍管理模块实现

在这个功能模块中主要完成的是对学生基本信息处理，可以新增学生账号或批量导入处理学生信息，如图 5-8 所示。



图 5-8 学籍管理图

在年级学籍管理中，实现对学生账号信息处理，如搜索学生信息、重新分班、生日、学生照片管理等等。可以对某个或部分学生的学籍信息维护处理，也可以通过导入批量数据方式处理学籍信息，下面介绍批量方式处理多个学生学籍信息，具体如图 5-9 所示。

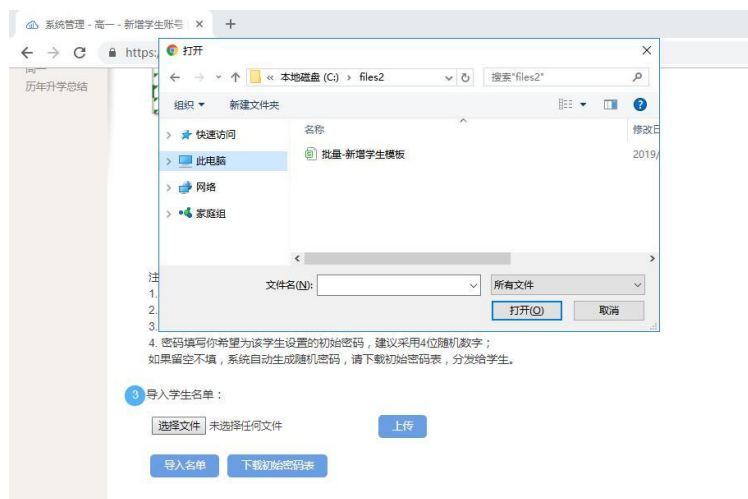


图 5-9 新增学生账号图

上述是批量导入数据时候需要调用相应的附件信息,可以在导入学生名单的按钮中,选择文件,获取到本地文件信息,用户在选择其中的文件之后的打开相应的附件信息,建立数据连接,得到导入数据结果页面,具体如图 5-10 所示。



图 5-10 更新成功图

在学籍管理中也可以对学生基本信息的更新处理,对于出现异常的学生学籍信息,在进行数据的处理时候,需要将其“批量新增学生”,完成导入之后就可以反馈更新学生记录的数据,实现对学生学籍信息的处理。在对学生信息进行更新的时候,需要与后台数据的连接与处理。详细的实现代码如下:

```
String ConnUrl="jdbc:sqlserver://localhost:1433;databaseName=news";
} catch (ClassNotFoundException ex)
...
String[] col = {"学号","姓名","性别","民族","年龄","年级","班级"}; //创建属性列名
```

5.2.2 成绩管理模块实现

在成绩管理中主要功能是对学生考试的成绩进行综合管理，包括对成绩数据的新增、查询，可以按照年级或班级的方式将成绩数据信息存储，也可以单独的按照学生考试成绩数据记录进行数据信息的添加，将符合要求的成绩存储到数据表中的。在成绩管理中主要实现的功能是管理已创建的考试，下载分析结果，成绩管理的页面如图 5-11 所示。

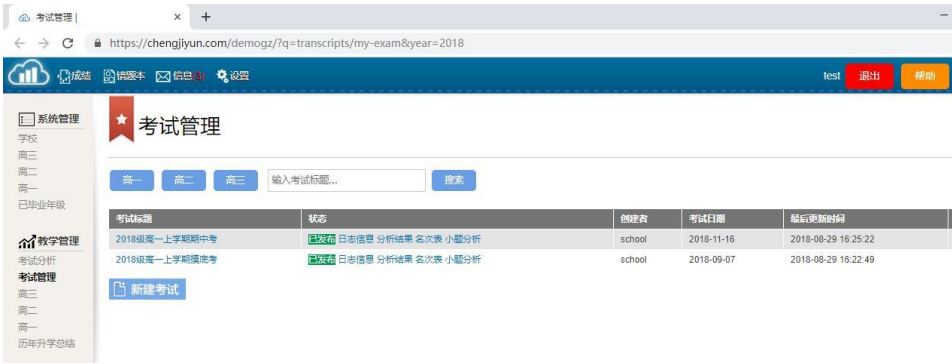


图 5-11 考试成绩管理图

在上述成绩管理页面中，可以根据考试标题进行不同时期不同阶段考试以往数据记录的搜索查询。经过检索之后将符合要求的考试成绩的信息显示在用户端中，包括了考试标题、考试状态、创建者、考试日期和最后更新时间等数据内容。用户也可以在成绩管理的页面中新建考试。创建属于新的考试，并且添加考试的成绩数据信息内容，具体如图 5-12 所示。

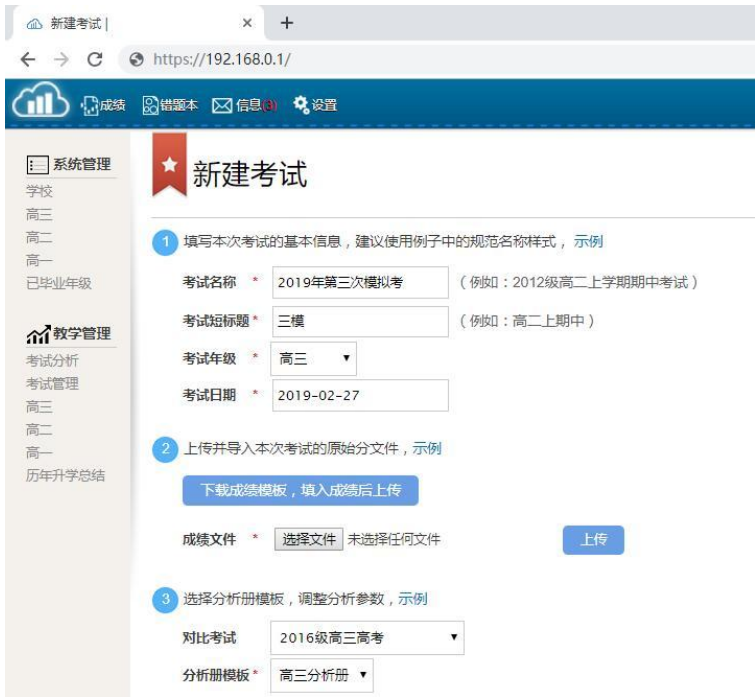
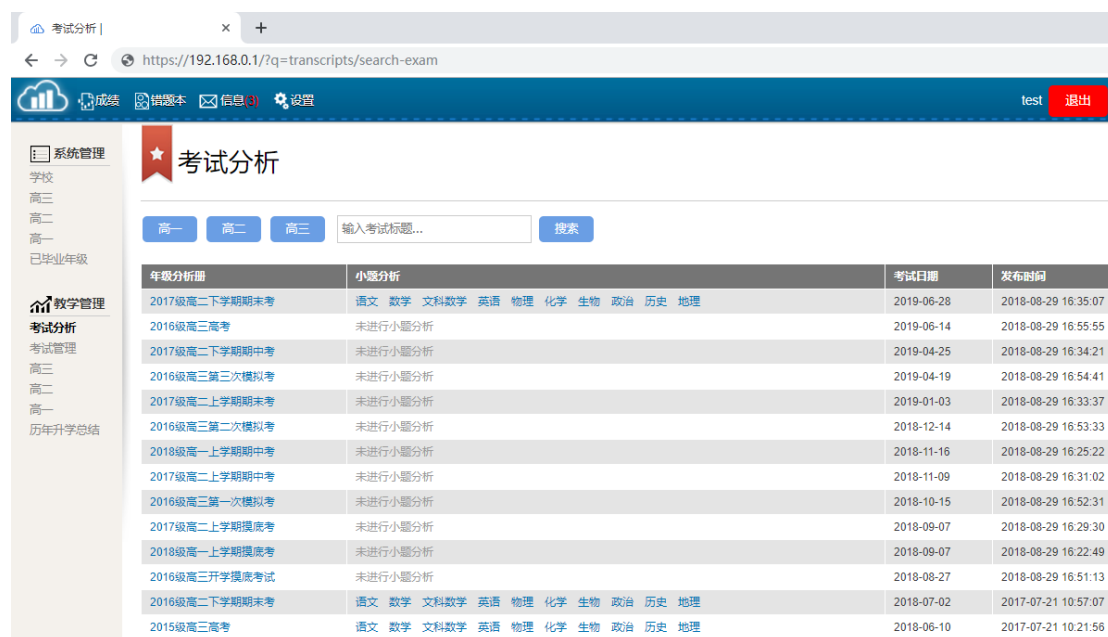


图 5-12 新建考试图

在上述新建考试页面中，用户需要填写考试基本信息，如考试名称、考试标题、考试年级、考试日期基本信息，通过创建 Examination.jsp 来完成数据的提交处理，将其数据提交到成绩管理业务中处理，在控制层中的对用户输入数据的 ExaminationTestAction 类对用户提交的数据信息进行验证，完成前期处理使得所提交的数据是符合要求的。之后符合要求的数据将其提交到业务逻辑层中处理，调用数据持久层中的 ExaminationDAOImpl 类完成对新建考试的数据访问和考试成绩数据对象添加。

### 5.2.4 考试分析模块实现

在考试分析中其主要完成的是对课程考试数据信息的汇总、统计和分析，对各类数据信息按照不同的类别进行统计，获取到按照用户需求的结果，将其显示到考试分析的页面结果中。在考试的分析中需要设置一定的分析类别参数，根据不同的类别参数就可以对原始成绩的数据处理，最终形成符合要求的考试分析结果报告内容。考试分析的页面如图 5-13 所示。



年级分析	小题分析	考试日期	发布时间
2017级高二下学期期末考	语文 数学 文科数学 英语 物理 化学 生物 政治 历史 地理	2019-06-28	2018-08-29 16:35:07
2016级高三高考	未进行小题分析	2019-06-14	2018-08-29 16:55:55
2017级高二下学期中考	未进行小题分析	2019-04-25	2018-08-29 16:34:21
2016级高三第三次模拟考	未进行小题分析	2019-04-19	2018-08-29 16:54:41
2017级高二上学期期末考	未进行小题分析	2019-01-03	2018-08-29 16:33:37
2016级高三第二次模拟考	未进行小题分析	2018-12-14	2018-08-29 16:53:33
2018级高一上学期期中考	未进行小题分析	2018-11-16	2018-08-29 16:25:22
2017级高二上学期期中考	未进行小题分析	2018-11-09	2018-08-29 16:31:02
2016级高三第一次模拟考	未进行小题分析	2018-10-15	2018-08-29 16:52:31
2017级高二上学期摸底考	未进行小题分析	2018-09-07	2018-08-29 16:29:30
2018级高一上学期摸底考	未进行小题分析	2018-09-07	2018-08-29 16:22:49
2016级高三开学摸底考试	未进行小题分析	2018-08-27	2018-08-29 16:51:13
2016级高二下学期期末考	语文 数学 文科数学 英语 物理 化学 生物 政治 历史 地理	2018-07-02	2017-07-21 10:57:07
2015级高三高考	语文 数学 文科数学 英语 物理 化学 生物 政治 历史 地理	2018-06-10	2017-07-21 10:21:56

图 5-13 考试分析图

在考试分析的功能中，可以按照年级的方式分析也可以按照学生考试小题进行分析。在用户选择某种分析方式之后，就可以进入到相应的页面中，如进入到“2017 级高二下学期期末考的语文”小题分析中，获取得到其相应的页面中，结果如图 5-14 所示。

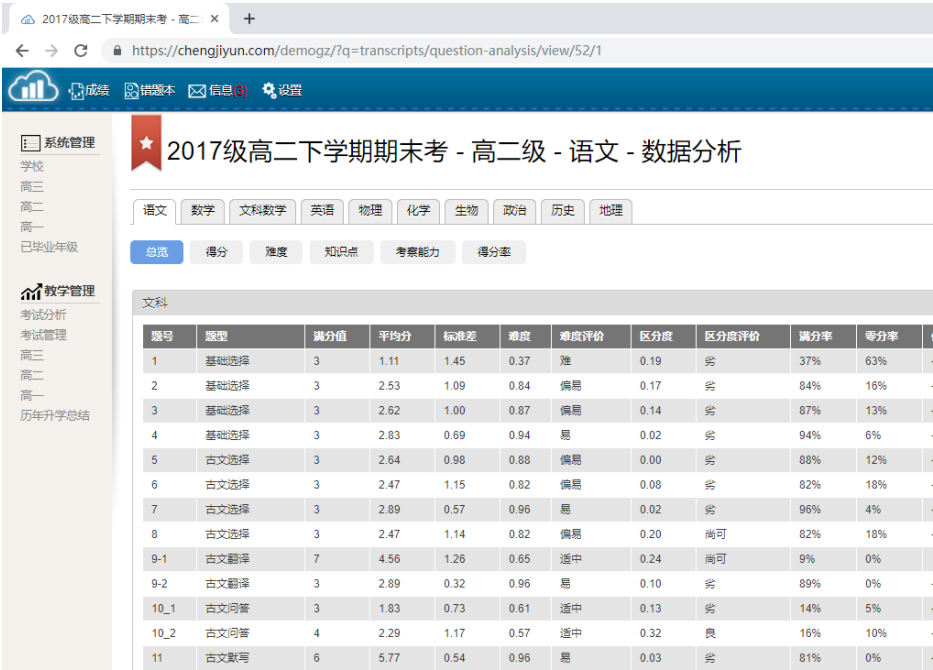


图 5-14 小题分析图

在上述的“2017 级高二下学期期末考的语文”小题分析中，可以获取得到对文科语文中的各个维度数据的分析结果数据，如得分、难度、知识点、考察能力、得分率等数据信息，将其数据汇总统计计算后的结果。对于分析结果既可以是数据结果方式显示，也可以形成一定的图表，更为直观方便，下面是得分率的结果曲线，其结果具体如图 5-15 所示。

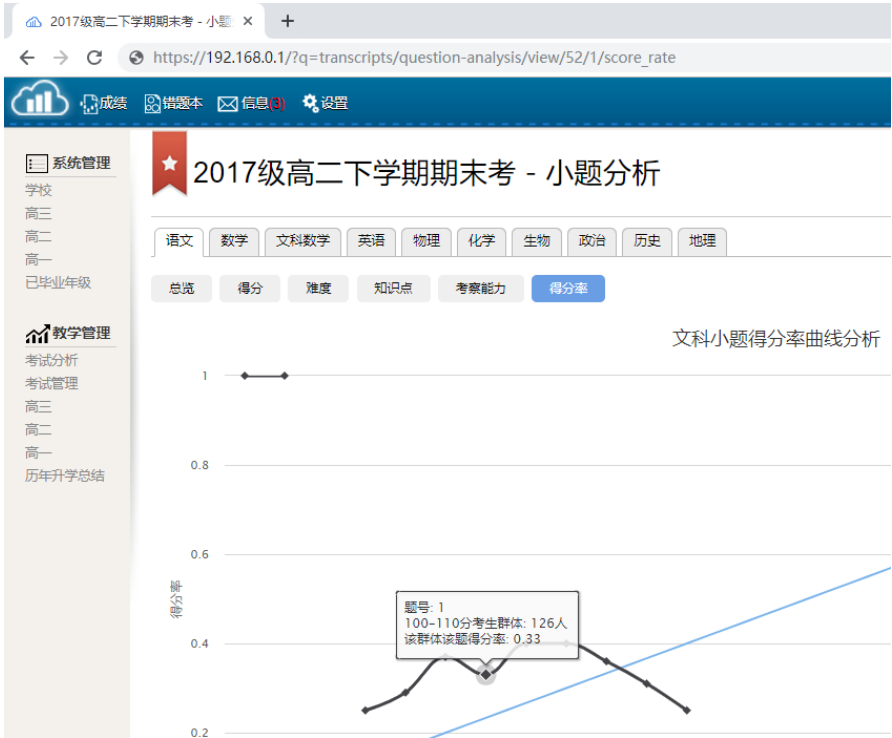


图 5-15 得分率曲线图



## 5.2.5 预测分析模块实现

在预测分析模块中，主要是对学生以往历史的数据记录进行有效的挖掘，充分利用数据挖掘技术，根据前面所建立的学生成绩预测模型，计算学生成绩预测决策树，进而可以对学生的答题情况进行合理的预测，下面是决策树计算中完成的诊断表的计算结果如图 5-16 所示。

高一（1）班学生SP诊断表						
序号	姓名	成绩	正答率	注意系数	所属区域	sp诊断结果
1	王佳莹	86	100.0	0	A	学习良好，稳定性高
2	林梓阳	75	90.0	0	A	学习良好，稳定性高
3	许灿佳	67	80.0	0.28	A	学习良好，稳定性高
4	夏烁	67	80.0	0.97	A'	成绩波动较大，稳定性差
5	卢英铎	68	80.0	0.42	A	学习良好，稳定性高
6	李德徽	74	80.0	1.39	A'	成绩波动过大，极度不稳定
7	徐旭填	67	70.0	0.14	B	稳定性高，但成绩不理想
8	涂景婕	65	60.0	0	B	稳定性高，但成绩不理想
9	王锐浩	64	60.0	0.95	B'	成绩稳定，但处于及格线上下
10	王梓霖	57	50.0	0.99	B'	虽稳定，但成绩多次不及格

图 5-16 决策树成绩分析

在上述的基于决策树的预测中，可以对部分学生的历史数据信息的挖掘，利用建立的预测模型完成的学习情况的诊断预测。也可以对某个人的考试成绩进行趋势的分析和预测。如对某个人的单次考试成绩的分析，如图 5-17 所示。



图 5-17 预测进退步变化图

根据对历史考试成绩数据的判断和计算对其考试成绩的分析，包括了学生在期末的考试成绩、学生在日常的单科的考试成绩，对于某一门的考试成绩的预测，如果出现不理想结果，可能就出现对其可能会不及格的预警。在对考试成绩的预测中，需要利用前面的决策树算法，起到了主要的作用，前面对考试数据信息的

挖掘业务中，通过对数据的分析，进行算法深度定制和优化，对学生的考试的记录进行特征提取和建模，实现对考试成绩的预测，从而发出考试预警内容，如图 5-18 所示。



图 5-18 预测结果图

通过前面的分析可以看到，将前面所涉及的学生考试预测模型和策略应用到分析建模中，通过数据集的测试，在一定程度上完成对不同类型的判断，对成绩的相关性做出分析和评估，为教学改革教学质量的提升等提供最基础的数据决策支撑，具有一定的价值。

## 5.3 数据访问的实现

### 5.3.1 数据库连接

在对系统数据库连接时候，采取的是托管 bean 的方式来完成，利用在应用程序与数据连接过程中的实体类，可以提高数据连接的可重用性和可维护性，使得数据的业务逻辑可以专注处理功能逻辑，可以实现相应的数据连接的开发目标。此外，在进行数据库连接过程中，需要提高自身的安全，设计安全可靠连接处理业务，实现数据访问和连接的安全性能。通过对客户端中数据连接限制处理，使得数据连接可以避免受到用户的恶意修改和操作，可以在一定程度上提高数据访问的安全性能。对于本系统中的数据访问连接，其实现主要是通过数据连接的 bean 类来完成，并且在这个类中对其进行数据访问连接的开启和关闭，其详细的实现代码如下。



```

public class DBUtil {
    private final static String driver = "com.mysql.jdbc.Driver";
    private final static String url = "jdbc:mysql://localhost:3306/test";
    private final static String user = "root";
    private final static String pass = "jieyi";
    static {
        try {
            Class.forName(driver);
        } catch (ClassNotFoundException e) {
            e.printStackTrace();
        }
    }
    public static Connection getConn() {
        Connection conn = null;
        try {
            conn = DriverManager.getConnection(url, user, pass);
        } catch (SQLException e) {
            // TODO Auto-generated catch block
            e.printStackTrace();
        }
        Return conn;
    }
    public static void close(Connection conn, PreparedStatement ps,
    ResultSet rs) {
        if (rs != null) {
            try {
                rs.close();
            } catch (SQLException e) {
                e.printStackTrace();
            }
        }
    }
}

```

通过上述的代码就可以完成对数据对象的持久化访问，与此同时，对于数据访问性能，在大数据量和大并发的场景业务下，需要确保数据连接的正确性，为此，可以设置相应的数据连接的最高值，一旦超过这个数据连接的值，那么将

其的请求存储到数据访问的队列中，在这个过程中，利用数据连接池的作用可以对其的数据访问进行配置和处理。

此外，在完成数据持久化访问后，可以对获取到的数据对象进行相应的操作处理，针对不同的数据对象进行一定的操作，实现良好的数据开发，实现数据访问和处理的代码重用，也可以对不同的业务功能的数据访问代码可以更好的维护，解决数据访问的故障。此外，也在一定程度上可以提高开发的效率，通过数据业务处理实体类进行数据操作的实现。

```
public static Integer executeUpdate(String dsName, String sql, String[]
params) {
    Connection conn = null;
    PreparedStatement pstmt = null;
    try {
        conn = getConnection(dsName);
        pstmt = conn.prepareStatement(sql);
        if (params != null) {
            for (int i = 0; i < params.length; i++) {
                pstmt.setString(i + 1, params[i]);
            }
        }
        return pstmt.executeUpdate();
    } catch (Exception e) {
        e.printStackTrace();
        return -1;
    } finally {
        close(conn, pstmt, null);
    }
}
```

通过上述代码，完成数据对象的操作，实现数据库的访问，并且在数据具有一定量的时候，需要大量用户的访问，需要提高数据处理的速度和效率，对数据对象建立相应的索引，可以更好的优化访问，实现数据的快速检索和处理，可以更好的完成快速访问数据的目的。

### 5.3.2 访问性能优化

数据库访问在面临大量并发用户访问的时候，需要通过不同的方式来提升数

据访问的性能，需要优化数据访问的方法，下面是本系统为解决系统数据访问采取的方法。

### 1. 采取数据库连接池的方式

在这个方式中，可以建立数据的连接池 Smartpool，在数据连接池中，对数据访问的对象和参数进行设置，在某一个的用户完成数据连接，实现数据对象访问的时候，数据的连接并没有完全的关闭，而是将其提交到数据连接池中，当其他用户或请求出现的时候，就可以利用这个数据连接对象，可以实现数据的控制连接和连接使用，这样就可以大大的降低数据访问的时间开销和资源开销，可以提高访问的处理效率。对于数据连接池来说，可以设置其中的最大连接数，当超过这个数据连接数量的时候，将会进行一定的限制和处理，使得连接失效，并且将其提交到数据连接的访问队列中，一旦出现空闲的资源，就可以调度此队列进行数据的访问处理。

此外，在数据连接池中，可以对数据对象建立索引，通过索引的方式可以对数据对象进行高效的访问，实现更好的数据表之间的连接。在数据的索引中，可以非常高效的提升数据访问的效率，降低访问的时间，如查询分组时间、排序时间等等，也是提升系统数据访问性能的非常重要的方面。

### 2. 数据缓存提升性能

在对数据访问的时候，在面临大数据量的读写的时候，需要将部分处于比较频繁访问的数据对象将其设计缓存，暂时性的存储到内存中，可以更加有利于数据的读写访问。在进行数据缓存的机制中，本系统将其设计为数据层与业务逻辑层之间，将数据缓存封装为数据访问类，可以更好的有效的查询数据对象，可以有助于数据的瞬时处理，数据的临时性处理，通过缓存机制，使得数据可以更好的访问完成读写业务。在实现缓存的技术中，在本系统的性能提升实现中，采取的是 opensymphony 组件，在这个基于 javaEE 的框架技术中完成访问数据的缓存。

### 3. 系统分布式部署

在这个机制中，将数据存储的服务器分布在不同的区域中，使得数据访问并不集中于某一个的服务器，使得在大规模的并发访问的时候，可以分流用户的数据流量。在数据存储的时候，可以将数据存储在多个的存储服务器中，一方面可以将数据读写的业务进行分担分流，另外一方面也可以提高系统的可用性，一旦其中的服务器出现故障，其余的服务器的系统数据也是完整的正常的，可以使得数据具有更好的安全性能。

## 5.4 本章小结

本章主要工作是介绍系统实现。首先是搭建系统开发的环境和运行平台，按照开发要求完成对主要功能的代码开发，页面设计等工作，对原始的数据进行统计分析，应用决策树算法到学生成绩数据中，发现学生成绩数据中的有价值的内容，完成预测分析业务功能，实现了学生成绩分析与预警系统的核心内容。接着是数据访问的实现。

## 第6章 学生成绩分析与预警系统的测试

在系统测试中主要包括了系统的功能测试和性能测试两个方面。在系统的功能测试中主要的工作是验证系统中的主要业务功能是否已经完成需求分析中的设立的功能目标，是否已经达到各项指标是功能测试的重点内容。

### 6.1 测试环境

功能测试，其通常的做法是采取黑盒测试的方式，就是利用系统测试的时候不考虑内部各个业务结构情况下，对在业务处理中所出现的各个页面与功能进行测试，类似于在完全封闭的盒子中的，对盒子内部结构和情况完全不知情下进行测试和检验。

性能测试，其通常采取的测试手段、方法是利用负载测试工具 LoadRunner，用过 LoadRunner 这个优秀的测试软件，可以非常快捷的模拟出一定较好的测试用户和数据，通过模拟对象对目标系统的访问和发出请求，对发出的数据和对象进行后台的监控，获取到相应的测试数据，对测试数据的汇总分析，监测目标系统的性能表现，并且在这个基础上对目标系统的性能做出分析。

系统测试首先是需要搭建一定测试环境，在搭建好的测试环境下进行各个项目和类别的测试，下面是本次系统测试的环境部署，具体如图 6-1 所示。

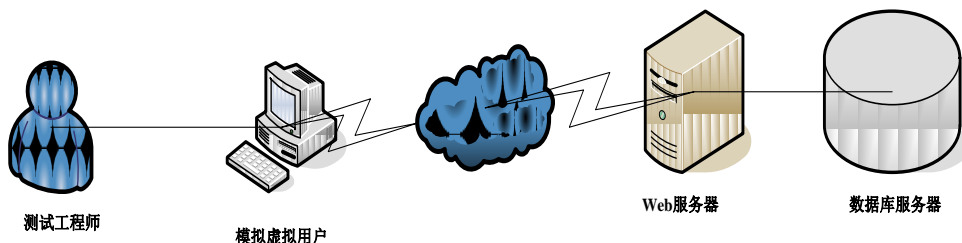


图 6-1 环境部署图

### 6.2 功能测试

系统在进行测试时候需要严格按照所涉及的用例进行模拟数据的测试，按照功能测试用例完成对表单中、页面中等各个测试对象进行数据输入和处理，系统对所输入的数据对象接收和运行，验证所开发出的系统在业务功能是否与预期的一致，是否是按照设计需求进行开发的，因此，在进行功能测试时候采取的严格的黑盒测试的方法和手段，完成部分的业务功能的测试。

在本系统的功能测试中，由于业务功能较多，功能较为复杂，下面仅仅对部分的业务功能进行测试，并且对测试结果进行说明和分析。

### (1) 成绩分析功能测试

成绩分析功能测试用例主要完成的是对是否可以按照用户的要求进行各类数据要求的汇总分析，如对某次考试的优良率、及格率、标准分分析，对其中的数据能够按照要求完成图表信息的显示，如形成饼图、折线图、雷达图等比较直观的结果。成绩分析测试用例具体如表 6-2 所示。

表 6-2 成绩分析功能测试用例表

序号	测试用例数据输入	期望结果	测试结果
1	用户进入到成绩分析页面，选择“2017 级高二下学期期末考一小题分析”	进入到页面中，显示结果	通过
2	选择“政治”，进入到“难度”分析选项卡，获取到难度分析结果	难度分析，文科各班得分数据分析，并且点击格子进入班级详细分析	通过
3	分析结果保存，点击“下载 Excel 表格”“保存图片”	进入到下载页面，将结果保存到本地	通过

按照测试用例，详细的测试结果如图 6-3 到图 6-5 所示。



图 6-3 成绩分析功能页面图

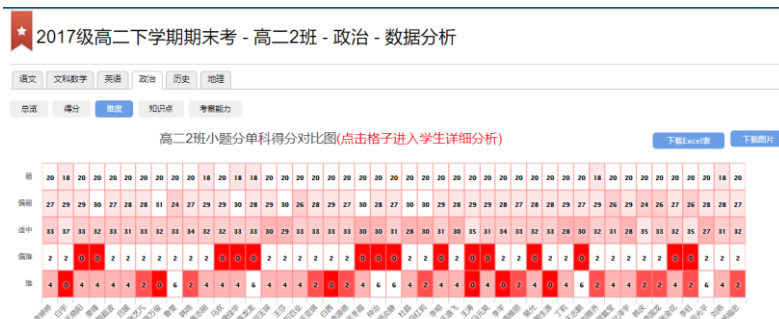


图 6-4 成绩分析班级分析功能页面图

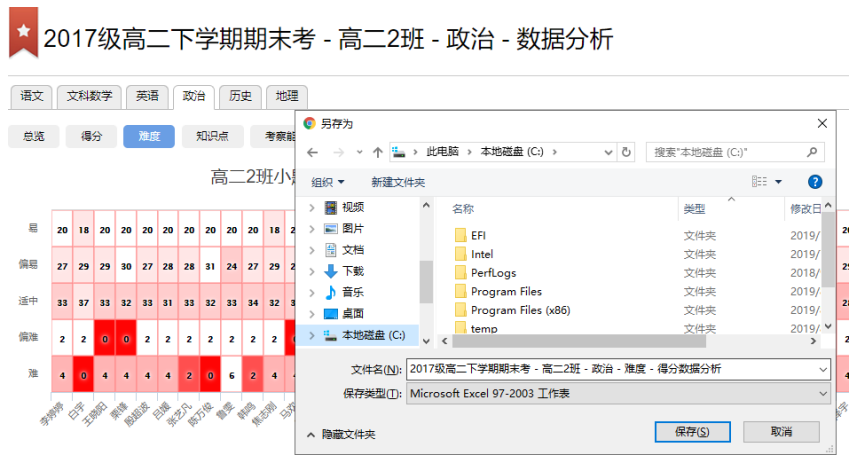


图 6-5 成绩分析结果下载页面图

(2) 决策树功能测试

决策树功能测试用例主要完成的是学生成绩分析，可以建立属于决策树中的信息增益等参数信息，可以对目标数据的成绩挖掘，并且处理挖掘和获取得到分析结果等。测试用例具体如表 6-6 所示。

表 6-6 决策树功能测试用例表

序号	测试用例数据输入	期望结果	测试结果
1	原始的数据可以导入到系统中，可以导入决策树 C4.5 算法	成绩数据的深度挖掘和分析	通过
2	可以在按照要求进行学生成绩分析	结果显示在列表中	通过

按照测试用例，详细的测试结果如图 6-7 所示。

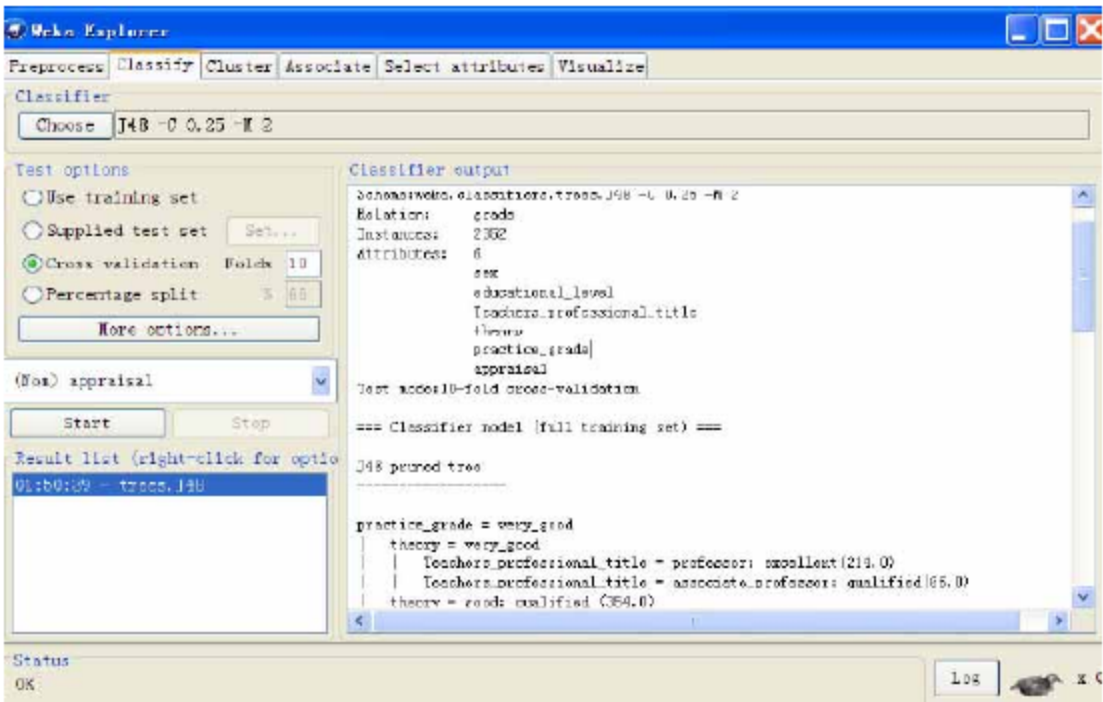


图 6-7 决策树算法运算图

为了验证本算法，在上述搭建的实验环境下，在 Weka 工具中进行三个年级数据集的对比测试，其结果如图 6-8 所示。

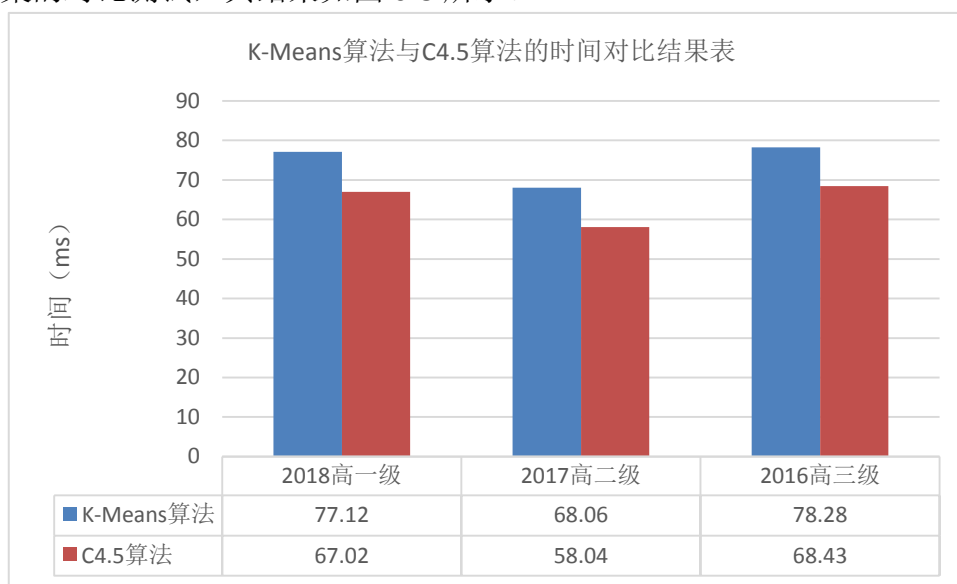


图 6-8 不同年级样本数下两个算法的时间花费对比图

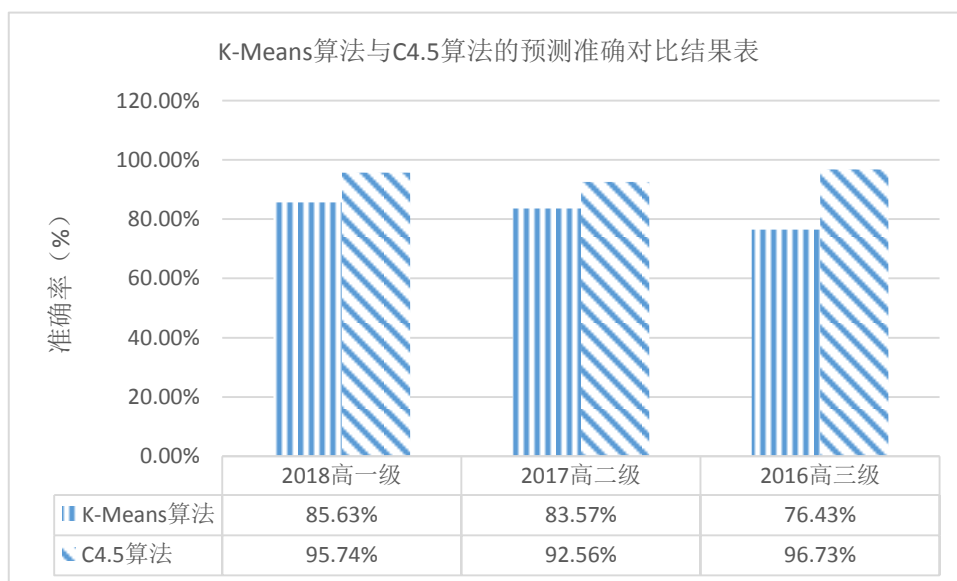


图 6-9 不同年级样本数下两个算法的预测准确率对比图

通过上述的不同算法测试，对数据挖掘的性能测试，可以看到，本文的 C4.5 算法具有较好的处理时间和预测准确率。

通过对前面的功能测试，按照系统测试中的黑盒测试手段和方法，对目标系统的主要业务功能进行针对性的测试。根据制定的测试的方法，获取测试的数据和结果，可以发现系统可能存在的业务功能问题，并且对业务功能进行分析，本系统的开发设计是符合需求的，开发完全达到预期目的。



### 6.3 性能测试

对于系统的性能指标在前面的系统需求分析已经明确, 根据需求分析内容按照测试规范要求, 通过负载测试工具 LoadRunner 对目标系统进行性能测试, 主要是模拟一定的用户数据, 访问目标系统在一定的时间范围内, 监测系统的各项指标数据内容, 并且按照测试的顺序进行逐项的测试。

在本次的性能测试中, 在测试环境下, 在客户端中安装部署 LoadRunner, 并且设置测试参数数据信息, 在一定的时间间隔内设置为 2 秒钟就启动 1 个的虚拟用户, 在极短时间范围内启动完 200 左右的虚拟用户模拟访问目标系统, 得到如下的测试结果如图 6-10 所示。

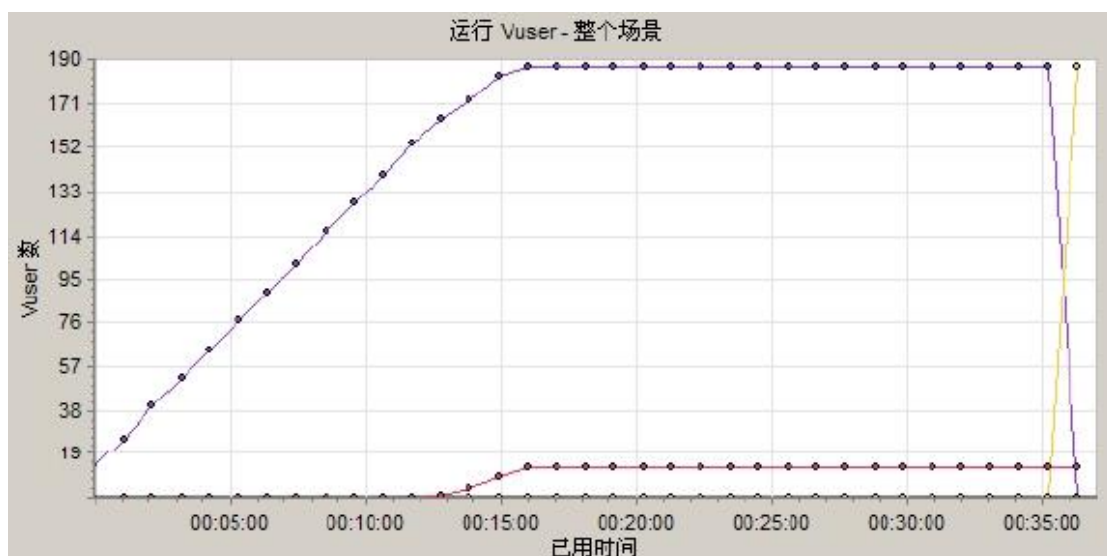


图 6-10 模拟 200 用户访问系统

下面是虚拟用户模拟访问目标系统的结果图, 可以看到, 在用户数增加到 200 个左右时候, 系统出现了错误数情况, 其结果如图 6-11 所示。

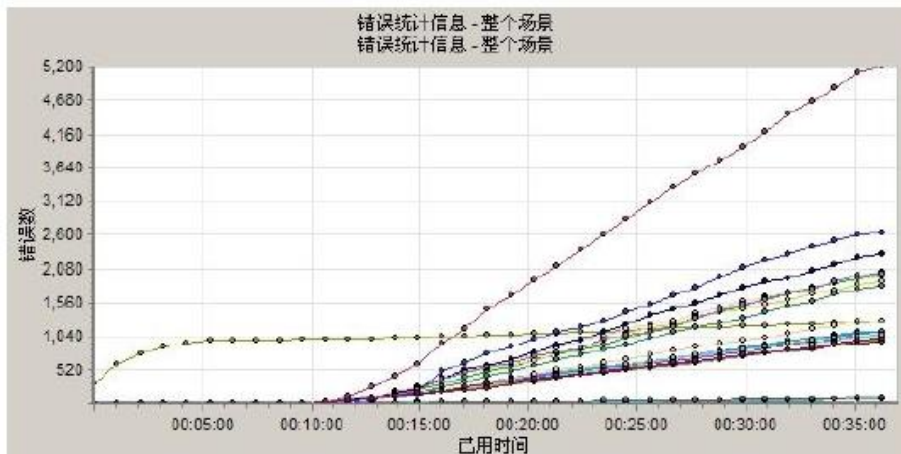


图 6-11 增加用户与出现错误结果图

通过前面的测试，根据 LoadRunner 工具测试结果，得到系统的负载结果如图 6-12 所示。

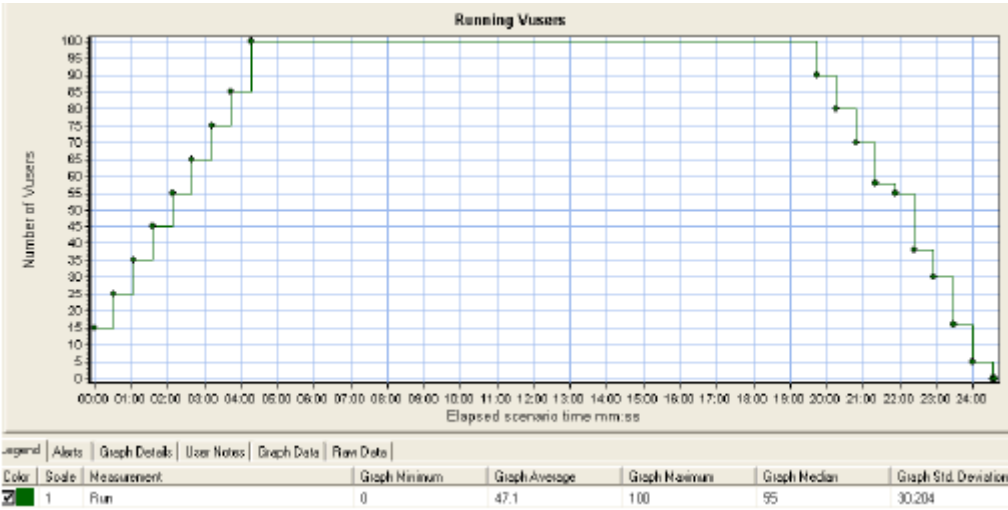


图 6-12 负载测试图

在上述的负载测试中，可以看到，在设置一定规则情况，利用 LoadRunner 工具对目标进行用例测试，主要是登录测试，可以看到在 200 用户数访问情况，系统基本运行正常，相对运行稳定，下面是每秒点击率-吞吐量的测试结果图，如图 6-13 所示。

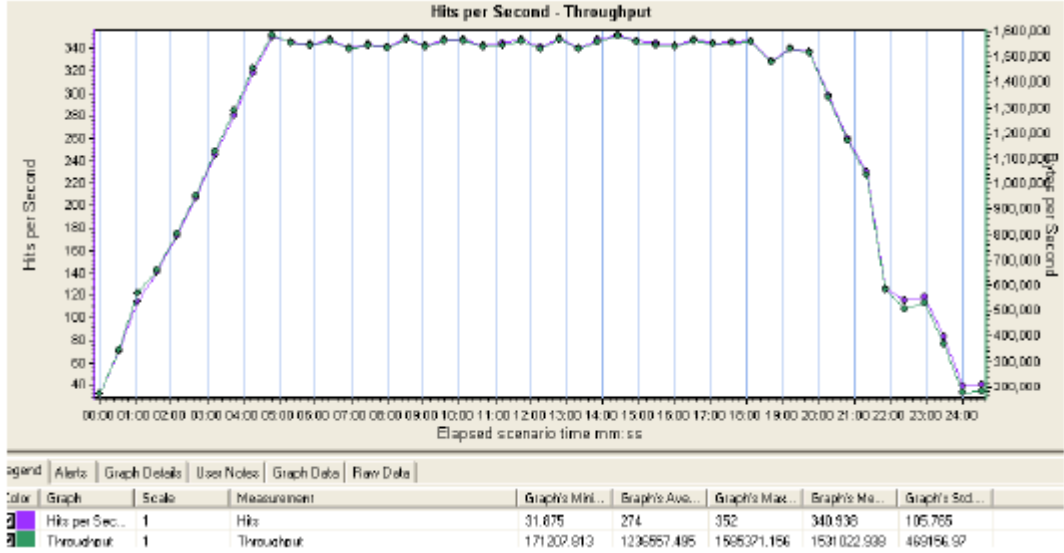


图 6-13 每秒点击率-吞吐量分析

在上图中，可以看到测试指标的值随着测试进行出现变化，并且随着点击数量不断增加，如果其中的点击的数量减少的话，那么其中的吞吐量也随着降低，可以说是系统可以满足用户的请求，对于用户所发送的请求能够及时处理和响应。

下面是测试的 CPU 的利用率的结果图，可以看到 CPU 的利用率可以说处于一个比较高的水平，具体即如果如图 6-14 所示。

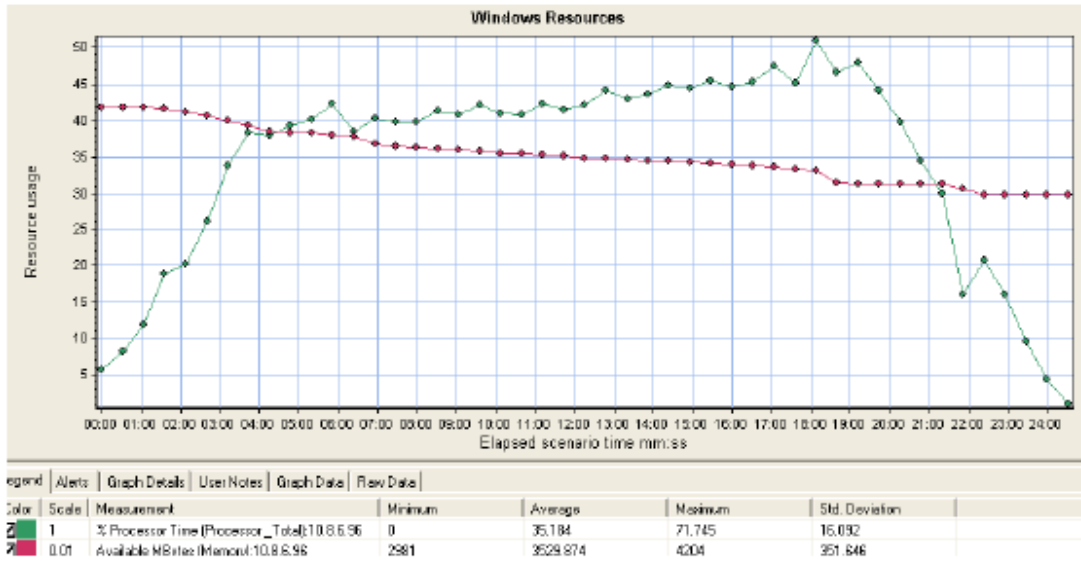


图 6-14 cpu 利用率结果图

通过前面的测试可以看到，整个系统的运行在性能方面是符合要求的，是可以按照用户设计需求运行，在各个测试指标方面表现良好，能够满足用户业务要求，通过性能测试验证是可行的，运行稳定。

6.4 本章小结

本章是系统测试，主要阐述了本系统的业务功能测试和性能测试。利用测试的结果和数据，对系统进行改进和优化，验证系统是符合要求的，是达到目的。

## 第7章总结与展望

### 7.1 总结

随着人工智能、大数据、互联网+等信息技术迅速发展,学校在日常教学活动中会产生相当多的基础数据,如学生成绩、学生作业、学生考勤等数据,利用上述数据并且经过一定挖掘处理,可以有效的分析学生的学习行为,挖掘数据之间的价值,并且能够从多个角度和方面展现学生的学习活动,利用大数据的挖掘技术实现学生成绩分析与预警。因此,开发应用信息平台十分必要。主要研究工作总结如下:

(1)学生成绩分析与预警系统需求分析。先分析各个方面的具体需求内容,然后对系统在运行过程中需要的系统运行参数、配置信息进行分析,得到非功能性的需求,经过分析完成学生成绩分析与预警系统具有系统管理、学籍管理、成绩管理、考试分析和预测分析等业务功能的需求分。

(2)学生成绩分析与预警系统的设计。按照需求分析,依据所完成的用户的需求进行相应的详细的设计,将系统划分为三个逻辑结构层次完成体系架构设计,接着是系统功能架构划分,将系统分为系统管理、学籍管理、成绩管理、考试分析和预测分析等不同业务功能模块,针对上述业务功能进行详细设计包括了实体类、时序图和活动图等统一建模设计,最后是数据库设计,按照系统对数据需求建模处理,对数据对象实体联系分析与建模,得到数据表结构。

(3)学生成绩分析与预警系统的实现。部署安装高效环境,开发系统界面,完成各个功能模块的开发。通过在 IDE 工具下完成系统主要业务模块服务的开放工作,在本系统中实现了主要的业务功能模块,包括了部分代码和界面实现,接着是数据库访问的实现,介绍数据库连接的实现和数据存储的实现。

(4)学生成绩分析与预警系统的测试。在完成上述的业务功能后需要对所开发得到的系统进行测试,搭建测试环境,设计测试用例对目标系统进行业务功能测试和性能测试,检验系统开发各项成果是否已经满足系统设计要求,根据测试结果对测试数据进行分析,验证系统已经完成既定要求的设计与开发工作。

### 7.2 展望

经过前面的分析和设计,完成了学生成绩分析与预警系统主要业务功能,可以通过其实现对学生成绩数据的挖掘分析,能够实现对成绩数据的有效分析,不过由于个人的原因,在时间精力上有不足的地方,挖掘分析上还处于探索阶段,

需要对此加以不断完善和改进，需要在未来的工作对此进行不断研究，不断分析不断取得新的成果。未来的工作可以从如下几个方面开展：

（1）对数据分析的力度上还需要加强，目前的数据仅仅是一个学校的基础数据，需要加入更多的数据，使得分析具有更有代表性。对于所采集的数据需要进一步的提高其质量，对部分的异常数据需要手工进行处理。

（2）目前的数据挖掘功能还有待于完善和开发。需要不断挖掘数据之间的内在价值，提升挖掘的准确性。

（3）采取更为先进的人工智能机器学习算法对成绩数据的分析和预测，对学生的分析不仅仅局限于成绩，可以进一步的延伸到其他的学习的行为上。

## 参考文献

- [1] 罗越, 陈国柱, 梅书宇. 对学生成绩进行数据挖掘应用分析[J]. 计算机工程, 2019, 50(09):266.
- [2] 刘博鹏, 樊铁成, 杨红. 基于数据挖掘技术的学生成绩预警应用研究[J]. 四川大学学报(自然科学版), 2019, 56(02):267-272.
- [3] 张贵元. 基于数据挖掘聚类算法的学生成绩分析[J]. 电脑知识与技术, 2019, 15(09):1-2.
- [4] 单耀, 王艺岚, 张家华. 基于学生成绩多变量数据挖掘的学习行为分析[J]. 计算机系统应用, 2019, 23(06):197-198.
- [5] 狄晓娇. 关联分析在学生成绩数据挖掘的应用[J]. 电脑知识与技术, 2018, 14(34):246-265.
- [6] 程代娣. Clementine 数据挖掘工具在计算机等级考试成绩中的应用[J]. 齐鲁工业大学学报, 2017, 31(06):52-56.
- [7] 陈硕. 基于数据挖掘技术的学生成绩分析系统的设计与实现[J]. 科技经济导刊, 2017, (34):22-23.
- [8] 李大志, 杜丽英. 基于数据挖掘的学生成绩预测的应用浅析[J]. 黑龙江科技信息, 2017, (07):156-156.
- [9] 黄青. 数据挖掘技术在学生成绩管理中的应用研究[J]. 电子世界, 2016, (23):38-38.
- [10] 万家华. 数据挖掘技术在学生成绩分析中的应用研究[J]. 哈尔滨师范大学自然科学学报, 2016, 32(06):38-41.
- [11] 栾红波, 文福安. 数据挖掘在大学英语成绩预测中的应用研究[J]. 软件, 2016, 37(03):67-69.
- [12] 刘爱萍. 基于关联规则数据挖掘技术的高校学生学习成绩分析[J]. 信息与电脑(理论版), 2015(22):83-85.
- [13] 彭好佑, 廖敬, 姚坚, 黄勇. 基于数据挖掘技术在成绩管理系统的应用研究[J]. 福建电脑, 2015, 31(03):97-98.
- [14] 傅则恒. 数据挖掘决策树技术在学生成绩分析中的应用研究[J]. 广东技术师范学院学报, 2015, 36(02):113-117.
- [15] 费建刚, 梁建国. 数据挖掘技术在学生成绩分析中的应用[J]. 电脑知识与技术, 2018, 19(24):5391-5393.
- [16] C. L. Philip Chen, etl. Data-intensive applications, challenges, tech

- niques and technologies: A survey on Big Data[J]. Information Sciences: An International Journal, 2014, 275:314-347.
- [17] Chen M S, Han J, Yu P S. Data Mining: An Overview from a Database Perspective[J]. IEEE Transactions on Knowledge and Data Engineering, 1996, 8(6):866-883.
- [18] Usama Fayyad, Gregory Piatetsky-Shapiro, Padhraic Smyth. The KDD process for extracting useful knowledge from volumes of data[J]. Communications of the ACM, 1996, 39(11):27-34.
- [19] Bhardwaj K, Pal S. Data Mining: A prediction for performance improvement using classification[J]. arXiv preprint, arXiv:1201.3418, 2012
- [20] Dorina Kabakchieva. Predicting Student Performance by Using Data Mining Methods for Classification[J]. Cybernetics and Information Technologies, 2013, 13(1):119-125
- [21] Shaobo Huang, Ning Fang. Predicting student academic performance in an engineering dynamics course: A comparison of four types of predictive mathematical models[J]. Computers & education, 2013, 61(Feb.):133-145
- [22] Raheela Asif, Agathe Merceron, Mahmood K. Pathan. Predicting Student Academic Performance at Degree Level: A Case Study[J]. International Journal of Intelligent Systems and Applications(IJISA), 2015, 7(1):118-124
- [23] 杜婧敏, 方海光, 李维杨, 全赛赛. 教育大数据研究综述[J]. 中国教育信息化, 2016, (19):89-93
- [24] 魏顺平. 学习分析技术:挖掘大数据时代下教育数据的价值[J]. 现代教育技术, 2013, 23(2):5-11.
- [25] 叶福兰. 基于数据挖掘的高校学生成绩预警状况分析[J]. 长春师范大学学报, 2013(10):45-49.
- [26] 刘斌, 陈依潼. 数据挖掘技术在学生成绩分析中的应用[J]. 电脑编程技巧与维护, 2014(16):45-46..
- [27] 王志娟, 班娅萌, 平金珍. 基于 AJAX 技术和 JAVAEE 的分页查询优化[J]. 信息通信, 2019(01):118-119.
- [28] 查嵩炜. JavaEE 环境下集群系统建设研究与应用[J]. 邮政研究, 2019, 35(02):10-13.
- [29] 王文明. 基于 JavaEE 架构的 SSH 框架[J]. 电子技术与软件工程, 2017(05):

- 69.
- [30] 傅莞龙, 张传武, 彭安金. 使用 Spring Data 和 JPA 在 JavaEE 系统中简化持久层[J]. 电子世界, 2017(06):88-89.
- [31] 黄贤英. 在 J2EE 平台上使用 EJB 开发应用. 计算机工程与应用, 2016, 27(12): 163-164
- [32] 乔岚. 基于 MyBatis 和 Spring 的 JavaEE 数据持久层的研究与应用[J]. 信息与电脑(理论版), 2017(08):73-76.
- [33] 李敏. UML 及 Struts2+Hibernate3 在 JavaEE 项目中的应用[J]. 福建师大福清分校学报, 2016(02):12-17.
- [34] Riaz Ahamed. Comprehensive Performance Analysis and Special Issues of Broadband Access Technologies for High Speed Data Communication[J]. Computer Sciences and Telecommunications . 2016: 21-63
- [35] Fan Liu, Xiaoxiao Shu, Yongchao Luo. Design of Power Marketing System Based on JavaEE[C]. Proceedings of the 2017 7th International Conference on Education, Management, Computer and Society (EMCS 2017), 2017.
- [36] un Yang. Design on Textbook Subscription System Based on JavaEE [C]. Proceedings of the 4th International Conference on Mechatronics, Materials, Chemistry and Computer Engineering, 2015.
- [37] 宦臣, 陈军, 沈群. 基于 JavaEE 架构的 SSH 框架的研究[J]. 计算机光盘软件与应用, 2014, 17(17):119-120.
- [38] 刘志鹏, 卫晨. Spring Data 在 JavaEE 系统中的应用与研究[J]. 电子世界, 2014, (16):278.
- [39] 赵学作. Docker 部署 SQL Server 数据库[J]. 网络安全和信息化, 2019, (07): 80-81.
- [40] 王珣, 左良利. 关系数据库支持的不确定时间序列存储[J/OL]. 计算机技术与发展, 2019, (11):1-6
- [41] 付熙徐, 龚希章. 一种轻量级的服务端防 SQL 注入攻击方法[J]. 盐城工学院学报(自然科学版), 2019, 32(02):28-32.
- [42] 仝瑞钦. 基于 SQL Server 的数据安全性维护策略研究[J]. 电脑编程技巧与维护, 2019(06):172-173.
- [43] 庄淼. SQL Server 数据库安全影响因素及优化设计措施[J]. 电脑知识与技术, 2019, 15(17):1-2.
- [44] ]闫守军, 巩凡, 岳翔宇. 网络设计中关系数据库技术的应用探讨[J]. 计算



- 机产品与流通, 2019(07):49.
- [45] 张俊萍, 邓怡, 刘天璇. 基于决策树与扩展相关性矩阵的故障诊断方法[J]. 测控技术, 2019(07):125-129.
- [46] 陈珊. 数据挖掘经典算法分析[J]. 电子技术与软件工程, 2019(15):128-129.
- [47] 李华, 张井玲, 刘婷婷. 大数据时代下数据挖掘技术的应用研究[J]. 现代信息科技, 2019(13):132-133.
- [48] 吕建驰. 机器学习算法在数据挖掘中的应用[J]. 电子世界, 2019,(13):62-63.
- [49] 迟殿委. 一种改进的决策树 ID3 算法的应用[J]. 现代计算机, 2019,(17):43-45.
- [50] 黄勇, 魏乐. 一种针对不均衡数据集的 SVM 决策树算法[J]. 成都信息工程大学学报, 2019, 34(03):274-277.
- [51] 崔羽飞, 魏进武, 潘思宇, 张溶芳, 王笑, 贾子翔. 一种基于决策树终端换机间隔的研究[J]. 电信科学, 2019, 35(07):165-172.
- [52] 崔懿心. 基于机器学习算法的社交数据挖掘与用户偏好的建模[J]. 电子技术与软件工程, 2019(14):174-175.
- [53] 杨瑞, 王萍, 索瑞霞, 银艳艳. 数据挖掘技术在 CRM 中的应用研究——基于决策树算法[J]. 中国管理信息化, 2019, 22(15):53-55.
- [54] 王志俊. 数据挖掘、OLAP 在决策支持系统中的应用机制探究[J]. 通讯世界, 2019, 26(07):190-191.
- [55] 罗秋瑾. 基于模糊度的决策树生成算法[J]. 云南民族大学学报(自然科学版), 2019, 28(03):285-288.
- [56] 韩存鸽, 叶球孙. 决策树分类算法中 C4.5 算法的研究与改进[J]. 计算机系统应用, 2019, 28(06):198-202.
- [57] 王诚, 王凯. 一种基于聚类约简决策树的改进随机森林算法[J]. 南京邮电大学学报(自然科学版), 2019, 39(03):91-97.
- [58] 李宏林. 应用朴素贝叶斯算法预测学生学习指标的改进研究[J]. 情报探索, 2015. 32(06): 22-26.

## 致谢

经过前段时期的紧张的准备，我的硕士论文的工作接近尾声，可以说是基本完成，在这个时期内，我需要感谢的人有很多，首要的是我的导师，无论在哪个时间点，我的老师都是全力支持我，对于老师的帮助，我心中无限感激和深深的佩服。我的论文的写作和完善是在指导老师的不懈帮助下完成的，可以说总开始到技术，我的老师都没有停止的指导，付出了很多的劳动，对于老师的这种无私奉献的精神，我自己也深受感动，激励我以后的工作生活中向我的老师学习。

我的论文的完成还需要感谢我的家人朋友，是大家在这段时间从各个方面配合我、支持我，给与我很多的鼓励，使得我可以全身心的完成论文的修改工作。