

# 云南大学本科毕业论文（设计） 开题报告书

题 目：	天气数据预测及可视化系统设计与实现
姓 名：	朱鹏宇
学 院：	软件学院
专 业：	软件工程
指导教师（职称）：	康雁 副教授
填写时间：	2024年1月3日

选题来源	教师科研课题
<p data-bbox="240 293 1353 360">一、选题依据</p> <p data-bbox="240 360 1353 1032">我国小农生产有几千年的历史，“大国小农”是我们的基本国情农情，小规模家庭经营是农业的本源性制度。人均一亩三分地、户均不过十亩田的小农生产方式,是我国农业发展需要长期面对的现实。发展多种形式适度规模经营，培育新型农业经营主体，是建设现代农业的前进方向和必由之路。需要注重发挥新型农业经营主体带动作用，培育各类专业化市场化服务组织，提升小农生产经营组织化程度，改善小农户生产设施条件，提升小农户抗风险能力，扶持小农户拓展增收空间，把小农生产引入现代农业发展轨道。</p> <p data-bbox="240 1032 1353 2040">随着我国农业信息化进程的快速推进，农业领域积累了大量的统计数据。如何进一步有效整合和利用现有的天气数据，挖掘数据潜在价值，成为农业信息化研究面临的重要课题。利用数据可视化技术可以将数据变换为可识别的图形符号、图像视频等信息，将不可见或难以直观显示的数据信息展现出来，为人们洞察事物潜在规律提供新的观察手段。当前的天气数据资源分散，展现方式较为单一，大多使用简单的可视化图表，缺少交互手段，用户无法在已有可视图表的基础上进行自定义交互可视分析。本文针对现有天气数据的特点和数据展示方式单一的问题，利用网络爬虫技术实现了对不同来源数据的采集，后端使用是 springboot 项目，前端主要是用的 vuejs 和 echartsjs 等技术设计实现了一个基于 B/S 架构的天气数据可视化系统。</p>	

基于天气数据的特点，以及其在数据可视化方面的需求，结合目前的计算机技术，进行 Web 可视化系统的设计与实现，收集整理的天气数据并进行可视化展示,对促进天气数据价值的挖掘与利用具有重要意义。

## 二、研究（设计）的学术或现实意义、主要内容、创新点

### 1.学术意义：

数据可视化通过借助人的视觉感知和大脑的认知辨识能力，将不可见或难以直接显示的数据转化为可感知的图形、符号、颜色、纹理等，向目标对象传递数据信息，帮助其快速高效地获取数据潜在的价值。用户可以通过对可视化图表的感知，使用可视化交互工具进行数据分析，获取知识，并进一步提升为智慧。将数据可视化方法运用于天气数据资源中，对于挖掘农业生产潜力、提高农业生产效率和促进农业可持续发展具有重要作用。

### 2.现实意义：

目前，我国资源分散，且根据时空特征将各项统计指标分别保存到不同类别数据中心，数据结构多样，天气数据展现方式单一，缺少针对图表的交互可视分析。因此，基于天气数据的特点，以及其在数据可视化方面的需求，结合目前的计算机技术，进行 Web 可视化系统的设计与实现，收集整理多源异构的天气数据并进行可视化展示，对促进天气数据价值的挖掘与利用具有重要意义。

### 3.主要内容:

(1)天气数据采集。通过研究现有网络爬虫技术的实现方式,针对互联网上的天气数据资源分散且结构不一致的特点,基于通用网络爬虫技术和框架,设计定向的网络爬虫,实现对数据资源的抓取。

(2)数据资源处理。使用随机森林,lstm 等数据模型预测时空序列数据技术基于不同数据集的特点和数据可视化需求,系统实现了对原始采集数据的清洗处理和转换整合,为数据的可视化展示提供了数据资源。

(3)可视化系统设计与实现。基于 Springboot 网络框架,设计并实现了天气数据的 Web 可视化系统。系统对采集处理后的数据资源进行了陈列展示,并提供了资源下载接口。系统利用 Tableau 可视化工具和 Echarts 库,实现了数据的图表可视化,并为图表提供动态交互功能。

### 4.创新点:

通过对多源异构的天气数据的收集处理,并基于处理的数据资源设计实现了一个能提供数据图表的 Web 可视化系统,系统实现了数据的多样化展示和图表的动态交互。

### 5.研究现状

可视化是指用于创建图形、图像或动画,以便交流沟通信息的任何技术和方法(Michael Friendly,2008)[1]。数据可视化研究起源于 20 世纪 50 年代,计算机的产生使得人类以全新的方式绘制图表,从而产生了计算机图形学[2]。“可视化”或它的全称“科学计算可视

化”(Visualization in Scientific Computing)一词是在 1987 年正式被提出,其根据是美国国家科学基金会召开的“科学计算可视化研讨会”的内容报告[3]。在短短 30 余年间,科学计算可视化发展成为一个十分活跃的研究领域,新的研究分支不断涌现,如出现了用以表现海量数据不同类型及其逻辑关系的信息可视化技术,以及将可视化与分析相结合的可视分析学研究方向[4]。计算机技术的快速发展和数据资源规模的增长,传统的数据可视化技术难以应对海量、高维、多源和动态数据的分析挑战,可视分析由此兴起[5]。可视分析通过将可视化和数据预测分析方法相结合,利用交互可视界面来分析、推理和决策,在增强可视化效果的同时为大规模数据集的分析利用提供了解决方案。现在又有了把“科学计算可视化”、“信息可视化”和“可视分析学”这三个分支整合在一起的新学科“数据可视化”。可视化的主要内容包括科学可视化(Scientific Visualization)、信息可视化(Information Visualization)以及可视分析(Visual Analytics)。数据可视化是可视化研究领域的新起点,极大地促进了学科交叉和整合[6-12]。

许多国内企业在数据可视化的商业应用方面也取得了显著成就,开发了许多成熟稳定的可视化产品或工具,为大范围的人们提供了更多体验数据可视化的机会[13]。百度可视化团队在公司业务需求基础上,开发了兼具可视化展示和交互功能的 Echarts 可视化库,并将其贡献到了开源社区。奇虎 360 公司依托于 12 亿终端设备产生的大数据进行分析,推出了大数据平台的可视化产品“骗子地图”

和 360 星图[14]。大数据魔镜通过大数据收集、整理和分析，为企业提供了商业数据可视化的解决方案[15]。阿里巴巴通过对其旗下的电子商务交易平台上产生的商业数据进行分析与可视化展示，为用户提供信息共享服务，同时提供基于云平台的可视化服务[16]。帆软公司研发的 FineBI 可视分析工具能帮助企业快速搭建面向全员的大数据分析平台，让每一个用户充分了解和利用他们的数据，自由释放数据潜能[17]。

在农业数据可视化方面，数据可视化技术不仅仅集中在天气领域主要集中在三维建模、模拟生长等科学可视化领域，在信息可视化领域的研究和应用还比较少[18-20]。岳延滨、朱艳等以油菜角果为研究对象，研究了油菜角果的三维建模技术，构建了基于形态特征参数的角果几何模型。张小斌等利用全国蓟马分布普查数据，构建了中国农业蓟马害虫分析可视化系统，为开展农业蓟马害虫分布与防控研究提供了技术支持[21]。伍艳莲、曹卫星研究了小麦的生理构造及其与群体之间的相互关系，实现了小麦从气管到群体的多层数据可视化，为构建可视化小麦生长系统奠定了技术基础[22]。朱孟帅根据农产品价格数据的特点，构建了基于时空维度的农产品在线可视化系统[23]。

### 三、研究（设计）的技术路线及预期目标

#### 1. 技术路线：

- (1) 数据采集，解决了大规模天气数据集的快速爬取问题，完成了千万级数据的采集需求。
- (2) 采集数据的清洗处理。在数据获取后使用正则表达式等筛选方法对数据进行格式化处理。
- (3) 可视化系统的设计与实现。基于 Springboot 框架以及 Html、CSS、JavaScript 等技术实现了 B/S 架构的农业统计数据可视化平台，使用 Echarts 可视化工具对收集的天气数据进行了可视化展示。

## 2. 预期目标：

通过对多源异构的天气数据进行采集处理，并利用采集处理的数据资源实现一个能提供多样数据展示效果和数据共享服务的 Web 可视化系统。

## 四、工作进度安排

### 时间 项目进度

2023.10.01— 2024.01.03 选定课题后熟悉课题，并进行项目的需求分析，查阅资料，完成开题报告并提交

2024.01.04— 2024.02.01 课题相关知识的学习掌握，明确实现过程中使用的方法与工具，为后期的实现过程奠定基础

2024.02.11— 2024.03.10 研究随机森林、LSTM 并对模型进行改进与优化，并初步完成优化模型的实现。

2024.03.11— 2024.03.25 利用优化模型对具有时空特性的天气

数据进行预测，对比其他传统模型。

2024.03.26— 2024.04.10 填写中期检查表，提交初期的项目进行验收

2024.04.11— 2024.04.20 进行毕业论文的撰写

2024.04.20— 2024.04.30 对课题论文进行检查和后期修正、完善，完成毕业论文的定稿

## 五、完成研究（设计）所需条件及落实措施

硬件配置：Processor: Intel(R) Core(TM) i5-7300HQ CPU 2.50GHZ;

Installed memory(RAM): 8.00 GB;

System type: 64-bit Operating System;

开发环境： Python3,Java8,JetBrains

IDEA,nodejs,npm,echartsjs,vuejs;

运行环境： Windows 10 及其以上版本，

所有软件安装和环境配置都已完成。

## 六、参考文献资料

[1] 周洪斌, 陈立平, 刘连浩. 基于 ECharts 的数据可视化应用 [J].

沙洲职业工学院学报,

2021, 24(1):7.DOI:10.3969/j.issn.1009-8429.2021.01.001.

[2] 王竟俨, 侯彦东. 基于 PyEcharts 的第三次农业普查数据可视

化分析 [J]. 河南科学,



2021, 039(004):517-523.

[3] 张晨祥, 宁璐, 冯明歌. 高校就业数据可视化管理系统设计与实现 [J]. 电脑编程技巧与维护, 2022(9):74-77.

[4] 李艳, 丁国强, 张庆. 网络招聘数据可视化系统的设计与实现 [J]. 信息与电脑, 2021, 033(001):112-115.

[5] 吴雅琴, 王晓东. 基于数据可视化血库管理系统的设计与实现 [J]. 内蒙古科技与经济, 2021(5):3.

[6] 杨国丹. 基于 VizQL 的汽车工业数据可视化系统的设计与实现 [J]. 现代计算机, 2023, 29(11):99-103.

[7] 文雪巍, 邢婷, 李鹏, 等. 基于网络爬虫疫情数据分析及可视化系统的设计与实现 [J]. 黑龙江工程学院学报, 2022, 36(5):6.

[8] 赵晨. 基于 HTML5 和 Web Service 的 WebGIS 开发与应用 [D]. 山西农业大学 [2024-05-16].

[9] 王冉冉, 李寒晓, 齐凯, 等. 基于” 双碳目标” 下养猪场多能互补系统设计与实现 [J]. 智能化农业装备学报 (中英文),

2023(4):26-32.DOI:10.12398/j.issn.2096-

7217.2023.04.004.

[10] 钟佳伶, 黎茂锋, 黄俊, 等. 气动数据可视化系统的设计与实现 [J]. 计算机测量与控

制, 2021.DOI:10.16526/j.cnki.11-4762/tp.2021.02.031.

[11] 罗行行, 杨茂强, 王清欢, 等. 新冠疫情数据可视化系统的设计与实现 [J]. 福建电脑,

2022, 38(4):3.

[12] 程妮, 尚宇波. 基于 JavaScript 的疫情数据可视化系统的设计与实现 [J]. 运城学院

学报, 2022, 40(6):6-9.

[13] 张建立, 杨栋, 王传刚, 等. 基于聚类的煤质数据可视化系统的设计实现 [J]. 电子元

器件与信息技术, 2021(011):005.

[14] 王蒙, 郭建磊, 董蕾. 长征数据三维可视化分析平台设计与实现 [J]. 软件, 2023,

44(7):131-133.

[15] 何龙祥, 葛继成, 王轻, 等. 基于数据仓库的医疗数据可视化系统的设计与实现 [J].

软件, 2021.DOI:10.3969/j.issn.1003-6970.2021.01.030.

[16] 褚建萍. 基于 Vue 的数据可视化系统研究 [J]. 电子技术与软件工程, 2022(18):234-

237.

[17] Jayasiri K C N , Thathsarani W R V K , De S D I ,et al.Design and Implementation of an Automated Hospital Management System with MERN Stack[J].International Journal of Engineering and Management Research, 2022.DOI:10.31033/ijemr.12.5.24.

[18] Han X , Liu J , Tan B ,et al.Design and Implementation of Smart Ocean Visualization System Based on Extended Reality Technology[J].Journal of web engineering, 2021(2):20.

[19] Arpaia P , Benedetto E D , Duraccio L .Design, implementation, and metrological characterization of a wearable, integrated AR-BCI hands-free system for health 4.0 monitoring[J].Measurement, 2021, 177(3):109280.DOI:10.1016/j.measurement.2021.109280.

[20] Jornet Monteverde J A , Soler Llorens J L .Design and Implementation of a Wireless Sensor Network for Seismic Monitoring of Buildings[J]. 2021.

[21] Liu Y , Guo J , Hou L ,et al.Design and Implementation of a Comprehensive Training Program for Software Engineering Talents’ Practical Ability[J]. 计算机教育, 2022(12).

[22] He Y , Zhang Y , Zhang Y ,et al.Design and Implementation of

Real-time Power Grid

WebGIS Visualization Framework Based on New Generation

Dispatching and Control

System[J].Journal of Physics: Conference Series, 2021,

2087(1):012073-.

[23] 方匡南, 吴见彬, 朱建平, 等. 随机森林方法研究综述 [J]. 统计与信息论坛, 2011,

26(3):7.DOI:10.3969/j.issn.1007-3116.2011.03.006.

[24] 马骊. 随机森林算法的优化改进研究 [D]. 暨南大学,2016.

[25] 李欣海. 随机森林模型在分类与回归分析中的应用 [J]. 应用昆虫学报, 2013,

50(4):8.DOI:10.7679/j.issn.2095-1353.2013.163.

[26] 王佳蕊, 孙勇, 孟祥东, 等. 一种基于 GCN-LSTM 的电力系统短期负荷预测方

法.CN202211388691.8[2024-05-16].

[27] 徐杨, 刘亚新, 汪涛, 等. 基于 LSTM 的三峡水库短期上游水位预测方法研究 [J]. 水

利水电快报, 2022, 43(10):6.

[28] 刘亚新, 樊启祥, 尚毅梓, 等. 基于 LSTM 神经网络的水电站短期水位预测方法 [J].

水利水电科技进展, 2019,

39(2):6.DOI:CNKI:SUN:SLSD.0.2019-02-013.

[29] 倪汉杰, 蒋仲廉, 初秀民, 等. 基于 DWT-LSTM 的航道水位智能预测模型研究 [J].

中国航海, 2021(044-002).

[30] 齐慧. 基于 KNN-LSTM 模型的常武地区地下水位时空预测方法研究 [J].[2024-05-

16].

#### 七、指导教师意见（需明确是否同意开题）

利用数据可视化技术可以将不可见或难以直观显示的数据信息展现出来,为人们洞察事物潜在规律提供新的观察手段。论文准备利用网络爬虫技术实现了对不同来源数据的采集,后端准备使用springboot项目,前端准备使用的vuejs和echartsjs等技术,技术路线合理可行。建议考虑相关预测算法实现天气预测。设计实现一个基于B/S架构的天气统计数据预测可视化系统具有一定的应用价值,同意开题。

指导教师: 康雁      2024年1月3日

#### 八、开题审核综合意见（需明确“同意开题”、“不同意开题”或“修改后开题”）

通过

审核负责人: 谢诚      2024年1月5日