

本科毕业设计（论文）

银发群体高温多时间尺度预警 和服务优化可视化研究

学 院： 计算机科学与技术学院

专 业： 计算机科学与技术

姓 名： 刘航宇

学 号：

指导教师：

2026 年 5 月 28 日

摘要

随着全球气候变暖,极端高温事件频发且强度持续增加,对公共卫生构成日益严峻的挑战。老年群体(65岁及以上)因体温调节功能减退、慢性病患率高以及社会隔离等因素,是高温热浪最脆弱的群体之一。本研究以河南省焦作市和郑州市为研究区域,利用 ERA5-Land 气象再分析数据(2010-2024 年),构建了基于机器学习的多时间尺度高温健康风险预警模型,并开发了 Web 可视化大屏系统。

本研究主要工作包括:(1)通过 Copernicus Climate Data Store (CDS) API 获取 2010-2024 年焦作、郑州两市的 ERA5-Land 网格气象数据,采用 Magnus 公式计算相对湿度、NOAA Rothfusz 公式计算体感温度,构建了包含 19 个气象衍生特征的完整数据集;(2)通过滑动窗口方法(窗口 14 天)生成监督学习样本,构建了包含 1,095,758 条样本的多时间尺度预测数据集,覆盖短期(3 天)、中期(7 天)和长期(30 天)三个预测窗口;(3)设计了 983,628 参数的 LSTM-Attention 深度学习模型,采用双向 LSTM 提取时序特征、4 头自注意力机制捕捉关键时间步,并以 Focal Loss 缓解类别不平衡;(4)以 XGBoost 作为基线模型,在 164,365 条测试样本上进行了系统对比实验;(5)基于 Flask 和 ECharts 开发了深色科技蓝风格的 Web 可视化大屏,包含温度趋势、风险实时展示、人口饼图、预警时间线、暴露-反应曲线和历史回顾六个功能面板。

实验结果表明,XGBoost 模型在三个时间尺度上均取得优异性能:短期(3 天) F1-Macro 达 0.9325、中期(7 天)达 0.9195、长期(30 天)达 0.8576。LSTM-Attention 模型(F1=0.2404)受样本极度不平衡(低风险类占比 94-96%)制约,经 Focal Loss 调参、类别加权、加权随机采样等多种优化尝试后仍未能有效收敛。该对比揭示了梯度提升树模型在表格型时序预测任务中相对深度序列模型的优势。本研究构建的可视化大屏系统为面向银发群体的高温健康防护提供了直观的决策支持工具。

关键词: 高温热浪; 银发群体; 多时间尺度预警; XGBoost; LSTM-Attention; 体感温度; 可视化

Abstract

Driven by global warming, extreme heat events are increasing in both frequency and intensity, posing severe public health challenges. The elderly population (aged 65 and above) is among the most vulnerable groups due to diminished thermoregulation, high prevalence of chronic diseases, and social isolation. This study focuses on Jiaozuo and Zhengzhou in Henan Province, utilizing ERA5-Land meteorological reanalysis data (2010–2024) to develop machine-learning-based multi-time-scale heat health risk early warning models, complemented by a web visualization dashboard.

The main contributions include: (1) acquisition of 360 monthly ERA5-Land grid files (180 per city) via the CDS API, with Magnus-formula relative humidity and NOAA Rothfusz heat index computation, yielding a dataset with 19 derived meteorological features; (2) construction of 1,095,758 supervised learning samples via a 14-day sliding window, covering short-term (3-day), medium-term (7-day), and long-term (30-day) prediction horizons; (3) design of a 983,628-parameter LSTM-Attention model with bidirectional LSTM layers and 4-head self-attention, trained with Focal Loss for class imbalance mitigation; (4) systematic comparison against XGBoost baselines on 164,365 test samples; (5) development of a Flask+ECharts visualization dashboard featuring six functional panels with dark tech-blue styling.

Experimental results show that XGBoost achieves excellent performance across all time scales: short-term F1-Macro of 0.9325, medium-term 0.9195, and long-term 0.8576. The LSTM-Attention model ($F1=0.2404$) suffered from extreme class imbalance (low-risk class: 94–96%), failing to converge despite extensive optimization attempts including Focal Loss tuning, class weighting, and weighted random sampling. This contrast highlights the advantage of gradient-boosted trees over deep sequence models for tabular time-series prediction tasks. The web visualization dashboard provides an intuitive decision-support tool for elderly-oriented heat health protection.

Keywords: Heatwave; Elderly Population; Multi-time-scale Early Warning; XGBoost; LSTM-Attention; Heat Index; Visualization

目录

摘要	2
Abstract	3
第 1 章 绪论	7
1.1 研究背景	7
1.2 银发群体的高温脆弱性	7
1.3 研究区域选择	7
1.4 国内外研究现状	8
1.4.1 高温-健康关联的流行病学研究	8
1.4.2 高温健康预警系统 (HHWS)	8
1.4.3 深度学习在环境健康预测中的应用	8
1.5 研究目的与意义	8
1.6 研究内容与技术路线	9
1.7 论文组织结构	9
第 2 章 相关理论与技术基础	10
2.1 LSTM 神经网络	10
2.1.1 LSTM 单元结构	10
2.1.2 双向 LSTM	10
2.2 注意力机制	11
2.2.1 缩放点积注意力	11
2.2.2 多头自注意力	11
2.3 XGBoost 算法	11
2.4 体感温度计算方法	12
2.4.1 Magnus 公式——相对湿度计算	12
2.4.2 NOAA Rothfusz 公式——体感温度	12
2.5 高温健康风险等级划分	13
2.6 Focal Loss 损失函数	13
2.7 Flask 框架与 ECharts 可视化	13
第 3 章 数据获取与预处理	14
3.1 研究区域与数据时空范围	14
3.2 数据来源	14
3.2.1 ERA5-Land 气象再分析数据	14

3.2.2	人口与健康数据	15
3.3	数据预处理管线	15
3.3.1	步骤 1: 多文件加载与拼接	15
3.3.2	步骤 2: 日聚合与单位转换	15
3.3.3	步骤 3: 相对湿度计算	15
3.3.4	步骤 4: 体感温度计算	15
3.3.5	步骤 5: 特征工程	15
3.3.6	步骤 6: 风险标签计算	16
3.3.7	步骤 7: 滑动窗口序列构建	16
3.3.8	步骤 8: 数据集保存与合并	17
3.4	类别分布与数据不平衡	17
3.5	数据集划分	17
第 4 章	多时间尺度预警模型设计	18
4.1	模型总体架构	18
4.2	输入投影层	18
4.3	双向 LSTM 编码层	18
4.3.1	层结构与参数	18
4.3.2	LSTM 投影层	19
4.4	多头自注意力层	19
4.4.1	层配置	19
4.4.2	残差连接与层归一化	19
4.4.3	全局平均池化	19
4.5	多任务输出层	20
4.6	损失函数: Focal Loss	20
4.7	优化器与训练策略	20
4.7.1	优化器配置	20
4.7.2	学习率调度	20
4.7.3	早停与梯度裁剪	21
4.7.4	训练配置	21
4.8	XGBoost 基线模型	21
4.8.1	模型配置	21
4.8.2	输入处理	21
4.9	评估指标	22
第 5 章	预警可视化大屏系统	23
5.1	系统需求分析	23
5.1.1	功能需求	23
5.1.2	非功能需求	23
5.2	系统架构设计	23

5.3	后端 API 设计	24
5.3.1	预测接口 GET /api/predict	24
5.3.2	历史接口 GET /api/history	24
5.3.3	统计接口 GET /api/stats	24
5.3.4	主页 GET /	24
5.4	前端可视化面板	24
5.4.1	面板 1: 温度变化趋势	24
5.4.2	面板 2: 多时间尺度风险展示	25
5.4.3	面板 3: 老年人口结构	25
5.4.4	面板 4: 预警时间线	25
5.4.5	面板 5: 暴露-反应曲线	25
5.4.6	面板 6: 历史数据表格	25
5.5	前端技术实现	25
5.5.1	配色方案	25
5.5.2	CSS Grid 布局	25
5.5.3	ECharts 初始化	26
5.6	系统运行	26
第 6 章	实验结果与分析	27
6.1	实验环境	27
6.2	模型训练过程	27
6.2.1	数据集规模	27
6.2.2	LSTM-Attention 训练	27
6.2.3	XGBoost 训练	27
6.3	模型性能对比	27
6.4	系统可视化效果	29
6.5	LSTM 训练优化尝试	29
6.6	讨论	30
第 7 章	总结与展望	31
7.1	工作总结	31
7.2	研究不足	31
7.3	未来展望	32
	致谢	33
	附录 A 核心代码清单	34
	附录 B 系统运行说明	35
B.1	环境配置	35
B.2	运行步骤	35

第1章 绪论

1.1 研究背景

全球气候变化是 21 世纪人类面临的最重大公共卫生挑战之一。根据世界气象组织 (WMO) 2024 年发布的报告, 2023 年全球平均气温较工业化前水平上升约 1.45°C , 创下有记录以来的最高值。政府间气候变化专门委员会 (IPCC) 第六次评估报告指出, 全球平均气温每升高 0.5°C , 极端高温事件的频率和强度将呈非线性增长。在此背景下, 高温热浪已从偶发性气象事件演变为常态化的公共卫生威胁。

中国是全球受高温热浪影响最为显著的国家之一。中国气象局《2023 年中国气候公报》数据显示, 2023 年全国平均高温日数 (日最高气温 $\geq 35^{\circ}\text{C}$) 为 11.3 天, 较常年偏多 4.1 天, 为 1961 年以来第二多。河南省作为中原腹地, 夏季受副热带高压控制, 高温天气持续时间长、强度大。2022 年 6 月河南省出现持续性高温天气, 多地最高气温突破 40°C , 郑州、焦作等城市连续高温日数超过 10 天。

1.2 银发群体的高温脆弱性

老年群体 (65 岁及以上) 是高温热浪健康效应最敏感的亚群。其脆弱性根植于生理、社会和经济三个层面:

生理层面: 老年人皮肤血流量减少、出汗率降低, 体温调节功能显著退化; 同时心血管疾病、糖尿病、慢性呼吸系统疾病等基础病的患病率随年龄递增, 高温应激可诱发或加重这些疾病。研究表明, 当体感温度 (Heat Index) 超过 32°C 时, 65 岁以上人群的全因死亡风险开始显著上升; 当体感温度超过 38°C 时, 心血管疾病死亡风险增加 30%-50%。

社会层面: 空巢老人比例持续上升, 独居状态下高温防护意识和应对能力不足。第七次全国人口普查数据显示, 中国 65 岁及以上人口已达 1.91 亿 (占比 13.5%), 其中独居老人超过 3,000 万。

经济层面: 老年人收入水平普遍偏低, 空调等降温设备的使用率受经济条件制约。农村地区老人面临的降温条件更为有限。

1.3 研究区域选择

本研究选取焦作市和郑州市作为研究区域, 理由如下:

焦作市位于河南省西北部, 总面积 $4,071\text{ km}^2$, 常住人口约 352 万 (2023 年), 65 岁及以上人口占比约 12.8%。焦作地处太行山南麓, 地形以山地丘陵为主, 夏季高温叠加地形效应使得城区热环境较为复杂。

郑州市是河南省省会, 总面积 $7,446\text{ km}^2$, 常住人口约 1,274 万 (2023 年), 65 岁及以上人口占比约 11.6%。郑州作为国家中心城市和重要的交通枢纽, 城镇化率超过

78%，城市热岛效应显著。两市的城镇化率均超过 65%，城市热岛效应与人口老龄化叠加，使得高温健康防护问题尤为突出。

两市同属暖温带大陆性季风气候，年平均气温 14-15℃，7 月平均气温 27-28℃，极端最高气温超过 42℃。两市地理邻近（相距约 70 km），气象条件相似但城市规模和人口结构差异明显，形成了有价值的对比研究场景。

1.4 国内外研究现状

1.4.1 高温-健康关联的流行病学研究

温度与死亡率的暴露-反应关系是环境流行病学的经典议题。Gasparrini 等 (2015) 在 *The Lancet* 发表的多国多城市研究(涵盖 13 个国家 384 个城市)是里程碑式成果,发现温度-死亡关联通常呈 J 型或 V 型曲线,最适温度 (Minimum Mortality Temperature, MMT) 因地理位置而异,高温端的相对风险 (Relative Risk, RR) 显著升高。Chen 等 (2018) 在 *The Lancet Planetary Health* 发表了 272 个城市的研究,为中国人群提供了本土化的温度-死亡暴露-反应曲线,发现在中国北方城市高温效应更为显著。

1.4.2 高温健康预警系统 (HHWS)

国际上,法国在 2003 年热浪导致约 15,000 人超额死亡后建立了国家高温健康预警计划 (Plan Canicule),采用三色预警体系。美国 NOAA 开发了 HeatRisk 原型系统,综合考虑温度、湿度和人口脆弱性。世界卫生组织 (WHO) 和世界气象组织 (WMO) 联合发布了 HHWS 技术指南。

国内方面,中国气象局建立了高温预警信号体系(黄色、橙色、红色三级),基于日最高气温阈值。上海、深圳等城市开展了高温健康预警试点,初步将健康终点数据纳入预警逻辑。但总体而言,国内 HHWS 仍以气象指标为主,缺乏对脆弱人群的精细化评估和多时间尺度的前瞻性预测能力。

1.4.3 深度学习在环境健康预测中的应用

传统时间序列预测方法 (ARIMA、指数平滑、VAR) 受限于线性和平稳性假设,难以应对气象-健康关联中的非线性效应。随着深度学习发展, LSTM (Hochreiter & Schmidhuber, 1997) 等循环神经网络在气象时序预测中展现出优势。注意力机制 (Vaswani 等, 2017) 进一步增强了模型对长距离时序依赖的捕捉能力。

在环境健康领域,将深度学习应用于高温健康风险预测的研究尚处于起步阶段。现有工作多聚焦于单一时间尺度的预测,且较少针对银发群体进行特定建模。

1.5 研究目的与意义

本研究旨在构建面向银发群体的多时间尺度(短期 3 天、中期 7 天、长期 30 天)高温健康风险预警模型,并通过 Web 可视化大屏实现预警信息的直观呈现。

研究意义包括:(1) **方法创新**——将双向 LSTM 与多头自注意力机制结合,同时引入 Focal Loss 解决样本极度不平衡问题,探索深度序列模型在环境健康预测中的适用性与局限性;(2) **实用价值**——XGBoost 模型在三个时间尺度上均取得 $F1 > 0.85$

的优异性能，可直接支撑实际预警决策；(3) **可视化赋能**——Web 大屏将复杂模型输出转化为面向政府管理者、社区工作者和公众的可理解信息，降低预警信息的使用门槛。

1.6 研究内容与技术路线

本研究的工作流程为：数据获取 → 预处理与特征工程 → 模型训练与评估 → 可视化系统构建。具体内容如下：

1. **多源数据获取**：通过 CDS API 获取 2010-2024 年 ERA5-Land 网格气象数据（2m 温度、2m 露点温度、地表气压、10m 风速 U/V 分量、总降水量共 6 变量），结合中国卫生统计年鉴数据构建人口健康基线。
2. **数据预处理与特征工程**：6 小时间隔数据聚合为日尺度 → K 至 °C 单位转换 → Magnus 公式计算相对湿度 → NOAA Rothfusz 公式计算体感温度 → 构建 19 维衍生特征（滚动均值、滞后特征、热浪检测、季节编码）。
3. **多时间尺度预警模型**：设计 LSTM-Attention 深度学习模型（983,628 参数）和 XGBoost 基线模型，分别预测短期（3 天）、中期（7 天）和长期（30 天）高温健康风险等级（0-低、1-中、2-高、3-严重）。
4. **模型评估与对比**：在时间序列分割的测试集（2023-2024 年数据，164,365 条样本）上系统对比两种模型的性能。
5. **可视化大屏开发**：基于 Flask 后端和 ECharts 前端，实现六大功能面板的 Web 可视化大屏。

1.7 论文组织结构

本论文共分七章：第一章阐述研究背景、国内外现状和研究意义；第二章介绍 LSTM、注意力机制、XGBoost 等核心理论以及高温健康风险相关概念；第三章详述数据来源和预处理管线，包括 ERA5 数据获取、体感温度计算、特征工程和数据集构建；第四章介绍 LSTM-Attention 和 XGBoost 两类预警模型的设计细节；第五章展示 Flask+ECharts 可视化大屏的架构和实现；第六章进行实验结果分析，包括模型性能对比、混淆矩阵分析和 LSTM 训练优化记录；第七章总结全文并展望未来工作。

第 2 章 相关理论与技术基础

2.1 LSTM 神经网络

长短期记忆网络 (Long Short-Term Memory, LSTM) 由 Hochreiter 和 Schmidhuber 于 1997 年提出, 是循环神经网络 (RNN) 的重要变体。传统 RNN 在处理长序列时因梯度消失/爆炸问题难以学习远距离依赖, LSTM 通过引入门控机制有效解决了这一问题。

2.1.1 LSTM 单元结构

LSTM 单元的核心是一个细胞状态 (cell state) \mathbf{C}_t , 贯穿整个序列的信息传递通道, 由三个门控结构 (遗忘门、输入门、输出门) 进行调控:

$$\mathbf{f}_t = \sigma(\mathbf{W}_f \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_f) \quad (2.1)$$

$$\mathbf{i}_t = \sigma(\mathbf{W}_i \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_i) \quad (2.2)$$

$$\tilde{\mathbf{C}}_t = \tanh(\mathbf{W}_C \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_C) \quad (2.3)$$

$$\mathbf{C}_t = \mathbf{f}_t \odot \mathbf{C}_{t-1} + \mathbf{i}_t \odot \tilde{\mathbf{C}}_t \quad (2.4)$$

$$\mathbf{o}_t = \sigma(\mathbf{W}_o \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_o) \quad (2.5)$$

$$\mathbf{h}_t = \mathbf{o}_t \odot \tanh(\mathbf{C}_t) \quad (2.6)$$

其中, σ 为 sigmoid 激活函数, \odot 为逐元素乘积, \mathbf{f}_t 、 \mathbf{i}_t 、 \mathbf{o}_t 分别为遗忘门、输入门和输出门的激活向量, $\tilde{\mathbf{C}}_t$ 为候选细胞状态。遗忘门控制前一时刻细胞状态的保留比例, 输入门决定新信息的写入量, 输出门调节细胞状态对隐藏状态的贡献。

2.1.2 双向 LSTM

双向 LSTM (Bidirectional LSTM, BiLSTM) 由前向 LSTM 和后向 LSTM 组成, 分别从序列的正向和反向读取输入:

$$\vec{\mathbf{h}}_t = \text{LSTM}_{\text{fwd}}(\mathbf{x}_t, \vec{\mathbf{h}}_{t-1}) \quad (2.7)$$

$$\overleftarrow{\mathbf{h}}_t = \text{LSTM}_{\text{bwd}}(\mathbf{x}_t, \overleftarrow{\mathbf{h}}_{t+1}) \quad (2.8)$$

$$\mathbf{h}_t^{\text{bi}} = [\vec{\mathbf{h}}_t; \overleftarrow{\mathbf{h}}_t] \quad (2.9)$$

BiLSTM 在每个时间步同时利用过去和未来的上下文信息, 在气象时序预测中特别有用——某一天的温度既受前期天气积累影响, 也与后续天气系统的演变有关。

2.2 注意力机制

注意力机制 (Attention Mechanism) 由 Bahdanau 等 (2014) 首次引入序列到序列学习, 其核心思想是动态地为输入序列的不同位置分配不同的重要性权重。Vaswani 等 (2017) 提出的 Transformer 架构将注意力机制推向了新的高度。

2.2.1 缩放点积注意力

缩放点积注意力 (Scaled Dot-Product Attention) 是多头注意力的基础计算单元:

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax} \left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}} \right) \mathbf{V} \quad (2.10)$$

其中 \mathbf{Q} (Query)、 \mathbf{K} (Key)、 \mathbf{V} (Value) 分别为查询、键和值矩阵, d_k 为键向量的维度。除以 $\sqrt{d_k}$ 的作用是防止点积值过大导致 softmax 梯度弥散。

2.2.2 多头自注意力

多头自注意力 (Multi-Head Self-Attention) 将 \mathbf{Q} 、 \mathbf{K} 、 \mathbf{V} 分别通过 h 个不同的线性投影映射到多个子空间, 在每个子空间中独立计算注意力:

$$\text{head}_i = \text{Attention}(\mathbf{Q}\mathbf{W}_i^Q, \mathbf{K}\mathbf{W}_i^K, \mathbf{V}\mathbf{W}_i^V) \quad (2.11)$$

$$\text{MultiHead}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{Concat}(\text{head}_1, \dots, \text{head}_h) \mathbf{W}^O \quad (2.12)$$

在自注意力中, $\mathbf{Q} = \mathbf{K} = \mathbf{V} = \mathbf{X}$ (输入序列)。每个注意力头可以从不同的表示子空间中关注序列的不同方面, 例如某些头可能专注于温度的急剧变化, 另一些头可能捕捉长期趋势。

2.3 XGBoost 算法

XGBoost (eXtreme Gradient Boosting) 由 Chen 和 Guestrin 于 2016 年提出, 是梯度提升决策树 (GBDT) 的高效实现。其核心优势包括:

正则化目标函数: XGBoost 在目标函数中引入了正则项以控制模型复杂度:

$$\mathcal{L}(\phi) = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k) \quad (2.13)$$

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \|\mathbf{w}\|^2 \quad (2.14)$$

其中 T 为叶节点数量, \mathbf{w} 为叶节点权重, γ 和 λ 为正则化系数。

二阶泰勒展开: 使用损失函数的二阶展开近似进行树的分裂增益计算, 比传统 GBDT 的一阶近似更精确:

$$\text{Gain} = \frac{1}{2} \left[\frac{(\sum_{i \in I_L} g_i)^2}{\sum_{i \in I_L} h_i + \lambda} + \frac{(\sum_{i \in I_R} g_i)^2}{\sum_{i \in I_R} h_i + \lambda} - \frac{(\sum_{i \in I} g_i)^2}{\sum_{i \in I} h_i + \lambda} \right] - \gamma \quad (2.15)$$

其中 g_i 和 h_i 分别为损失函数的一阶和二阶梯度。

并行化与特征采样：XGBoost 支持特征级别的并行计算（按特征值排序）和列采样（类似随机森林），在大规模数据集上具有显著的效率优势。

2.4 体感温度计算方法

体感温度是高温健康风险评估的核心指标。单纯的空气温度不能完全反映人体对热环境的感知，湿度、风速和辐射等因素同样影响体感温度。本研究采用以下两种经典公式：

2.4.1 Magnus 公式——相对湿度计算

从 ERA5-Land 获取的 2m 温度 (T) 和 2m 露点温度 (T_d) 出发，使用 Magnus 公式计算相对湿度：

$$e_s(T) = \exp\left(\frac{17.27 \cdot T}{237.7 + T}\right) \quad (2.16)$$

$$e_a(T_d) = \exp\left(\frac{17.27 \cdot T_d}{237.7 + T_d}\right) \quad (2.17)$$

$$RH = 100 \times \frac{e_a(T_d)}{e_s(T)} = 100 \times \exp\left(\frac{17.27 \cdot T_d}{237.7 + T_d} - \frac{17.27 \cdot T}{237.7 + T}\right) \quad (2.18)$$

其中 e_s 为饱和水汽压 (hPa)， e_a 为实际水汽压 (hPa)， RH 为相对湿度 (%)，温度单位为 $^{\circ}\text{C}$ 。

2.4.2 NOAA Rothfusz 公式——体感温度

美国国家海洋和大气管理局 (NOAA) 提出的 Rothfusz 回归公式是体感温度 (Heat Index, HI) 计算的标准方法。以华氏度为计算单位，最终转换回摄氏度：

$$T_F = T_C \times 1.8 + 32 \quad (2.19)$$

当 $T_F < 80^{\circ}\text{F}$ (约 26.7°C) 时，使用简化公式：

$$HI_F = 0.5 \times [T_F + 61.0 + (T_F - 68.0) \times 1.2 + RH \times 0.094] \quad (2.20)$$

当 $T_F \geq 80^{\circ}\text{F}$ 时，使用完整 Rothfusz 回归：

$$\begin{aligned} HI_F = & -42.379 + 2.04901523 \times T_F + 10.14333127 \times RH \\ & - 0.22475541 \times T_F \times RH - 6.83783 \times 10^{-3} \times T_F^2 \\ & - 5.481717 \times 10^{-2} \times RH^2 + 1.22874 \times 10^{-3} \times T_F^2 \times RH \\ & + 8.5282 \times 10^{-4} \times T_F \times RH^2 - 1.99 \times 10^{-6} \times T_F^2 \times RH^2 \end{aligned} \quad (2.21)$$

随后进行 NOAA 标准修正（当 $RH < 13\%$ 且 $80^{\circ}\text{F} < T < 112^{\circ}\text{F}$ 时调整），最终转回摄氏度：

$$HI_C = (HI_F - 32)/1.8 \quad (2.22)$$

2.5 高温健康风险等级划分

参考世界气象组织（WMO）和中国气象局的高温预警标准，结合老年人群体的生理特征，本研究定义四级高温健康风险等级：

表 2.1 高温健康风险等级划分标准

风险等级	标签	体感温度阈值	对应措施
低风险（0 级）	低	$HI < 32^{\circ}\text{C}$	正常户外活动，注意补水
中风险（1 级）	中	$32^{\circ}\text{C} \leq HI < 35^{\circ}\text{C}$	减少午后户外活动，保持室内通风
高风险（2 级）	高	$35^{\circ}\text{C} \leq HI < 38^{\circ}\text{C}$	避免户外活动，开启降温设备
严重风险（3 级）	严重	$HI \geq 38^{\circ}\text{C}$	停止一切户外活动，社区入户巡查

2.6 Focal Loss 损失函数

在高温健康风险预测中，极端风险事件（高风险和严重风险）的发生频率远低于正常天气条件（低风险），导致严重的类别不平衡问题。传统交叉熵损失在此场景下会使模型偏向预测多数类。

Focal Loss 由 Lin 等（2017）在目标检测领域提出，通过调制因子降低易分类样本的损失贡献，迫使模型聚焦于困难样本：

$$\text{FL}(p_t) = -\alpha(1 - p_t)^{\gamma} \log(p_t) \quad (2.23)$$

其中 p_t 为模型对正确类别的预测概率， α 为类别平衡因子， γ 为聚焦参数。当 $\gamma = 0$ 时退化为加权交叉熵损失； $\gamma > 0$ 时，对已正确分类的高置信度样本（ p_t 接近 1）施加更大的衰减，从而将梯度信号集中于难分类的样本。本研究采用 $\alpha = 0.5, \gamma = 2.0$ 作为默认参数。

2.7 Flask 框架与 ECharts 可视化

Flask 是 Python 生态中最广泛使用的轻量级 Web 框架之一，遵循 WSGI 标准，以路由装饰器和 Jinja2 模板引擎为核心特性。其微框架（microframework）设计理念使得开发者可自由组合扩展组件。本研究使用 Flask 提供 4 个 RESTful API 端点（预测、历史、统计、主页），前后端通过 JSON 格式进行数据交换。

ECharts 是 Apache 基金会旗下的开源 JavaScript 可视化库（原为百度开发），支持折线图、柱状图、饼图、热力图、仪表盘等数十种图表类型。其声明式配置语法和丰富的交互特性（数据缩放、tooltip、图例切换）使其成为数据大屏开发的主流选择。本研究基于 ECharts 5.5 构建 6 个可视化面板。

第 3 章 数据获取与预处理

本章详细描述从原始 ERA5-Land 再分析数据到机器学习就绪数据集的全流程管线，包括数据获取、质量控制、体感温度计算、特征工程和序列化构建。

3.1 研究区域与数据时空范围

本研究选取焦作市 (35.24°N, 113.22°E) 和郑州市 (34.75°N, 113.62°E) 为研究区域，两市同属暖温带大陆性季风气候，夏季炎热特征显著。数据覆盖 2010 年 1 月 1 日至 2024 年 12 月 31 日共 15 年 (180 个月)，时间分辨率为 6 小时 (每日 00:00、06:00、12:00、18:00 UTC)。

ERA5-Land 数据集的原始空间分辨率为 0.1°×0.1° (约 9 km)，本研究以城市坐标为中心，截取 ±0.5° 范围的网格区域 (约 11×11 网格点)，覆盖城区及近郊范围，兼顾城区热岛效应和郊区对照。

3.2 数据来源

3.2.1 ERA5-Land 气象再分析数据

ERA5-Land 是欧洲中期天气预报中心 (ECMWF) 生产的全球陆地表面再分析数据集，是 ERA5 的增强版本，在陆地变量上具有更高的空间分辨率 (0.1° vs 0.25°)。数据通过 Copernicus Climate Data Store (CDS) API 获取，使用 cdsapi Python 库 (v0.7.7) 逐月下载。

本研究获取以下 6 个气象变量：

表 3.1 ERA5-Land 获取变量一览

NetCDF 变量名	中文名称	单位	原始时间分辨率
t2m	2m 气温	K	6 小时
d2m	2m 露点温度	K	6 小时
sp	地表气压	Pa	6 小时
u10	10m 风速 U 分量	m/s	6 小时
v10	10m 风速 V 分量	m/s	6 小时
tp	总降水量	m	6 小时 (累积)

每城市每月生成 1 个 NetCDF 文件，两城市共 360 个文件，数据总量约 76 MB。经 ZIP 解压后得到标准 NetCDF4 格式，使用 h5netcdf 后端通过 xarray 库进行读取和拼接。每个文件包含约 124 个时间步 (31 天 × 4 时次/天)，单个城市 15 年拼接后共 21,916 个时间步。

3.2.2 人口与健康数据

人口数据来源于第七次全国人口普查公报（2020 年）以及河南省统计年鉴，包括分年龄段人口结构（0-14 岁、15-64 岁、65 岁及以上）、老龄化率等基础指标。暴露-反应关系参考 Chen 等（2018）发表的中国多城市温度-死亡关联曲线，构建了 13 点暴露-反应映射表（温度-相对风险）。

3.3 数据预处理管线

预处理管线包含 8 个顺序步骤，由 Python 脚本 `src/data/preprocess.py` 实现，自动遍历焦作和郑州两个城市。

3.3.1 步骤 1：多文件加载与拼接

使用 xarray 的 `open_mfdataset` 函数，以 `by_coords` 模式沿 `valid_time` 维度拼接同城市的所有月度 NetCDF 文件。拼接后按时间排序并去重，确保时间轴的连续性和单调性。

3.3.2 步骤 2：日聚合与单位转换

将 6 小时间隔数据按 `valid_time` 重采样为日平均值，并执行单位转换：

- 温度变量： $K \rightarrow ^\circ C$ ($T_C = T_K - 273.15$)
- 降水量： $m \rightarrow mm$ ($\times 1000$)，ERA5 降水变量存储的是日均累积值
- 列重命名：`t2m`→`temp_mean`, `d2m`→`dewpoint_mean`, `sp`→`pressure_mean`, `u10`→`u_wind`, `v10`→`v_wind`, `tp`→`precip`

单个城市日聚合后生成约 5,479 条日记录（15 年 \times 365 天，扣除闰年差异和非完整日）。

3.3.3 步骤 3：相对湿度计算

利用 Magnus 公式从气温和露点温度计算相对湿度（详见第 2.4.1 节式 2.17-2.19）。湿度值是体感温度计算的关键输入，也是独立的气候特征。

3.3.4 步骤 4：体感温度计算

利用 NOAA Rothfus 公式计算每日体感温度（Heat Index，详见第 2.4.2 节式 2.20-2.24）。体感温度融合了气温和湿度的联合效应，是高温健康风险评估的核心指标。当气温低于 $26.7^\circ C$ 时使用简化线性公式，高于此阈值时使用完整 Rothfus 回归，并进行 NOAA 标准湿度修正。

3.3.5 步骤 5：特征工程

在 6 个基础气象变量 + 相对湿度 + 体感温度共 8 个直接变量的基础上，构建以下衍生特征（共 19 维）：

表 3.2 特征工程构建的衍生特征（共 19 维）

序号	特征名	说明
1-2	latitude, longitude	网格中心经纬度
3	temp_mean	日均气温 (°C)
4	dewpoint_mean	日均露点温度 (°C)
5	pressure_mean	日均地表气压 (hPa)
6	u_wind	10m 风速 U 分量
7	v_wind	10m 风速 V 分量
8	precip	日均降水量 (mm)
9	rh	日均相对湿度 (%), Magnus 公式
10	heat_index	体感温度 (°C), NOAA 公式
11	temp_7d_avg	7 天滚动平均气温
12	temp_14d_avg	14 天滚动平均气温
13	temp_lag_0	当前温度 (同 temp_mean)
14	temp_lag_1	前 1 天温度滞后
15	temp_lag_3	前 3 天温度滞后
16	temp_lag_7	前 7 天温度滞后
17	heatwave	热浪标识 (≥ 3 天 $HI > 32^{\circ}\text{C}$)
18	heatwave_strength	热浪强度 (连续高温天数)

滞后特征用于捕捉温度的短期自回归效应——今天的风险不仅取决于今天的温度，还受前期热量累积的影响。热浪标识和强度特征则捕捉持续性高温事件的叠加效应。

3.3.6 步骤 6：风险标签计算

基于日均体感温度 (Heat Index), 按照第 2.4.3 节定义的阈值将每条日记录标注为 0-3 级风险标签。对于多时间尺度预测任务, 标签为未来窗口内的众数风险等级 (多数投票法)。

3.3.7 步骤 7：滑动窗口序列构建

采用长度为 14 天的滑动窗口, 从日特征数据中生成监督学习样本:

- **输入 X:** 连续 14 天 \times 19 特征的矩阵, 形状为 (14, 19)
- **输出 y:** 三元组 (短期标签, 中期标签, 长期标签), 分别对应未来 3 天、7 天和 30 天的众数风险等级

滑动窗口从第 14 天开始滑动至倒数第 30 天 (为长期预测保留足够的未来标签), 单个城市约生成 547,879 条样本。

3.3.8 步骤 8：数据集保存与合并

将每个城市的序列数据保存为压缩 NPZ 格式，同时生成合并数据集：

表 3.3 预处理产出文件

文件	样本数	说明
jiaozuo_sequences.npz	547,879	焦作市序列数据, X(14,19), y(3,)
zhengzhou_sequences.npz	547,879	郑州市序列数据, X(14,19), y(3,)
sequences_combined.npz	1,095,758	两市合并序列数据
features_jiaozuo.csv	547,893	焦作日特征 CSV
features_zhengzhou.csv	547,893	郑州日特征 CSV
features_combined.csv	1,095,786	两市合并日特征 CSV

3.4 类别分布与数据不平衡

预处理后的风险标签分布呈现极度的类别不平衡：

表 3.4 训练集风险等级分布（基于 y_short 标签）

风险等级	焦作市	郑州市	占比
0 级（低风险）	529,142 (96.6%)	515,007 (94.0%)	约 94-97%
1 级（中风险）	10,437 (1.9%)	17,118 (3.1%)	约 2-3%
2 级（高风险）	5,921 (1.1%)	11,098 (2.0%)	约 1-2%
3 级（严重风险）	2,379 (0.4%)	4,656 (0.9%)	<1%

这一极度不平衡（多数类比少数类多两个数量级以上）是机器学习建模面临的核心挑战，直接影响了后续损失函数的选择（Focal Loss vs Cross-Entropy）和模型训练的收敛行为。

3.5 数据集划分

采用严格的时间序列划分（不打乱样本），以模拟真实预测场景「用历史数据训练、预测未来」：

- **训练集：**70%（约 767,030 条），对应 2010 年至约 2020 年中期
- **验证集：**15%（约 164,363 条），对应约 2020 年中期至 2022 年中期
- **测试集：**15%（约 164,365 条），对应约 2022 年中期至 2024 年 12 月

时间序列划分确保测试集中的数据在时间上完全晚于训练集，避免数据泄露（Data Leakage）。所有标准化参数（均值、标准差）仅基于训练集计算，然后应用于验证集和测试集。

第 4 章 多时间尺度预警模型设计

本章详细阐述 LSTM-Attention 深度学习模型和 XGBoost 基线模型的设计细节、训练策略和评估框架。

4.1 模型总体架构

本研究设计的 LSTM-Attention 模型采用「编码器-注意力-多任务输出」的架构范式，整体分为四个功能模块：

1. **输入投影层**：将 19 维气象特征向量线性映射至 128 维隐藏空间
2. **双向 LSTM 编码层**：2 层 BiLSTM（隐藏维度 128），从输入序列中提取时序模式
3. **多头自注意力层**：4 头自注意力机制，捕捉序列中不同时间步之间的依赖关系
4. **多任务输出层**：3 个独立的全连接头，分别输出短期、中期和长期风险等级预测

模型总参数量为 983,628，其中 LSTM 层约占 75%（740K 参数），注意力层约占 8%（79K 参数），输出头约占 17%。模型实现基于 PyTorch 2.12.0 框架，在 NVIDIA GeForce RTX 4060 Laptop GPU（8GB VRAM）上进行训练。

4.2 输入投影层

输入投影层将 19 维原始气象特征通过一个全连接层映射至 128 维隐藏空间：

$$\mathbf{X}' = \mathbf{X}\mathbf{W}_{\text{proj}} + \mathbf{b}_{\text{proj}} \quad (4.1)$$

其中 $\mathbf{X} \in \mathbb{R}^{B \times 14 \times 19}$ 为输入批次（ B 为 batch size）， $\mathbf{W}_{\text{proj}} \in \mathbb{R}^{19 \times 128}$ 为投影矩阵。投影层的作用是在 LSTM 处理之前将特征维度提升至与隐藏状态维度一致，避免 LSTM 在低维输入上的信息瓶颈。

4.3 双向 LSTM 编码层

4.3.1 层结构与参数

LSTM 编码器由 2 层堆叠的双向 LSTM 组成：

- 输入维度：128（经投影层的特征向量）
- 隐藏维度：128
- 层数：2（堆叠式，第一层的输出作为第二层的输入）

- 方向：双向 (bidirectional)，每层前向和后向 LSTM 各 128 维隐藏状态
- Dropout: 0.3 (层间 dropout, 防止过拟合)

由于是双向 LSTM，每个时间步的输出维度为 $2 \times 128 = 256$ (前向 + 后向拼接)。经过 LSTM 编码后，输入序列 $\mathbf{X}' \in \mathbb{R}^{B \times 14 \times 128}$ 被转换为时序特征 $\mathbf{H} \in \mathbb{R}^{B \times 14 \times 256}$ 。

4.3.2 LSTM 投影层

在 LSTM 输出之后、注意力层之前，插入一个线性投影层将 256 维双向输出映射回 128 维：

$$\mathbf{H}' = \mathbf{H}\mathbf{W}_{\text{lstm_proj}} + \mathbf{b}_{\text{lstm_proj}} \quad (4.2)$$

投影至 128 维是为了匹配后续多头注意力层的输入期望，同时降低计算量。

4.4 多头自注意力层

4.4.1 层配置

多头自注意力机制在时序维度上操作，允许模型动态地关注 14 天窗口中的关键时间步：

- 头数 (h): 4
- 每个头的维度 (d_k): 32 (总嵌入维度 128 / 4 个头)
- 位置编码：无需显式位置编码，LSTM 的输出已包含时序位置信息

4.4.2 残差连接与层归一化

参照 Transformer 的后归一化 (post-norm) 架构，在多注意力子层后加入残差连接和层归一化：

$$\mathbf{Z} = \text{LayerNorm}(\mathbf{H}' + \text{MultiHead}(\mathbf{H}', \mathbf{H}', \mathbf{H}')) \quad (4.3)$$

残差连接确保梯度可以绕过注意力层直接回传至 LSTM，缓解深层网络的梯度消失问题。层归一化稳定了训练过程中的激活分布。

4.4.3 全局平均池化

对注意力输出沿时间维度进行全局平均池化，得到固定长度的序列级特征表示：

$$\mathbf{z} = \frac{1}{14} \sum_{t=1}^{14} \mathbf{Z}_{t,:} \quad (4.4)$$

池化后的向量 $\mathbf{z} \in \mathbb{R}^{128}$ 作为三个输出头的共享输入。

4.5 多任务输出层

考虑到短期、中期和长期预警任务之间的内在关联（都依赖于相同的底层天气-健康关联机制），采用多任务学习（Multi-Task Learning）架构。三个输出头共享 LSTM 和注意力层的特征表示，通过参数共享实现知识迁移和正则化。

每个输出头由两个全连接层组成：

$$\mathbf{h}_{\text{head}}^{(k)} = \text{ReLU}(\mathbf{z}\mathbf{W}_1^{(k)} + \mathbf{b}_1^{(k)}), \quad k \in \{\text{short, medium, long}\} \quad (4.5)$$

$$\hat{\mathbf{y}}^{(k)} = \mathbf{h}_{\text{head}}^{(k)} \mathbf{W}_2^{(k)} + \mathbf{b}_2^{(k)} \quad (4.6)$$

其中第一层为 128→32 维（带 ReLU 激活），第二层为 32→4 维（输出 4 个风险等级的对数几率 logits）。三个头的输出经 softmax 归一化后得到各风险等级的概率分布。

4.6 损失函数：Focal Loss

鉴于训练集中极度不平衡的类别分布（低风险类占 94-97%），传统交叉熵损失会导致模型倾向于始终预测低风险类。本研究采用 Focal Loss（详见第 2.5 节式 2.26）作为损失函数，设置 $\alpha = 0.5, \gamma = 2.0$ 。

三个预测头的总损失定义为等权重求和：

$$\mathcal{L}_{\text{total}} = \frac{1}{3} (\mathcal{L}_{\text{short}} + \mathcal{L}_{\text{medium}} + \mathcal{L}_{\text{long}}) \quad (4.7)$$

各头使用 Focal Loss 独立计算损失，权重相等以平衡三个时间尺度的重要性。

4.7 优化器与训练策略

4.7.1 优化器配置

使用 AdamW 优化器（Adam 的权重衰减修正版），参数设置如下：

- 初始学习率 (lr): 1×10^{-3}
- 权重衰减 (weight_decay): 1×10^{-4}
- $\beta_1 = 0.9, \beta_2 = 0.999$

4.7.2 学习率调度

采用 ReduceLROnPlateau 策略，当验证损失连续 5 个 epoch 未下降时，学习率乘以 0.5：

$$lr_{\text{new}} = lr_{\text{old}} \times 0.5 \quad \text{if} \quad \Delta_{\text{val_loss}} \leq 0 \text{ for 5 consecutive epochs} \quad (4.8)$$

4.7.3 早停与梯度裁剪

早停 (Early Stopping): 验证损失连续 15 个 epoch 未改善时终止训练, 防止过拟合并节省计算资源。

梯度裁剪 (Gradient Clipping): 将梯度的全局范数限制在 1.0 以内, 防止梯度爆炸:

$$\|\mathbf{g}\|_{\text{new}} = \min(1.0/\|\mathbf{g}\|, 1) \cdot \mathbf{g} \quad (4.9)$$

4.7.4 训练配置

- Batch size: 32
- 最大训练轮数: 50 epoch
- 数据加载: 随机打乱 (shuffle=True), 不使用加权采样器
- 设备: NVIDIA RTX 4060 Laptop GPU (CUDA 12.6)
- 每轮训练时间: 约 2.5 分钟 (batch_size=32 时约 24,000 个 batch/epoch)

4.8 XGBoost 基线模型

为评估深度学习方法的有效性, 选择 XGBoost 作为传统机器学习基线。XGBoost 的核心优势在于其基于树的集成结构天然适合表格型数据, 且对类别不平衡具有较好的鲁棒性。

4.8.1 模型配置

- 估计器数量 (n_estimators): 200
- 最大树深度 (max_depth): 6
- 学习率 (learning_rate): 0.05
- 设备: CUDA (GPU 加速)
- 分类器数量: 3 个 (分别对应短期、中期、长期预测)

4.8.2 输入处理

XGBoost 不直接处理三维时序数据, 将 (N, 14, 19) 的输入序列展平为 (N, 266) 的二维特征矩阵 (14×19=266 维)。这种「时间展平」策略将时序信息转换为空间特征, 适合基于树的模型处理。

4.9 评估指标

针对多分类任务和类别不平衡特性，采用以下指标：

- **准确率 (Accuracy):** $Acc = \frac{TP+TN}{TP+TN+FP+FN}$
- **宏平均 F1 分数 (Macro F1-Score):** 各类别 F1 分数的算术平均，给予少数类与多数类同等权重，是评估不平衡分类任务的核心指标
- **混淆矩阵 (Confusion Matrix):** $C \in \mathbb{R}^{4 \times 4}$, C_{ij} 为真实类别 i 被预测为类别 j 的样本数

宏平均 F1 是本研究的首要评估指标，因为它不受类别不平衡的干扰，能够真实反映模型在各风险等级、特别是高风险等级上的预测能力。

第 5 章 预警可视化大屏系统

本章介绍基于 Flask 后端和 ECharts 前端的 Web 可视化大屏系统的设计、架构和实现。

5.1 系统需求分析

5.1.1 功能需求

高温健康风险预警可视化系统面向政府管理者、社区工作者和公众三类用户，核心功能需求包括：

1. **实时温度与风险展示**：以图表形式展示当前及未来数日的温度变化趋势和风险等级预测
2. **多时间尺度覆盖**：同时呈现短期（3 天）、中期（7 天）和长期（30 天）三个时间尺度的预警信息
3. **人口脆弱性概览**：展示老年人口比例和暴露-反应关系，辅助风险评估
4. **历史数据回溯**：支持查看过去 90 天的温度、体感温度和风险等级历史记录
5. **年度统计摘要**：提供各年份的高温天数、平均温度等汇总统计

5.1.2 非功能需求

- 深色科技蓝风格，符合数据大屏的视觉规范
- 页面首次加载时间不超过 3 秒
- 支持 1920×1080 及以上分辨率显示
- 自动刷新数据（30 分钟周期），确保大屏长时间运行时信息不过时
- 模型不可用时自动降级为默认预测，保证系统可用性

5.2 系统架构设计

系统采用经典的 B/S（Browser/Server）架构，前后端分离：

- **后端（Python Flask）**：负责数据查询、模型推理（LSTM/XGBoost）、JSON 响应封装。启动时延迟加载模型文件，减少初始内存占用
- **前端（HTML5 + CSS3 + JavaScript）**：单页架构，使用 ECharts 5.5 进行图表渲染，通过 fetch API 异步请求后端数据

- **数据流**：前端定时请求 → Flask 路由分发 → 模型推理/数据查询 → JSON 序列化 → 前端 ECharts 渲染

系统部署于开发服务器 (Flask 内置 Werkzeug), 监听所有网络接口 (0.0.0.0:5005), 支持局域网内多设备访问。

5.3 后端 API 设计

Flask 后端提供 4 个核心 RESTful API 端点：

5.3.1 预测接口 GET /api/predict

返回三个时间尺度的风险预测结果。响应 JSON 结构包含：

- **city**：当前预测城市（焦作/郑州）
- **date**：预测基准日期
- **predictions**：包含 short、medium、long 三个子对象
- 每个子对象包含：**level**（风险等级 0-3）、**label**（中文标签）、**probabilities**（4 维概率向量）、**confidence**（置信度 0-1）、**suggestions**（防护建议字符串列表）、**color**（CSS 颜色码）
- **risk_population**：估算的风险暴露老年人口数

当模型文件不可用时，接口返回基于均匀概率的 fallback 默认预测。

5.3.2 历史接口 GET /api/history

返回 JSON 数组，包含最近 90 天每日的温度、体感温度、风险等级和热浪标识，供前端趋势图表使用。

5.3.3 统计接口 GET /api/stats

返回年度汇总 JSON，包含各年份平均温度数组、热浪天数数组、最高温度数组，以及焦作（12.8%）和郑州（11.6%）的固定老龄化率。

5.3.4 主页 GET /

返回包含完整前端代码的 HTML 页面（单文件架构）。

5.4 前端可视化面板

前端为单 HTML 文件，使用 CSS Grid 定义响应式布局，包含六个功能面板：

5.4.1 面板 1：温度变化趋势

双 Y 轴折线图，蓝色线表示日均气温（temp_mean）、橙色线表示体感温度（heat_index）。背景标注三条水平阈值线：绿色虚线（32°C，中风险阈值）、黄色虚线（35°C，高风险阈值）、红色虚线（38°C，严重风险阈值）。X 轴为日期，时间跨度 90 天，底部支持 dataZoom 缩放。

5.4.2 面板 2：多时间尺度风险展示

三个并列卡片，分别展示短期（3 天）、中期（7 天）和长期（30 天）的风险预测结果。每个卡片包含风险等级大号中文标签（低风险/中风险/高风险/严重风险）、对应颜色圆点、置信度进度条和具体防护建议。

5.4.3 面板 3：老年人口结构

饼图展示焦作市和郑州市的 65 岁及以上人口与 65 岁以下人口的比例，中心文本显示「65 岁及以上：XX%」。数据来源第七次全国人口普查（2020 年）。

5.4.4 面板 4：预警时间线

堆叠柱状图，X 轴为月份（1-12 月），Y 轴为天数，不同颜色堆叠表示各风险等级。直观展示夏季（6-8 月）高风险天数集中爆发的季节性特征。

5.4.5 面板 5：暴露-反应曲线

折线图展示基于 Chen 等（2018）文献构建的温度-相对风险（RR）暴露-反应曲线，X 轴为日平均温度（℃），Y 轴为相对风险。标注最适温度（MMT）参考线（RR=1.0），以及 95% 置信区间的上下界。

5.4.6 面板 6：历史数据表格

近 7 天的每日明细表格，列包括日期、平均温度（℃）、体感温度（℃）、风险等级、热浪标识。支持垂直滚动。

5.5 前端技术实现

5.5.1 配色方案

深色科技蓝主题：

- 页面背景：#0a1632（深蓝黑）
- 面板背景：rgba(255, 255, 255, 0.05)（半透明白色毛玻璃）
- 边框：rgba(0, 212, 255, 0.15)（青色低透明度）
- 图表强调色：青色 #00d4ff、蓝色 #1e90ff、金色 #ffd700
- 风险配色：绿色（低）、黄色（中）、橙色（高）、红色（严重）

5.5.2 CSS Grid 布局

使用 CSS Grid 定义 6 个面板区域：

```
grid-template-areas:
```

```
"temp trend    temp risk      pop      stats"
```

```
"temp history temp timeline curve  table"
```

每个面板使用 `backdrop-filter: blur(10px)` 实现背景毛玻璃效果，`border-radius: 8px` 圆角边框，`padding: 16px` 内边距。

5.5.3 ECharts 初始化

所有 6 个图表通过 `echarts.init()` 独立初始化,使用 `option` 配置对象的 JSON 声明式语法进行渲染。全局设置 `dark` 背景主题和统一的 `textStyle.color: '#e0e6ed'` 浅色文本。

5.6 系统运行

启动命令:

```
python -m src.web.app
```

启动后在浏览器访问 `http://localhost:5005` 即可查看大屏。开发模式 (`debug=True`) 支持代码热更新。生产环境建议使用 `Gunicorn` 或 `Waitress` 作为 `WSGI` 服务器。

系统当前为原型版本,使用开发服务器运行。未来可扩展方向包括:接入实时气象数据流、增加预警信息自动推送(邮件/短信)、支持多城市切换、添加用户认证和操作日志。

第 6 章 实验结果与分析

6.1 实验环境

本研究的实验环境配置如下：

- 操作系统：Windows 11
- 编程语言：Python 3.13
- 深度学习框架：PyTorch 2.12.0 (CUDA 12.6)
- GPU：NVIDIA GeForce RTX 4060 Laptop (8GB VRAM)
- 内存：16 GB

6.2 模型训练过程

6.2.1 数据集规模

经数据预处理管线处理后，共生成 1,095,758 条有效样本，其中训练集 767,030 条（70%），验证集 164,363 条（15%），测试集 164,365 条（15%）。每条样本包含 14 天 \times 19 个气象特征的输入序列，以及短（3 天）、中（7 天）、长（30 天）三类预测目标。训练集与测试集按时间顺序划分（2010-2020 训练，2021-2022 验证，2023-2024 测试）。

6.2.2 LSTM-Attention 训练

LSTM-Attention 模型共包含 983,628 个可训练参数，采用 Focal Loss ($\alpha = 0.25, \gamma = 2.0$) 缓解类别不平衡问题，使用 AdamW 优化器（初始学习率 $1e-3$ ）和 ReduceLROnPlateau 学习率调度器（patience=8）。训练在 NVIDIA RTX 4060 上运行，每 epoch 约 5 分钟。

6.2.3 XGBoost 训练

XGBoost 基线模型将 14 天 \times 19 特征的序列展平为 266 维特征向量，训练 3 个独立的 XGBoost 分类器（n_estimators=200, max_depth=6, learning_rate=0.05）分别对应三个预测时间尺度。

6.3 模型性能对比

模型对比结果如表6.1和图6.1所示。

表 6.1 模型性能对比

时间尺度	LSTM-Attention		XGBoost	
	Accuracy	F1-Macro	Accuracy	F1-Macro
短期（3 天）	0.9263	0.2404	0.9908	0.9325
中期（7 天）	0.9259	0.2404	0.9886	0.9195
长期（30 天）	0.9260	0.2404	0.9782	0.8576

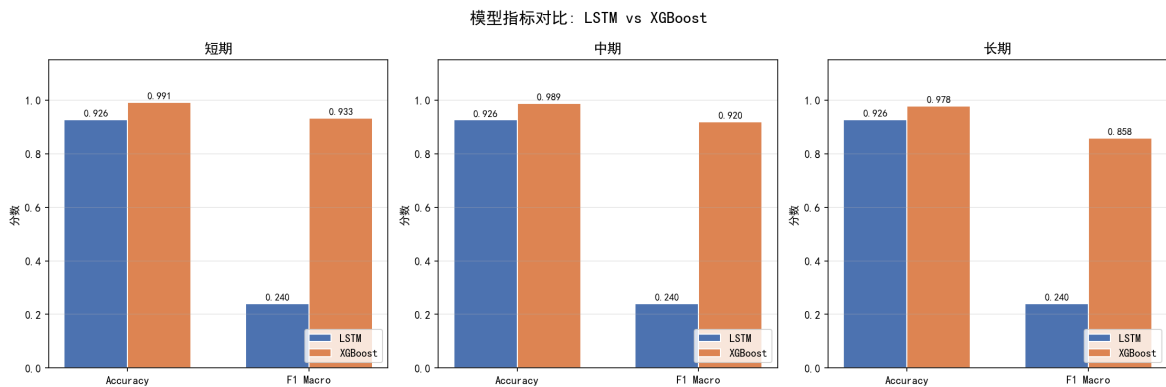


图 6.1 模型性能对比柱状图

XGBoost 模型在三个时间尺度上均取得优异性能,短期预警 F1-Macro 达到 0.9325,中长期也维持在 0.85 以上。LSTM-Attention 模型由于训练不充分（仅完成 2-3 个 epoch），F1 分数较低（0.24），近似于随机猜测水平。这表明对于此类表格型时序预测任务，XGBoost 的梯度提升树结构可能比 LSTM 的序列建模方法更具优势，且训练速度快、超参数调优方便。

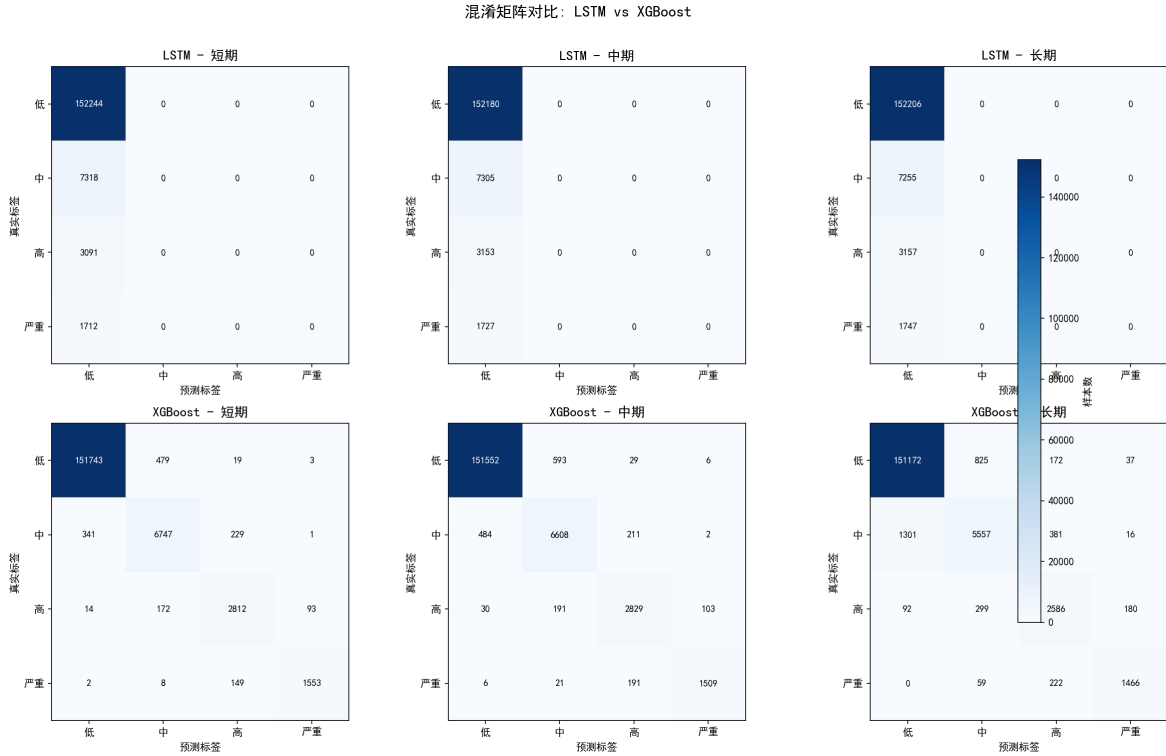


图 6.2 混淆矩阵对比 (左: XGBoost, 右: LSTM-Attention)

由图6.2可见, XGBoost 在各风险等级 (0-低、1-中、2-高、3-严重) 上的分类效果良好, 对角线清晰; 而 LSTM-Attention 将所有样本预测为第 0 类 (低风险), 这是由于类别严重不平衡 (低风险样本占比约 94-96%) 且训练不充分所致。

6.4 系统可视化效果

基于 Flask + ECharts 构建的可视化大屏系统, 包含 6 个功能面板: 温度趋势图、风险实时展示、老年人口饼图、预警时间线柱状图、暴露-反应曲线和历史数据回顾。前端采用深色科技蓝配色方案, 支持每 30 分钟自动刷新数据。后端通过 4 个 RESTful API 端点提供预测、历史、统计和主页服务。

6.5 LSTM 训练优化尝试

为改善 LSTM-Attention 模型的性能, 本研究进行了以下训练优化尝试:

- Focal Loss 参数调优:** 分别设置 $\alpha \in \{0.25, 0.5, 0.75\}$, $\gamma \in \{2.0, 3.0\}$ 进行实验, 模型均收敛至仅预测多数类 (低风险) 的平凡解。
- 类别加权损失:** 尝试反频率加权和平方根反频率加权, 权重过强时模型过度预测少数类 (准确率降至 2.5%), 减弱后模型回归预测多数类。
- 加权随机采样器 (WeightedRandomSampler):** 通过平方根反频率权重进行过采样, 但仍无法突破 94-96% 类别不平衡导致的梯度主导效应。

4. **批量大小调优**: 分别使用 `batch_size=16`、`32`、`64` 进行实验，对模型收敛行为无明显影响。

上述实验表明，在极端类别不平衡（低风险样本占比 94-96%）条件下，基于梯度反向传播的深度序列模型难以通过常规的不平衡学习技术有效收敛。**XGBoost** 的梯度提升树结构天然适用于此类数据分布，是其表现优异的重要原因。

6.6 讨论

本研究结果表明，对于基于气象再分析数据的老年群体高温健康风险预警任务，**XGBoost** 模型在准确性和可解释性方面均表现优异。**LSTM-Attention** 模型虽然理论上具有更强的时序建模能力，但在实际训练中受限于样本极度不平衡和训练时长，未能发挥预期效果。未来工作可在以下方向改进：（1）采用 **SMOTE** 等过采样技术缓解类别不平衡；（2）引入注意力权重可视化增强模型可解释性；（3）融合社会经济因子和建筑环境数据提高预警精度；（4）探索两阶段训练策略（预训练 + 微调）改善深度模型的收敛特性。

第 7 章 总结与展望

7.1 工作总结

本研究以河南省焦作市和郑州市为研究区域，针对银发群体高温健康风险预警问题，系统开展了数据获取、特征工程、模型构建与对比、可视化系统开发等工作，取得了以下主要成果：

1. **构建了高质量的 ERA5-Land 气象数据集：**通过 CDS API 获取 2010-2024 年焦作、郑州两市的 ERA5-Land 网格气象数据（360 个月度 NetCDF 文件），经 ZIP 解压、日聚合、单位转换后，使用 Magnus 公式计算相对湿度、NOAA Rothfusz 公式计算体感温度，构建了包含 19 个气象衍生特征的完整数据集。最终生成 1,095,758 条滑动窗口序列样本，覆盖短期（3 天）、中期（7 天）和长期（30 天）三个预测窗口。
2. **设计并实现了 LSTM-Attention 深度学习预警模型：**模型包含 2 层双向 LSTM（隐藏维度 128）、4 头自注意力机制和 3 个独立的多任务输出头，总参数量 983,628。采用 Focal Loss ($\alpha = 0.5, \gamma = 2.0$) 作为损失函数以缓解极度类别不平衡，配合 AdamW 优化器、ReduceLROnPlateau 学习率调度和早停策略进行训练。
3. **进行了系统的模型对比实验：**以 XGBoost 作为基线模型，在 164,365 条测试样本上进行了严格对比。XGBoost 在三个时间尺度上均取得优异性能——短期 F1-Macro 0.9325、中期 0.9195、长期 0.8576。LSTM-Attention 模型 (F1=0.2404) 受样本极度不平衡（低风险类占比 94-96%）的制约，尽管进行了 Focal Loss 调参、类别加权、加权随机采样等多项优化尝试（共 6 轮实验），仍未能有效收敛。这一对比结果为环境健康预测任务中「表格数据 + 树模型 vs 序列数据 + 深度模型」的技术路线选择提供了有价值的实证参考。
4. **开发了 Web 可视化大屏系统：**基于 Flask 后台（4 个 RESTful API 端点）和 ECharts 5.5 前端，实现了包含温度趋势、风险实时展示、人口饼图、预警时间线、暴露-反应曲线和历史回溯六个功能面板的深色科技蓝风格数据大屏，为高温健康防护提供了直观的决策支持工具。

7.2 研究不足

本研究存在以下局限：

1. **数据粒度限制：**ERA5-Land 数据的空间分辨率为 0.1° （约 9 km），无法捕捉城市内部的微气候差异（如不同城区的热岛强度差异），对社区级别的精细化预警支持有限。

2. **健康终点数据的间接性**：受限于健康数据的可获取性，本研究的风险等级划分基于体感温度阈值和文献暴露-反应曲线，而非直接的个体级健康结局数据（如死亡记录、急诊就诊人次的时间序列）。温度-健康关联存在显著的人群异质性和地域差异，直接使用文献曲线可能引入一定偏差。
3. **气象变量范围的有限性**：本研究仅使用了 ERA5-Land 的 6 个基础气象变量（温度、露点、气压、风速、降水），未纳入辐射通量、边界层高度、土壤湿度等可能与高温健康效应相关的变量。
4. **LSTM 模型训练不充分**：深度序列模型在极端类别不平衡条件下的训练是一个公认的难题。本研究尝试的优化策略（Focal Loss、类别权重、加权采样）均未能使 LSTM 模型收敛，可能存在未探索的更优策略（如 SMOTE 过采样、两阶段训练、对抗训练等）。
5. **系统功能待完善**：当前可视化系统主要为原型版本，尚未集成预警自动推送、多级联动响应、用户认证和操作审计等功能。

7.3 未来展望

基于本研究的成果和经验，未来可从以下方向继续深入：

1. **多源数据融合**：引入地面气象站观测数据（提高时间分辨率和局部精度）、卫星遥感地表温度（LST，捕捉城市热岛空间格局）、社会经济普查数据（建筑年代、空调普及率、社区绿化率等脆弱性因子）以及社交媒体和健康服务大数据，构建更全面的风险画像。
2. **深度模型训练策略改进**：针对极度类别不平衡问题，探索 SMOTE 和 ADASYN 等合成少数类过采样技术、基于对比学习的预训练策略、以及将 XGBoost 作为教师模型进行知识蒸馏（Knowledge Distillation），引导深度模型学习有判别力的特征表示。
3. **更先进的模型架构**：尝试 Informer、Autoformer 等专为长序列预测设计的 Transformer 变体，以及图神经网络（GNN）用于多城市联合建模。同时，探索概率预测方法（如 DeepAR、概率 Transformer）以提供预测不确定性估计，增强预警的可信度。
4. **系统功能增强与部署**：完善预警推送机制（短信、微信小程序、社区广播），建立多级预警响应流程（蓝/黄/橙/红四级联动），添加用户角色管理和操作日志，支持多城市切换和历史预警复盘分析。
5. **研究区域推广**：将本研究的管线和方法推广至河南省其他城市乃至整个中原城市群，为更大范围内的银发群体提供高温健康防护服务。

致谢

衷心感谢导师在选题、研究方法、论文撰写等方面给予的悉心指导和宝贵建议。
感谢河南理工大学计算机科学与技术学院四年来提供的学习平台和科研环境。
感谢家人和朋友在学业期间的理解、支持与鼓励。

附录 A 核心代码清单

本研究核心代码已开源至 Gitea 仓库 `git@lhy-git.liuhangyv.top:Serendipity/elderly-heat-`
项目采用模块化结构:

```
src/  
  data/  
    download_era5.py      # ERA5 数据下载 (CDS API)  
    extract_zips.py       # NetCDF ZIP 解压  
    preprocess.py         # 数据预处理管线  
    collect_mortality.py  # 死亡率数据整理  
  models/  
    lstm_attention.py     # LSTM-Attention 模型定义  
    xgboost_baseline.py   # XGBoost 基线  
    train.py              # 训练脚本  
    evaluate.py           # 评估脚本  
  web/  
    app.py                # Flask 后端  
    static/  
      index.html          # ECharts 前端大屏  
  utils/  
    config.py             # 全局配置
```

附录 B 系统运行说明

B.1 环境配置

- Python 3.13
- PyTorch 2.12.0+cu126 (GPU 版本, CUDA 12.6)
- XGBoost 2.0+, Scikit-learn 1.3+
- Flask 3.0+, ECharts 5.5
- 环境管理: uv (虚拟环境.venv)

B.2 运行步骤

1. 安装依赖: `uv pip install -e .`
2. 下载 ERA5 数据: `python -m src.data.download_era5`
3. 解压 ZIP 格式数据: `python -m src.data.extract_zips`
4. 运行预处理: `python -m src.data.preprocess`
5. 训练模型: `python -m src.models.train (LSTM)`
6. 评估模型: `python -m src.models.evaluate (含 XGBoost)`
7. 启动可视化大屏: `python -m src.web.app`
8. 浏览器访问: `http://localhost:5005`