长沙师范学院 2024 届本科毕业论文（设计）开题报告

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 课题题目 | 基于 Excel 的在线分析处理系统设计与实现 | | | | |
| 学生姓名 | 周泉 | 学号 | 2020542340244 | 专业班级 | 20物联网工程402班 |
| 指导教师 | 陈耀东 | 职称 | 副教授 | 学历 | 博士研究生 |
| 1. **文献综述** 2. **研究背景**   随着信息技术的迅速发展，数据已经成为当今社会不可或缺的资源。企业、政府机构和组织等各个领域产生的大量数据，数据如果能够被有效地收集、分析和利用，将为决策提供重要支持，推动业务的发展和创新。  在现代企业和组织中， Excel 表格作为一种通用的数据管理和分析工具被广泛使用，被广泛应用于企业、学术界、科研机构等各个领域。Microsoft 公司推出的 Excel，被当今世界上普遍认为是功能强、技术新、使用方便的优秀电子表格软件，是一种“表格”形式的数据综合处理与分析系统**[1]**。它的数据组织格式同时支持结构化的数据存储，使得用户能够清晰地组织和管理大量数据，结构化的数据存储方式便于进行各种数据分析和处理操作。Excel 软件提供了直观、易用的界面，使得用户可以轻松地进行数据录入、编辑和分析。同时，Excel 还自带丰富的公式和函数，提供了强大的计算和数据处理能力，满足用户对数据分析的灵活需求。  基于这些优点，大多数企业、机构和个人都使用 Excel 进行数据管理和分析。同时各行业许多专业课程和培训课程都涵盖了 Excel 的使用，使得用户更容易接触和掌握这一工具。Excel 还可以轻松与其他 Microsoft Office 套件中的软件（如 Word、PowerPoint）进行集成，使得用户在不同的任务中可以方便地使用Excel进行数据处理和分析。  由于 Excel 的广泛应用，各种操作系统和办公软件都提供了对Excel文件的良好兼容性，使得 Excel 文件的分享和交流更为便捷，反过来提高了 Excel 的普及程度。然而，随着数据量和复杂性的不断增加，仅仅依赖传统的 Excel 工作表往往无法满足对数据深层次分析的需求。  **（1）国内研究现状**  国内各企业广泛采用商业智能（ Business Intelligence, 简称 BI）平台进行数据分析，其中包括立方、亿信华辰、魔镜、帆软等多个知名平台。这些平台在可视化分析、数据挖掘和报表功能方面提供了丰富的解决方案，然而在Excel表格分析智能化方面存在改进空间。一些大型企业推出了强大的数据处理和分析软件，例如腾讯云的数据分析工具和阿里云的数据计算服务，这些工具主要专注于大数据领域，相对较少关注 Excel 表格这一常见的数据源。在学术界，北京大学深圳研究生院/信息工程学院袁粒老师团队创建的ChatExcel是一款较为新兴的 Excel 分析工具，尽管仍处于刚起步阶段（发布于2023年三月）。  国内 BI 市场调研：  **帆软：**于2022年12月正式发布 FineBI 6.0新版本，着重提高产品的易用性和团队协作能力。这一更新将用户体验推向新高度，同时强调团队协作在数据分析中的关键作用。  **永洪科技：**上线永洪 BI V10.0 新版本，注重模版化的报告和组件开发能力。新增报告演示模式、数据分析和 3D 模型能力，进一步强化了企业级管理。  **百度：**在渠道生态建设上进行了大力投入，成功提高了来自渠道伙伴的收入。这表明百度致力于通过建设强大的渠道生态系统来推广其BI 解决方案，以更好地服务广大用户。  **浪潮通软：**专注于政企行业客户，具备强大的实施和定制开发能力。通过预制的几百个分析模型，满足不同行业、不同层级客户的多样化需求。  **思迈特软件：** Smartbi v10.5新版本新增和优化了多个功能，包括数据模型、自助 ETL 、交互仪表盘、电子表格等。引入了“指标管理”和“对话式分析”两大功能，提高了产品的易用性、美观性、安全性和智能性。  **亿信华辰：**深耕政府、能源和金融行业，为用户提供一体化解决方案，包括数据治理、大数据平台、数据仓库和商业智能等。这显示出对不同行业复杂数据需求的专注。  这些发展呈现出国内BI 市场不断演进的趋势，企业通过不断更新产品和追求更高水平的商业智能服务，努力满足用户不断增长的数据分析需求。  IDC（International Data Corporation 国际数据公司）数据显示，2022下半年中国商业智能软件市场规模为5.1亿美元，由于疫情线下产品交付与验收受到严重影响，以及人民币对美元汇率下降等原因，同比增速仅为7.6%，是过去五年来的最低水平。2022全年中国商业智能软件市场规模达到8.8亿美元，同比增长13.8%，远低于2022上半年的预测数据。IDC 预计，从2023年开始，企业 IT 投资将逐步恢复，将在2024年回归正常水平。根据 IDC 最新一期的预测数据，到2027年，中国商业智能软件市场规模将达到19.7亿美元，未来5年市场年复合增长率（CAGR）为17.5%。  国内 BI 厂商在2023年产品重点投入在易用性提升和大模型能力融合。各厂商结合自身产品特点，在数据填报、数据准备、数据可视化和高级分析等功能模块，通过可视化配置、AI智能辅助或页面优化等方式，进一步降低了产品的使用难度，提高产品交互体验。同时，由于2023年AI 大模型的热潮，基于技术的发展和用户侧的推动，各 BI 厂商也在积极探索大模型技术与 BI 产品融合的功能和场景，通过与第三方大模型集成的方式与重点客户共同实践大模型在数据分析场景下的行业应用。    （图1 中国 BI 软件市场预测）  **（2）国外研究现状**  国外许多商业智能软件公司致力于创新 BI 工具，例如 Tableau、 Microsoft Power BI、QlikView 等。它们提供了丰富的可视化和分析功能，但在结构化数据如Excel表格的处理上，仍有一些挑战需要解决。国外一些研究团队致力于自动化数据清理和预测分析的研究，以提高数据分析的效率和准确性。这为专业 Excel 表格分析软件的发展提供了一些启示。  总而言之，目前国内外对于专业 Excel 表格分析的软件的研究相对较为有限。大多数研究集中在大数据处理、商业智能平台等领域，对于 Excel 表格这一常见但具有挑战性的数据源，缺乏专注的研究。在人工智能和深度学习技术不断发展的背景下，未来可能会有更多的研究关注如何将这些技术应用于 Excel 表格分析，提高智能化水平和用户体验。  2020年国内外12款自助式BI工具最全测评  （图2 国内外 BI 软件优缺点对比）   1. **应用背景**   在企业和组织的日常运营中，会产生大量的数据，为了得到准确可靠的数据，会使用不同的数据处理软件对数据进行分析处理，而在对数据进行数值修约和可疑数据的取舍处理时，必须遵守相应标准规范的约定，这样处理得到的检测结果才具有代表性**[2]**。在处理数据时 Excel 操作直观简单，能够大大提高数据的处理效率和准确率。然而，随着业务数据的不断增长和复杂化，传统的 Excel 表格分析面临一些挑战：  （1）数据量庞大：随着业务的扩张， Excel 表格中的数据量庞大，手动分析变得耗时且容易出错。  （2）复杂的数据结构：Excel表格中可能包含多个工作表、复杂的公式和多层次的数据结构，增加了分析的复杂度。  （3）实时性要求：企业需要及时了解最新的业务数据，但传统的 Excel 分析往往无法满足实时性的要求。  （4）数据质量不稳定： Excel 表格中常常存在数据质量不稳定的问题，包括缺失值、异常数据等，需要进行数据清理和处理。  随着人工智能和数据科学领域的不断发展，利用先进的技术解决 Excel 表格分析面临的挑战成为刻不容缓的任务。同时，市场对于智能化、高效性和用户友好的 BI 工具的需求也在不断增加。在此背景下，我们决定开发一款专业 Excel 表格分析的智能 BI 软件，以满足企业和组织对于更先进、更智能的数据分析工具的迫切需求。该软件将充分利用人工智能、机器学习等技术，以及实时数据处理和高级可视化工具，为用户提供更为便捷、直观、准确的数据分析体验。  **3、关键技术简介**  （1）人工智能技术  人工智能（Artificial Intelligence，简称 AI）是指通过计算机系统模拟、复制和执行人类智能的能力**[3]**。它是基于计算机科学的分支，是现代社会中的新兴学科，主要包括机器学习、知识图谱、语言理解以及问答系统等，将机器赋予人的思维方式、方法以及技能等关键要素，从而替代传统人工思考、工作，独立设计规划等智能行为**[4]**。  如同蒸汽时代的蒸汽机、电气时代的发电机、信息时代的计算机和互联网，人工智能正成为推动人类进入智能时代的决定性力量**[5]**。由美国OpenAI 公司开发的 ChatGPT 属于“生成式 AI ”，与传统的“分析式 AI”不同，生成式人工智能更加灵活和智能化，不仅能理解和解释数据的含义，还能够根据输入的数据和资料，进行预测并输出全新文本**[6]**。人工智能内容生成技术（AIGC）利用技术赋予机器部分人类思维能力,以生成有价值内容。智慧性是 AIGC 产生有价值内容的基础，具备人类思维与智慧交互能力是其发展趋势**[7]**。AI 可以用于智能解析 Excel 数据，理解数据的结构和关系。通过自然语言处理（NLP）和模式识别技术，系统可以更智能地理解表格的内容，再生成出易于用户理解和整理的数据。  （2）Spring Boot 框架  Spring Boot 是一个用于快速开发基于Spring框架的 Java 应用程序的开源框架**[8]**。它简化了Spring应用的搭建过程，提供了一种约定大于配置的方式，使得开发者可以更专注于业务逻辑而不必过多关心配置细节。  使用Spring Boot作为项目后端框架，可以极大地简化应用的开发和维护过程。它提供了丰富的功能和模块，支持构建各种类型的应用，从传统的单体应用到分布式的微服务架构。Spring Boot的约定大于配置的设计理念和丰富的生态系统使得开发者能够更加高效地进行后端开发，集中精力于业务逻辑的实现。  （3）Vue 框架  Vue.js 是一款轻量级的JavaScript框架，用于构建用户界面**[9]**。它的核心库专注于视图层，易于集成到其他库或现有项目中。 Vue 的设计灵感来自 Angular 和 React ，但它更加轻量、灵活，使得前端开发更加简单、直观。  使用Vue框架作为前端的开发工具，可以使得前端开发更加简单、灵活且高效。其清晰的设计思想、强大的生态系统以及响应式数据绑定等特性，使得开发者能够更容易地构建出交互性强、用户体验良好的前端应用。Vue 的组件化和模块化开发方式，使得团队能够更好地协作，项目更易于维护和扩展。  **二、选题依据、主要研究内容、研究思路及方案**   1. **选题依据**   在信息化、数据化繁荣发展的时代，信息网络的数据量激增，同时存在类似于智能终端代写文章等垃圾数据，这些数据充斥在信息网络中，用户获得数据的有效性极大地降低，获取有用信息的时间成本增加**[10]**。但在另一方面，人工智能又提升了数据收集管理能力和数据挖掘利用水平，用户通过端口将数据提供给人工智能服务系统，交由人工智能从庞大的数据集中精准找到目标数据，并对其进行深度挖掘分析，从中挖掘出有利用价值的信息**[11]**。由人工智能引领的新一轮科技革命和产业变革方兴未艾。人工智能呈现深度学习、跨界融合、人机协同、群智开放、自主操控等新特征，正在对各行业产生重大而深远的影响**[12]**。  对于用户而言，大量的Excel表格需要手动进行数据分析和解析，耗时且容易出错，尤其在大数据量的情况下，传统的分析方式效率低下；同时传统的数据分析工具在可视化方面可能受限，难以提供高度定制化的可视化图表，通过引入专业的可视化工具和图表库，结合AI生成的数据，提供更丰富、直观的数据可视化效果；部分 Excel 表格中常常包含有缺失值、异常值等问题，手动清理繁琐且容易遗漏。我们可以借助AI模块，自动进行数据清理，识别和处理表格中的异常数据，提高数据的质量和准确性。除此之外，传统数据分析工具可能存在繁琐的操作和学习曲线，限制了用户在数据分析中的工作效率。通过自动化和智能化的设计，提高整体工作效率，让用户更专注于数据本身的理解和业务决策。  综上所诉，此课题具备应有的市场需求性以及技术可行性。   1. **主要研究内容**   （1）研究所需技术与应用需求调研  为了实现整个系统，首先需要学习主流的人工智能技术，包括但不限于深度学习、机器学习等。同时，掌握用于搭建系统的相关技术，例如 Spring Boot 框架和前端开发技术。思考这些技术在整个系统实现中的运用，尤其关注它们在创新性的 Excel 数据处理应用方面的潜在用途。  （2）研究基于AIGC的数据分析平台框架  使用Spring Boot 框架快速搭建出数据分析平台的整体框架是实现智能的图表分析能力的基础。在此平台中创建用户模块，用于存储和管理用户信息。创建业务处理模块，通过调用第三方 AI API，处理用户上传的 Excel 图表，并满足用户的分析需求。同时，提供清晰的接口供用户使用，实现系统的高度可扩展性和灵活性。  （3）设计并实现易于用户使用的交互性软件  使用自主搭建的数据分析平台提供的接口，设计出直观而清晰的图形化界面，偏于用户上传具体的Excel图表和分析诉求。通过渲染 AIGC 的内容，用户可以获得直观而详细的图表以及分析结论。该交互软件旨在简化用户的操作流程，使用户能够轻松理解和利用 AIGC 的强大功能，从而提高用户体验。  **3、研究方法**  （1）调查法：通过互联网对市场大多数产品进行实际体验，分析优缺点以及用户体验，整理数据用于研究。  （2）模块实现法:从整体拆分整个项目，划分出细致模块，实现高内聚低耦合的编程思路。  （3）文献研究法：通过调查文献来获得资料，从而全面地、正确地了解掌握所要研究问题。  **4、研究思路及方案设计**  研究思路：  本课题使用前后端分离的思想，利用 Java Spring Boot 技术快速搭建出数据分析平台，在平台的业务处理层中调用AI接口实现主要功能,同时用 MySQL来存储用户数据（例如账号、密码、生成的信息等）。设计出简洁便利的可交互性软件，调用数据分析平台的接口处理具体业务，渲染出易于用户处理的图表信息，提供优质的 Excel 表格的分析和处理功能。    （图3 系统结构图）  预留模块：（1）用户管理模块：存储用户账号信息 （2）数据存储模块：存储用户上传的图表及分析诉求 （3）图表处理模块：具体业务实现，即处理用户上传的图表 （4）可视化交互模块：便于用户使用的图形化界面   |  |  | | --- | --- | | 第一章  绪论 | * 1. 选题背景与意义   2. 研究现状      1. 国外研究现状      2. 国内研究现状   3. 论文主要内容与章节安排 | | 第二章  相关技术概述 | 2.1 关键技术  2.1.1 人工智能技术  2.1.2 Spring Boot 框架  2.1.3 Vue 框架  2.2 小结：相关技术发展前景和主要问题 | | 第三章  系统设计 | 3.1 需求分析  3.1.1 功能需求分析  3.1.2 性能需求分析  3.2 系统框架设计  3.2.1 基于AIGC的数据分析平台系统框架设计  3.2.2 用户可视化交互软件设计  3.3 系统模块设计  3.3.1 用户管理模块设计  3.3.2 数据存储模块设计  3.3.3 图表处理模块设计  3.4 小结 | | 第四章  系统实现 | 4.1 实现技术概述  4.2 功能实现  4.2.1 用户管理模块实现  4.2.2 数据存储模块实现  4.2.3 图表处理模块实现  4.2.4 用户可视化界面模块实现  4.3 小结 | | 第五章  总结与展望 | 5.1 总结  5.2 展望 |   **三、工作进度及具体安排**  2023.10.10 — 2023.10.31 毕业设计选题  2023.11.01 — 2023.11.6 文献检索、需求调研  2023.11.06 — 2023.11.28 撰写开题报告与开题报告答辩PPT  2023.11.28 — 2023.11.28进行开题报告答辩  2023.11.29 — 2024.03.09 实施设计、撰写毕业设计报告初稿  2024.03.09 — 2024.05.04 验证设计、按指导修改毕业设计报告  2024.05.04 — 2024.05.20 完成毕业设计报告，审查修改  2024.05.21 — 2024.06.12完成毕业设计答辩、材料归档  **四、进度安排**  2023.10.21-2023.11.18  命题、选题、下达任务书  2023.11.19-2023.12.03  撰写开题报告、完成开题  2023.12.03-2024.03.09  开始设计、毕业设计报告写作，完成中期检查  2024.03.10-2024.05.07  完成设计方案、完成设计报告终稿、送审查重  2024.05.08-2024.06.14  完成毕业设计答辩、材料归档  **五、参考文献**  [1]殷留义, 李咏波, 李林, 刘金凤, 马莹, 勾世谈, 王筠, 薛莹莹, 徐永燕. Excel和Aspen Plus软件在《化工热力学》汽液平衡数据计算中的应用[J]. 广东化工, 2023, 50 (17): 168-170.  [2]黄美景.Excel在试验检测数据处理中的应用[J].大众标准化,2023,(12):137-139.  [3]费一鸣.基于人工智能技术的大数据分析方法研究进展[J].电子测试,2022,36(06):  [4]高纯. 关于人工智能在计算机信息技术中的标准化应用研究[J]. 大众标准化, 2022, (23): 177-179.  [5] 谭铁牛.人工智能的历史、现状和未来[J].求是,2019-02-18,(04).  [6]郑世林, 陶然, 杨文博. ChatGPT等生成式人工智能技术对产业转型升级的影响[J]. 产业经济评论.  [7]刘逸伦;黄微;张晓君;李贺南. AIGC 赋能的科技情报智能服务：特征、场景与框架[J/OL].现代情报,1-12[2023-11-25].  [8]崔臣;宋甲旭.基于Spring Boot的校园二手交易系统研究[J].无线互联科技,2023,20(18):31-34.  [9]李国印;杨军;宁玉富;孙浩.基于Spring-Boot和Vue框架在HE染色病理图像中浸润性肺腺癌细胞初步诊断筛选系统的研究与设计[J].长江信息通信,2023,36(10):14-16.  [10]徐攀.信息网络中大数据与人工智能技术的应用分析[C]//航空工业测控技术发展中心;中国航空学会测试技术分会.第十九届中国航空测控技术年会论文集.中国飞行试验研究院;,2022:3.DOI:10.26914/c.cnkihy.2022.053233  [11]林伟. 人工智能数据安全风险及应对[J]. 情报杂志, 2022, 41 (10): 105-111+88.  [12]周立宏. 基于AIGC的芒果XR虚拟制作实践[J]. 影视制作, 2023, 29 (07): 33-37. | | | | | |
| 指导教师意见：    指导教师：  年 月 日 | | | | | |
| 开题专家组意见  专家组组长签名：    年 月 日 | | | | | |