

# 中国科学技术大学

# 硕士学位论文



## 数据驱动的呼吸机使用辅助决策方 法研究

作者姓名： 席嘉滢

学科专业： 控制科学与工程

导 师： 赵云波教授

完成时间： 二〇二六年五月二十三日



University of Science and Technology of China

A dissertation for master's degree



# **Data-Driven Ventilator Assisted Decision-Making Method Research**

Author: Xi Jiaying

Speciality: Control Science and Engineering

Supervisor: Prof. Zhao Yunbo

Completion date: May 23, 2026



## 中国科学技术大学学位论文原创性声明

本人声明所呈交的学位论文，是本人在导师指导下进行研究工作所取得的成果。除已特别加以标注和致谢的地方外，论文中不包含任何他人已经发表或撰写过的研究成果。与我一同工作的同志对本研究所做的贡献均已在论文中作了明确的说明。

作者签名：\_\_\_\_\_

签字日期：\_\_\_\_\_

## 中国科学技术大学学位论文授权使用声明

作为申请学位的条件之一，学位论文著作权拥有者授权中国科学技术大学拥有学位论文的部分使用权，即：学校有权按有关规定向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅，可以将学位论文编入《中国学位论文全文数据库》等有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。本人提交的电子文档的内容和纸质论文的内容相一致。

控阅的学位论文在解除后也遵守此规定。

公开 控阅（\_\_\_\_年）

作者签名：\_\_\_\_\_

导师签名：\_\_\_\_\_

签字日期：\_\_\_\_\_

签字日期：\_\_\_\_\_



## 摘要

利用呼吸机进行机械通气是危重症患者维持生命的重要治疗手段，该过程中准确及时地对呼吸机进行各项设定是保障治疗效果的前提要求。当前呼吸机在使用中主要由医生人工设定，但受限于部分医生专业水平不足、操作中生理和精神状态不佳等主客观条件，在临床实践中往往难以达到理想的机械通气效果。为此，如果能在临床实践中实时为医生提供呼吸机设定的准确辅助决策建议，将有望辅助医生做出更优的呼吸机设定决策，具有重要的临床应用价值。

面对上述需求，本文围绕机械通气的呼吸机启用时机、使用时长及推荐参数三个呼吸机设定中的核心问题，展开了较为系统的研究，具体工作和创新点如下：

(1) 针对呼吸机启用时机的决策问题，提出了一种融合临床约束规则的集成决策模型，为患者使用呼吸机的时机提供了科学的决策依据。在现有数据驱动方法基础上，将插管指标以先验规则形式引入该模型中，完成对患者的快速筛查，在此基础上结合患者的多维度生理状态以及集成模块，进而实现对呼吸机需求的概率预测，根据设定的阈值输出呼吸机的启用决策。

(2) 针对呼吸机使用时长的预测问题，提出了一种融合静态与动态特征的时序预测模型，实现了对呼吸机使用时长较高精度的预测。该模型综合考虑患者入院时的生理状态与治疗过程中的动态演变指标，通过创新性地结合特征回归筛选模块及改进的时序预测模块，从而提取患者的状态特征和时序特征，有效提高呼吸机使用时长的预测精准度。

(3) 针对呼吸机输出参数的推荐问题，提出了一种嵌入专家知识的参数推荐模型，提供了性能表现较好的参数推荐方案。该模型在深度离线强化学习模块中创新性地引入一种融合动脉血气分析的增强型优化结构，进而实现临床先验知识的嵌入，完成数据驱动与临床先验的有效结合，最终生成具有一定临床适用性的治疗策略推荐。

上述模型和方法在大型公开重症监护数据集 MIMIC-III 上进行了实验验证。结果表明，所提出的三种模型在多个评价指标上优于所选对比模型，表现出较好的性能。本文工作为呼吸机全程管理提供了系统性的数据驱动解决方法，研究成果有望辅助临床医生在危重症救治中做出更精准、更科学的决策，提升治疗质量与资源利用效率。

**关键词：**呼吸机；机械通气；机器学习；时序预测；强化学习；专家知识

## ABSTRACT

Mechanical ventilation is a vital life-sustaining therapeutic intervention for critically ill patients, wherein accurate and timely parameter settings constitute a prerequisite for ensuring therapeutic efficacy. Currently, ventilator settings in clinical practice are primarily determined manually by physicians. However, constrained by subjective and objective factors such as disparities in professional competence and suboptimal physiological or mental states during operation, achieving ideal mechanical ventilation outcomes remains challenging. Consequently, providing real-time, accurate decision support for ventilator settings holds significant clinical value by assisting physicians in optimizing setting decisions.

Addressing these needs, this thesis conducts systematic research on three core issues regarding ventilator settings: the timing of initiation, duration of use, and recommended parameters. The specific contributions and innovations are as follows:

(1) Addressing the decision-making problem regarding the timing of ventilator initiation, an integrated decision model incorporating clinical constraint rules is proposed. This model provides a scientific basis for determining the timing of ventilator usage. Building upon existing data-driven methods, intubation indicators are introduced as prior rules to achieve rapid patient screening. Subsequently, by integrating the patient's multi-dimensional physiological status with an ensemble module, the model predicts the probability of ventilator demand and outputs the initiation decision based on a set threshold.

(2) Addressing the prediction problem of ventilator usage duration, a time-series prediction model fusing static and dynamic features is proposed to achieve more precise prediction. This model comprehensively considers the patient's physiological state at admission and the dynamic evolution of indicators during treatment. By innovatively integrating a feature regression screening module and an improved time-series prediction module, the model effectively extracts patient state features and temporal features, thereby enhancing the prediction accuracy of ventilator usage duration.

(3) Addressing the recommended problem of ventilator parameter, a parameter recommendation model embedding expert knowledge is proposed to provide superior parameter recommendation schemes. In this model, an enhanced optimization structure integrating arterial blood gas analysis is innovatively introduced into a deep offline reinforcement learning module. This facilitates the embedding of clinical prior knowl-

edge, achieving an effective fusion of data-driven approaches and clinical priors, and ultimately generating treatment strategy recommendations with clinical applicability.

The aforementioned models and methods were experimentally validated on the large-scale public Intensive Care Unit database, MIMIC-III. The results demonstrate that the proposed models outperform selected comparative models across multiple evaluation metrics, exhibiting superior performance. This thesis provides a systematic data-driven solution for the whole-process management of ventilators. The findings are expected to assist physicians in making more precise and scientific decisions during the treatment of critically ill patients, thereby improving treatment quality and resource utilization efficiency.

**KEY WORDS:** Ventilator, Mechanical Ventilation, Machine Learning, Time Series Prediction, Reinforcement Learning, Expert Knowledge

## 目录

<b>第 1 章 绪论 .....</b>	<b>1</b>
1.1 课题研究背景及意义 .....	1
1.2 国内外研究现状 .....	2
1.2.1 呼吸机启用时机决策研究现状 .....	2
1.2.2 呼吸机使用时长预测研究现状 .....	3
1.2.3 呼吸机输出参数推荐研究现状 .....	4
1.2.4 时序预测算法研究现状 .....	5
1.2.5 离线强化学习算法研究现状 .....	6
1.3 研究内容及组织架构 .....	7
1.3.1 研究内容 .....	7
1.3.2 组织架构 .....	8
<b>第 2 章 相关基础知识 .....</b>	<b>10</b>
2.1 呼吸机 .....	10
2.2 动脉血气分析 .....	11
2.3 数据集 .....	12
2.3.1 MIMIC-III 数据集 .....	12
2.3.2 eICU-CRD 数据集 .....	13
2.4 回归模型 .....	14
2.4.1 Random Forest .....	14
2.4.2 XGBoost .....	15
2.5 聚类模型 .....	16
2.6 神经网络 .....	16
2.7 时序预测 .....	18
2.7.1 LSTM 模型 .....	18
2.7.2 基于 Transformer 的时序模型 .....	18
2.8 强化学习 .....	19
2.8.1 Q 学习和深度 Q 学习 .....	19
2.8.2 离线强化学习算法 .....	20
2.8.3 离线策略评估算法 .....	21
<b>第 3 章 基于临床约束和集成学习的呼吸机启用决策方法 .....</b>	<b>24</b>
3.1 数据处理 .....	24

3.2	模型设计 .....	26
3.2.1	分类模块 .....	27
3.2.2	类别不平衡处理模块 .....	28
3.2.3	临床约束模块 .....	30
3.3	实验设计与结果分析 .....	30
3.3.1	实验设置 .....	30
3.3.2	评价指标 .....	31
3.3.3	实验结果 .....	32
3.4	本章小结 .....	33
<b>第 4 章</b>	<b>基于特征回归筛选和时序深度建模的呼吸机使用时长预测方法 .....</b>	<b>34</b>
4.1	数据处理 .....	34
4.2	模型设计 .....	34
4.2.1	特征回归模块 .....	36
4.2.2	时序预测模块 .....	37
4.3	实验设计与结果分析 .....	39
4.3.1	实验设置 .....	39
4.3.2	评估指标 .....	42
4.3.3	实验结果 .....	45
4.4	本章小结 .....	50
<b>第 5 章</b>	<b>基于专家知识和强化学习的呼吸机参数推荐方法 .....</b>	<b>52</b>
5.1	数据处理 .....	52
5.2	模型设计 .....	52
5.2.1	聚类模块 .....	54
5.2.2	RL 问题定义 .....	54
5.2.3	离线强化学习模块 .....	56
5.3	实验设计与结果分析 .....	57
5.3.1	实验设置 .....	57
5.3.2	评估指标 .....	60
5.3.3	实验结果 .....	61
5.4	本章小结 .....	66
<b>第 6 章</b>	<b>总结与展望 .....</b>	<b>68</b>
6.1	研究总结 .....	68
6.2	研究展望 .....	68

参考文献 .....	70
致谢 .....	77
在读期间取得的科研成果 .....	78

## 插图清单

图 1.1	研究内容框架图 .....	8
图 2.1	呼吸机整体示意图 .....	10
图 2.2	呼吸机的人机交互界面示意图 .....	10
图 2.3	患者使用呼吸机进行机械通气的示意图 .....	11
图 2.4	随机森林流程图 .....	15
图 2.5	K-means 聚类过程图 .....	17
图 2.6	神经网络结构图 .....	17
图 2.7	LSTM 结构图 .....	18
图 2.8	强化学习结构图 .....	20
图 3.1	基于患者生理状态的呼吸机启用决策模型总体结构图 .....	27
图 3.2	集成学习模型流程图 .....	28
图 3.3	模型的 ROC 曲线对比图 .....	32
图 3.4	模型的 PR 曲线对比图 .....	32
图 3.5	特征重要性排序 .....	33
图 4.1	呼吸机的使用时长预测模型结构图 .....	36
图 4.2	MRA-LSTM 模块结构图 .....	38
图 4.3	各回归模型的拟合情况图 .....	45
图 4.3	各回归模型的拟合情况图 (续) .....	46
图 4.4	表现优秀的回归模型的交叉验证对比图 .....	47
图 4.5	部分预测模型的训练损失曲线 .....	48
图 4.6	RMRA-LSTM 模型拟合情况图 .....	49
图 5.1	人口统计学分布 .....	52
图 5.2	基于专家知识和强化学习的呼吸机参数推荐模型总体结构图 .....	53
图 5.3	离线强化学习框架图 .....	57
图 5.4	轮廓系数值随聚类数量变化的图 .....	62
图 5.5	聚类结果图 .....	63
图 5.6	K-CQL 算法训练过程中奖励值随训练步数的变化曲线图 .....	64
图 5.7	各模型输出的策略的机械通气动作分布图 .....	66
图 5.8	K-BCQ 和 K-CQL 在 ID 和 OOD 状态下的初始状态估计值 .....	67

## 附表清单

表 3.1	部分初始数据示例表 .....	26
表 3.2	深度学习模型的各层规模参数 .....	29
表 3.3	对比模型的 AUC-ROC 及 AUC-PR 值 .....	32
表 4.1	所提模块的时序注意力与 Transformer 自注意力对比 .....	41
表 4.2	实验环境配置参数 .....	42
表 4.3	各回归模型的性能评分 .....	47
表 4.4	各预测模型在 MIMIC-III 数据集的实验结果 .....	49
表 4.5	RMRA-LSTM 模型的消融实验结果 .....	50
表 5.1	实验环境配置参数 .....	59
表 5.2	DeepVent, DeepVent+, K-CQL- 和 K-CQL 的模块比较 .....	60
表 5.3	聚类算法的实验对比结果 .....	61
表 5.4	K-CQL 对比实验中各模型在 MIMIC-III 数据集的平均初始状态值 ....	63
表 5.5	K-CQL 对比实验中各模型在 eICU-CRD 数据集的平均初始状态值 ....	64
表 5.6	K-CQL 消融实验中各模型的平均初始状态值 .....	64
表 5.7	基于 DR 算法评估模型的输出策略的平均初始状态值 .....	65

## 第1章 绪论

本章首先简要概述了呼吸机使用辅助决策相关的背景知识与研究意义，接着对国内外相关研究现状进行难点总结和讨论，最后介绍了本文主要研究内容及组织结构。

### 1.1 课题研究背景及意义

利用呼吸机进行机械通气是重症监护病房 (Intensive Care Unit, ICU) 中为危重症患者提供生命支持的核心措施之一<sup>[1]</sup>，其广泛应用于呼吸衰竭、严重感染及术后恢复等临床场景，主要作用在于维持患者的气体交换功能、减轻呼吸肌负担并为基础疾病的治疗争取时间。然而，在临床实践中，如何合理设置呼吸机启动时机、使用时长及输出参数以确保既能提供足够的氧合支持，又能尽可能减少肺损伤等并发症<sup>[2]</sup>是一个重要的挑战。

目前呼吸机的使用主要依赖医生经验来决定启用时机、使用时长和输出参数，但这种方法在个体化精准治疗方面仍存在局限性。危重症患者通常伴有多种并发症，其病情变化迅速，难以用固定规则或统一标准来判断呼吸机的启停以及输出参数<sup>[3]</sup>，另一方面，医生若专业水平有限或因身心疲劳导致认知负荷加重，所作出的决策往往存在偏差，难以达到理想效果。因此仅凭医生经验进行呼吸机的一系列设置，稍有不慎可能导致机械通气不足、机械通气过度或者参数调整不当的问题，过早或过晚进行机械通气都可能对患者产生不良影响：过早插管可能增加气道损伤、呼吸机相关性肺炎等风险，而过晚插管则可能导致低氧血症恶化，加重器官功能损害<sup>[4]</sup>。此外，呼吸机的参数设置直接影响肺保护策略的实施，若设置不合理，可能引发通气相关性肺损伤，延长机械通气时间，从而影响患者的预后，甚至增加患者的死亡风险<sup>[5-6]</sup>。随着人工智能和医学数据的发展，数据驱动的呼吸机使用辅助决策方法研究有望辅助医生做出更合理的呼吸机设定决策。

首先，基于患者的生理状态预测患者的呼吸机启用时机具有重要意义。一方面，该预测结果能够为临床医生在患者病程早期阶段实施干预措施提供重要的决策依据，通过对生命体征等多元数据的综合评估，提前识别呼吸功能恶化的高风险患者，从而避免因延误上机而导致的缺氧、器官功能损伤等严重后果。另一方面，该预测为个性化治疗路径提供了支持，例如结合预测结果可以提前规划气道管理策略，从而提升治疗的整体性与连贯性。

其次，准确预测患者的呼吸机使用时长具有重要的临床和现实意义。一方

面,它使得医生可以提前了解患者可能需要的呼吸机使用时长,这对于促进医院医疗资源的科学配置、提升床位周转效率具有积极意义,确保呼吸机设备和高强度护理的人力资源得到合理分配,避免资源浪费或短缺。另一方面,预测结果可以为医生提供科学依据,便于判断患者病情的好坏,支持个性化治疗方案的制定,如气管造口的时间和口腔营养的启动<sup>[7-8]</sup>。因此,准确预测呼吸机使用时长不仅有助于提升医疗质量和效率,对于患者康复进程的推进以及医疗体系的长远良性发展,同样具有深远意义。

最后,探讨合适的策略进行呼吸机的参数推荐具有重要意义。最佳的呼吸机的参数设置随着个体的差异有所不同,通常是未知的<sup>[3]</sup>,需要医生根据患者实际情况进行手动调整,而医生在选择这些设置时的主客观因素如经验知识、身心状态对自己决策的准确性有直接影响,参数设置有偏差可能会加重病情,引发肺损伤、隔膜功能障碍、肺炎和氧毒性<sup>[2]</sup>,甚至会导致死亡。因而,对于多数临床医务人员而言,如何基于患者具体状况科学确定呼吸机的输出参数,是一项兼具复杂性与技术难度的工作。该过程需结合多方面临床信息进行综合分析并决策,实施难度较大。为了防止并发症并提供最佳护理,有必要对患者进行个性化的呼吸机参数设置推荐,结合当前前沿的人工智能技术是一种有效的手段。

综上所述,数据驱动的呼吸机使用辅助决策方法研究不仅能够改善患者预后、降低并发症风险,也为危重症患者的呼吸机治疗流程智能化提供了研究基础,是推动呼吸机使用向预防性、精准化方向发展的关键一环。本文所提模型可作为重症监护辅助决策模块,与呼吸机、监护仪、电子病历系统对接,实时输出上机建议、时长预估、参数推荐,不改变临床流程、不替代医生决策,具备实际部署价值。

## 1.2 国内外研究现状

### 1.2.1 呼吸机启用时机决策研究现状

Giang 等人<sup>[9]</sup>的研究表明使用患者的电子健康记录数据的机器学习模型有望用于呼吸机相关性肺炎的诊断。Zhu 等人<sup>[10]</sup>的研究表明可以成功建立模型来预测通气患者的死亡率,其中极限梯度提升 (eXtreme Gradient Boosting, XGBoost) 模型在七种机器学习模型中表现最好。Ashrafi 等人<sup>[11]</sup>的研究结果表明,先进的集成学习、细致的特征选择和有效的类不平衡处理可以显著增强呼吸机相关性肺炎引起的创伤性脑损伤病例的早期发现,其中 XGBoost 在实验所用的六种机器学习模型中表现出了最好的效果。Wu 等人<sup>[12]</sup>构建了一个实用的呼吸机相关性肺炎预测列线图,医生可以使用它来评估急性呼吸窘迫综合征患者发生呼吸机相关性肺炎的风险,并在早期阶段启动预防策略。

相较于机械通气参数推荐与时长预测等方向,目前针对呼吸机启用时机决策的研究仍相对有限。现有研究主要集中于呼吸衰竭风险评估、机械通气需求预测及 ICU 患者病情严重程度分析等相关任务,而直接面向呼吸机启用时机智能决策的研究较少。现有的研究并未聚焦于呼吸机启用时机决策的研究,且现有模型普遍未将权威临床指南、绝对插管指征等强临床约束作为核心决策前置条件,多以纯数据驱动方式输出概率,缺乏对临床禁忌、急救规范的显性编码。

## 1.2.2 呼吸机使用时长预测研究现状

准确预测呼吸机的持续通气时长的一种可行性方法是使用人工智能 (Artificial Intelligence, AI) 方法,例如机器学习 (Machine Learning, ML)。ML 是一个探索性的过程<sup>[13-14]</sup>,其核心在于对数据进行深入分析与洞察,以发现其中隐藏的结构、规律和关联。ML 算法分为两类:有监督和无监督<sup>[15]</sup>。监督学习算法可以检测潜在解释特征与已知目标结果之间的关系,适合用于 ICU 来预测临床结果<sup>[15-17]</sup>。目前已经有一些研究基于患者的状态通过机器学习预测呼吸机使用过程中的一些相关参数,例如最佳脱机时间<sup>[18-19]</sup>,呼吸机压力<sup>[20]</sup>。

Parreco 等人<sup>[21]</sup>利用机器学习模型预测患者是否需要超过 7 天的机械通气支持,验证了机器学习方法在机械通气持续时长评估中的应用潜力,但尚未实现对机械通气使用时长的连续值预测。Wang 等人<sup>[22]</sup>利用机器学习方法开发了一个可解释的模型,用于预测急性呼吸窘迫综合征 (Acute Respiratory Distress Syndrome, ARDS) 患者的机械通气的持续时长,研究表明极限梯度提升 (eXtreme Gradient Boosting, XGBoost) 模型在测试数据集中表现最稳定且准确,被选为最佳模型,其中血管增压剂、酸碱度和序贯器官衰竭评分对机械通气的持续时长预测影响最大。Sayed 等人<sup>[23]</sup>通过三种稳健的监督机器学习技术,包括轻量级梯度提升机算法 (Light Gradient Boosting Machine, LightGBM)、随机森林 (Random Forest, RF) 以及 XGBoost 用于预测机械通气持续时长,结果表明监督式机器学习可以在 ARDS 发生后的不同时期,在 ICU 中对机械通气的持续时长进行早期且准确的预测,监督式机器学习模型对优化 ICU 资源利用和成本降低具有重要意义。Schwager 等人<sup>[24]</sup>基于 RF 预测患者的机械通气时长,取得了不错的成效。Villar 等人<sup>[25]</sup>的研究表明临床医生预测机械通气时长的能力非常有限,但即便采用多层感知器等机器学习技术及诊断 72 小时数据,持续时长延长仍难以早期预测,故而还需要更多研究来确定预测机械通气持续时长的标志。

但是以上所用的预测方法都是基于患者当前的临床状态,并没有考虑利用时序预测的方法基于患者的当前及历史的生理指标进行呼吸机的持续使用时长预测。即现有模型普遍存在时序特征利用不充分、长时依赖捕捉弱等问题,容易造成预测误差。

### 1.2.3 呼吸机输出参数推荐研究现状

目前呼吸机的参数推荐方法通常依赖于比例积分微分控制, 已知是次优的<sup>[26]</sup>。Siu 等人<sup>[27]</sup>设计了闭环自适应控制器用于根据动脉二氧化碳分析自动调整呼吸机模式。Ai 等人<sup>[28]</sup>引入闭环呼吸起搏器, 通过比例积分控制器适应各种通气条件。然而, 传统的自动控制方法评估的数据特征数量有限, 在处理高维、大规模的临床数据时效率较低<sup>[29]</sup>。机器学习方法在处理高维、大量数据方面展现了其优越性, 因此, 它们正在成为医学领域处理大量高维数据的有效工具。Zhu 等人<sup>[10]</sup>利用 k 近邻、逻辑回归、决策树和极限梯度提升等机器学习方法来预测通气患者的死亡率。Akbulut 等人<sup>[30]</sup>和 Oruganti 等人<sup>[31]</sup>利用深度神经网络来计算频率、潮气量和吸入氧气分数。然而, 这些机器学习方法主要关注呼吸机参数变化与患者生命体征之间的关系, 而没有考虑呼吸机对患者的长期影响。与经典的机器学习方法相比, 强化学习算法从与环境的交互中学习, 以实现长期目标奖励最大化<sup>[32]</sup>。强化学习算法可以模拟临床中如何通过与患者的持续互动来推荐呼吸机参数, 这已经有望超越临床标准, 为在此背景下使用强化学习提供了强有力的证据。此外, 强化学习算法的机制是最大化累积奖励, 这有助于其在考虑长期影响的优化问题中表现出优越的性能<sup>[33]</sup>。Peine 等人<sup>[34]</sup>利用 Q 学习推荐呼吸机的部分参数的设置, 例如呼气末正压、吸入氧气分数和潮气量。Kondrup 等人<sup>[35]</sup>首次考虑将深度学习和强化学习相结合调整呼吸机参数。Chen 等人<sup>[36]</sup>使用长短期记忆网络 (Long Short-Term Memory, LSTM) 设计了一个仿真环境, 并基于 Soft Actor-Critic (SAC) 算法进行具体参数的建议。Yuan 等人<sup>[37]</sup>将 Transformer 与保守 Q 学习 (Conservative Q Learning, CQL) 结合提出了一个可以对患者病情进行诊断并预测最佳呼吸机参数的模型。Zhang 等人<sup>[1]</sup>等人利用强化学习算法从固定数据集中识别最优策略, 并首次考虑呼吸机参数推荐中的安全问题, 以确保治疗过程能够在安全的前提下实现疗效的最大化。

研究表明强化学习在呼吸机的机械通气方面的研究确实显示出了巨大的潜力, 与其他算法相比, 它被证明适合解决时间序列的决策推荐问题, 故而可以应用于时间序列相关的呼吸机参数设置问题<sup>[36,38]</sup>。

Peine 等人<sup>[34]</sup>和 Kondrup 等人<sup>[35]</sup>基于马尔可夫决策过程 (Markov Decision Process, MDP) 的理论框架进行问题建模, 选取了病人的包括人口统计学信息、生命体征、实验室指标和液体相关方面的多个生理指标作为状态, 其中 Kondrup 使用 APACHE-II 的变化作为奖励函数的一环<sup>[35]</sup>。临床上医生通过血气分析进行呼吸机调参, 因此, 有必要开发更合适的贴近于实际情况的基于强化学习的呼吸机参数推荐模型从而提高临床可用性。Chen 等人<sup>[36]</sup>在进行 MDP 的建模时, 考虑了血气分析状态设定, 但是没有考虑中间奖励值的合理设置。Wang 等人<sup>[39]</sup>对

患者进行初步诊断根据疾病种类等信息与呼吸机参数相关联，构建血气分析数学模型将患者的血气分析参数值是否在正常范围内作为衡量机械通气目标是否达到的依据。Ma 等人<sup>[40]</sup>基于数据驱动和知识驱动的机器学习方法实时预测重症监护患者的器官衰竭，即引入了医生的先验知识。

但是以上方法均没有全面考虑临床知识在模型全流程中的深度融合，尤其缺乏将患者血气分析指标作为呼吸机参数设置核心依据的显性化、结构化嵌入，未能把临床指南中的专家经验转化为模型可执行的约束与目标，导致模型学习与决策过程脱离真实临床调参逻辑，推荐的参数方案虽在数据层面表现较优，却难以满足危重症患者个体化通气治疗的规范性与实用性要求。

#### 1.2.4 时序预测算法研究现状

时序预测是通过分析历史时间序列数据的内在规律，对未来时刻数值或状态进行推断的核心技术，广泛应用于气象预报、金融分析、工业监控及医疗健康等多个领域，为各类场景的决策制定提供重要数据支撑<sup>[41]</sup>。

时序预测算法主要分为传统统计、浅层机器学习、深度学习三大方向，其中深度学习凭借强大的特征提取能力，成为当前复杂时序数据预测的核心技术。

传统算法以自回归滑动平均模型 (Auto-Regressive Moving Average Model, ARMA)、自回归积分滑动平均模型 (Autoregressive Integrated Moving Average Model, ARIMA) 等为代表，核心依赖数据平稳性假设，具有原理简单、可解释性强的优势，但无法适配复杂非线性时序数据，在长序列预测中精度不足<sup>[42]</sup>。

浅层机器学习算法包括支持向量机，随机森林等，可捕捉时序数据中的简单非线性关系，且可解释性较强，但依赖人工构造特征，对长序列、高维数据的适配能力较弱。

深度学习是当前时序预测的研究热点，LSTM、门控循环单元 (Gate Recurrent Unit, GRU) 通过门控机制缓解梯度消失问题，可有效捕捉长时依赖<sup>[43]</sup>；Transformer 及其变体<sup>[44-45]</sup>则凭借独特结构，更适配长序列、多维度时序预测场景。

混合模型通过融合不同类型算法的优势，兼顾各模型长处，有效弥补单一模型的局限性，是近年来的重要研究方向。

尽管时序预测算法已取得显著进展，但在临床应用中仍存在关键不足：一是时序特征利用不充分，难以挖掘多尺度、多维度特征的协同作用；二是长时依赖捕捉能力有限，LSTM 存在长时依赖衰减，Transformer 类模型计算复杂度高、适配小样本临床数据能力弱。

### 1.2.5 离线强化学习算法研究现状

离线强化学习的核心挑战在于解决静态数据集与学习策略之间的分布偏移问题，由此引发的价值过高估计是算法设计的关键难点。针对这一问题，近年来学术界涌现了大量研究成果，主要集中在以下几个层面：

#### (1) 价值函数正则化与不确定性估计

该类算法侧重于修正价值函数的估计方式，以缓解分布外动作的过高估计。Kumar 等人提出的 CQL<sup>[46]</sup> 是该领域的基石性工作，通过引入保守正则化项强制降低分布外动作的 Q 值估计。在此基础上，近年来的研究进一步结合了不确定性量化。例如，Kostrikov 等人提出的隐式 Q 学习 (Implicit Q-Learning, IQL) 算法<sup>[47]</sup> 创新性地利用期望回归来避免对未见动作的价值评估，仅从数据集中的动作中提取价值信息，在无需显式行为策略的情况下实现了优异的性能。

#### (2) 策略约束与生成式模型

传统的策略约束算法通过显式限制策略偏差来保证安全性，而近年的研究趋势是将生成模型引入离线强化学习，以增强策略的表达能力。Fujimoto 等人提出的 Batch-Constrained Q-learning (BCQ) 算法<sup>[48]</sup> 通过生成模型逼近行为策略，将动作限制在数据集覆盖的范围，之后该团队提出的 TD3+BC 算法<sup>[49]</sup> 通过在标准强化学习目标中加入简单的行为克隆项，以极简的方式实现了策略约束，证明了简单的正则化在离线设定下的有效性。更进一步，Chen 等人提出了扩散可信 Q 学习 (Diffusion Trusted Q-Learning, DTQL) 模型<sup>[50]</sup>，利用扩散模型强大的多模态建模能力来拟合行为策略，通过将策略学习转化为去噪扩散过程，不仅有效解决了分布偏移问题，还能捕捉数据集中复杂的多模态行为分布。

#### (3) 基于序列建模的方法

不同于上述基于动态规划的方法，近年兴起了一类将离线强化学习转化为序列建模问题的新范式。Chen 等人提出的决策 Transformer<sup>[51]</sup> 摒弃了传统的价值函数估计，而是利用 Transformer 架构对状态、动作和回报序列进行建模，通过自回归预测直接生成高回报动作。该方法有效规避了传统离线强化学习中的价值高估问题，为解决分布偏移提供了全新的视角。

尽管离线强化学习算法已取得一定进步，但在应用于呼吸机参数推荐等复杂临床场景时，仍存在以下不足：

一是离线强化学习方法对训练数据分布具有较强依赖性。当策略生成的数据分布偏离训练数据覆盖范围时，模型在分布外 (Out-of-Distribution, OOD) 状态一动作对上的估计误差可能增大，从而影响策略学习的稳定性与安全性。尤其在医疗场景中，由于患者状态复杂且数据分布往往存在长尾现象，模型可能在缺乏充分数据支持的状态区域产生不可靠的动作输出，因此如何提高离线强化学习

在分布外状态下的鲁棒性，仍是当前研究中的重要挑战。

二是缺乏对领域知识的深度利用。大多数前沿离线强化学习算法属于纯数据驱动模式，难以将临床指南、生理学机制（如血气分析逻辑）等专家先验知识有效地嵌入模型训练过程中。这导致模型的学习目标与真实的临床决策逻辑可能存在偏差，生成的策略虽然数值上较优，但缺乏临床可解释性或违背医疗常识。

## 1.3 研究内容及组织架构

### 1.3.1 研究内容

针对以上提到的研究不足，本研究从数据驱动与临床知识融合的角度出发，充分利用危重症患者多维度生理时序数据，分别构建启用时机决策、使用时长预测、输出参数推荐三种模型，形成覆盖呼吸机全流程的辅助决策方法，为临床医生提供科学、安全的决策支持。

由于机械通气治疗需严格遵循临床指南与急救规范，本研究所有模型均以不替代医生决策、仅提供辅助建议为原则，在满足临床安全约束与诊疗逻辑的前提下开展方法设计，重点解决启动时机判断不准、时长预测精度不足、参数推荐缺乏临床知识嵌入等关键问题。

本文具体包括以下三部分研究内容：

(1) 针对呼吸机启用时机难以精准判断、现有模型缺乏临床硬约束的问题，提出融合临床约束规则的集成决策模型，实现科学、安全的呼吸机启用决策。首先，对 MIMIC-III 重症数据集进行多维度生理特征提取与缺失值预处理；其次，将权威指南中的绝对插管指征编码为临床先验规则，实现高危患者快速筛查；然后，构建集成学习分类模块并引入类别不平衡处理策略，提升对少数类样本的识别能力；最后，通过 ROC、PR 曲线等指标验证模型。

(2) 针对呼吸机使用时长预测精度不足、时序特征利用不充分的问题，提出融合静态特征与动态时序的预测模型 RMRA-LSTM，实现呼吸机使用时长的较高精度动态预测。首先，构建包含患者入院基线与动态时序指标的数据集；其次，通过随机森林回归模块完成关键特征筛选与初步时长预测；然后，设计多尺度 CNN、残差 LSTM 与时序注意力机制，精准捕捉病情动态演变规律；最后，在数据集上完成对比实验与消融实验验证模型的有效性。

(3) 针对呼吸机推荐参数缺乏临床知识、安全性与适用性不足的问题，提出嵌入专家知识的参数推荐模型 K-CQL，输出符合临床规范的呼吸机参数推荐。首先，基于动脉血气分析指标使用 K-means 聚类实现患者亚群划分，将临床先验嵌入数据结构；其次，设计血气分析与 APACHE-II 评分相结合的复合奖励函数，

兼顾短期稳定与长期预后；然后，采用保守 Q 学习抑制分布外动作与 Q 值高估，提升策略安全性；最后，通过 FQE、DR 等离线评估方法验证所提模型的有效性。

### 1.3.2 组织架构

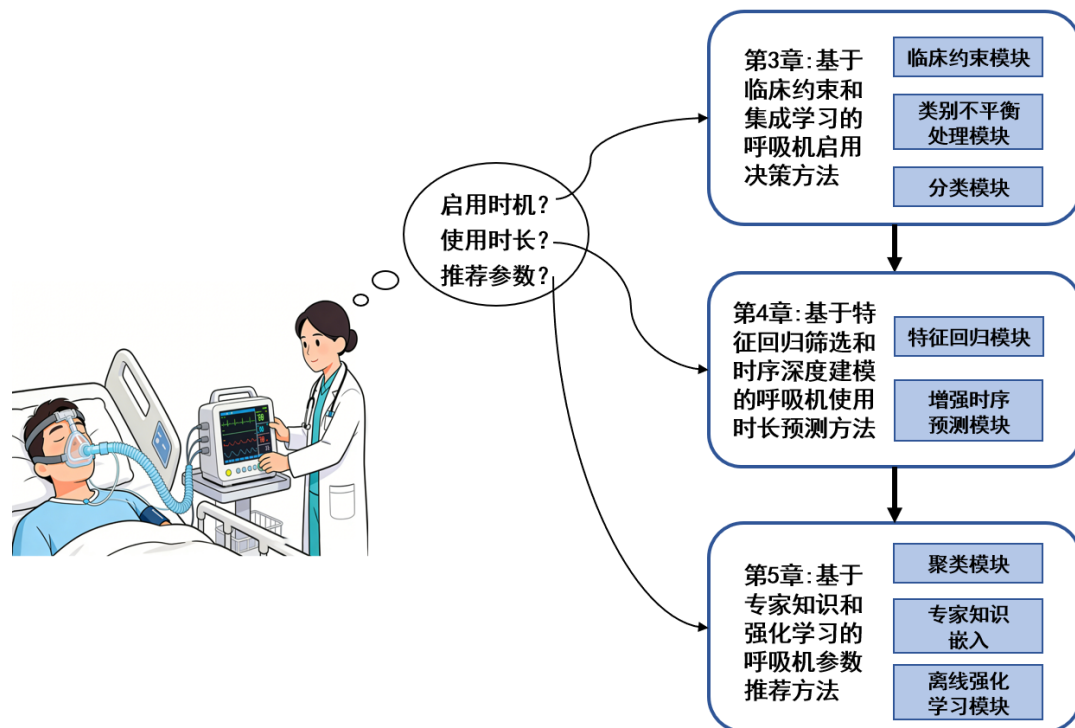


图 1.1 研究内容框架图

论文的研究内容框架如图1.1所示，组织结构共有六章，具体描述如下：

**第1章：绪论。**介绍呼吸机相关的研究背景，阐明对危重症患者进行机械通气辅助决策的意义。同时对现有呼吸机的启用情况决策、使用时长预测和重要参数推荐研究现状予以总结，基于现有研究中存在的问题，给出论文的主要研究内容和组织架构。

**第2章：相关基础知识。**介绍论文涉及到的基础知识以及使用的一些技术方法，包括呼吸机，动脉血气分析，MIMIC-III 数据集，回归模型，聚类模型，神经网络，时序预测以及强化学习。

**第3章：基于临床约束和集成学习的呼吸机启用决策方法。**首先对公开的医学数据集深入分析进行一系列的处理，然后设计临床约束规则将明确、无争议的插管指征编码为确定性规则，接着设计基于机器学习的二分类决策模块，最后通过测试集数据进行对比实验，验证模型的有效性和优越性。

**第4章：基于特征回归筛选和时序深度建模的呼吸机使用时长预测方法。**首先构建结合患者静态生理状态及动态时序变化的数据集，接着设计相关的回归模块和时序预测模块，最后通过对比实验、消融实验及各项指标评估验证模型有

效性和优越性。

第5章：基于专家知识和强化学习的呼吸机参数推荐方法。首先构建融合医学知识的聚类模块，接着基于临床先验知识定义离线强化学习相关的状态，动作，奖励，最后通过对比实验、消融实验及多指标评估验证该模型的有效性、优越性和安全性。

第6章：总结与展望。对论文的研究内容进行全面梳理与归纳，并在此基础上深入剖析当前研究工作中存在的局限与不足之处，以及进一步的改进方向。

## 第2章 相关基础知识

本章主要介绍了该论文的相关基础知识，包括呼吸机，动脉血气分析，MIMIC-III 数据集，回归模型，聚类模型，神经网络，时序预测以及强化学习，为该论文的后续研究提供理论基础。

### 2.1 呼吸机

呼吸机作为一种关键的生命支持设备（如图 2.1及图 2.2所示），其设计和工作原理基于对人体呼吸生理的深入理解。



图 2.1 呼吸机整体示意图



图 2.2 呼吸机的人机交互界面示意图

人体呼吸过程包含三个核心环节：外呼吸（即肺与外界环境之间的气体交换）、气体在血液中的运输以及内呼吸（组织细胞与血液之间的气体交换）<sup>[52]</sup>。当患者因疾病、创伤或麻醉等原因导致自主呼吸功能不全或衰竭时，呼吸机通过机械方式介入外呼吸环节，模拟并辅助或替代人体的自然通气功能（如图 2.3所示）。其核心机制在于，呼吸机在患者的肺泡与外界大气环境之间人为建立一个可控的压力梯度，从而驱动气体流入或流出肺部，完成强制的人工通气过程。

现代呼吸机是一个复杂的集成系统，其构成远不止提供气源的动力部分。一个完整的呼吸机系统主要包括气源（如医用空气与氧气）、一套精密的气体输送与调控系统，以及至关重要的电子监控与反馈系统。气体输送系统负责将混合气体以设定的压力、容量和流量模式送达患者气道，并通常集成有加湿、加热及细菌过滤功能，以保护气道黏膜并降低感染风险。而电子监控系统则是现代呼吸机智能化与安全性的基石，它持续监测多项关键参数，如气道压力、潮气量、呼吸频率、氧浓度等，并通过内置算法实时调整输出或发出警报，确保通气过程的安全。

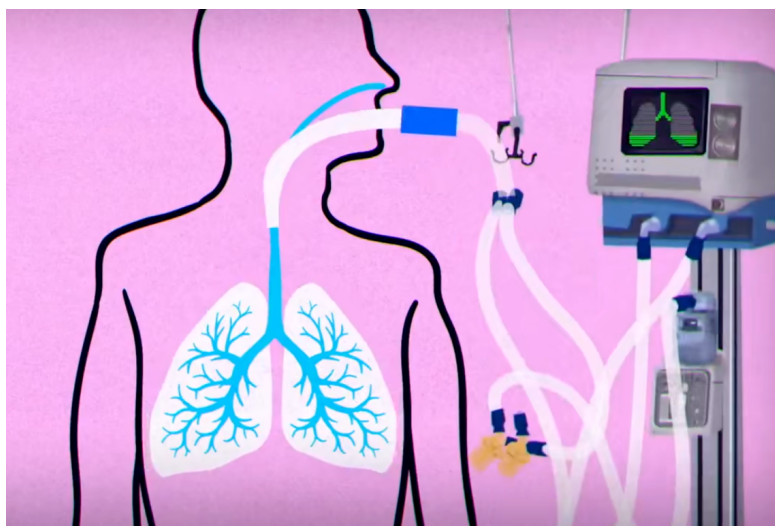


图 2.3 患者使用呼吸机进行机械通气的示意图

全、精准与适配。

在临床应用中，呼吸机的主要目标是维持有效的肺泡通气和氧合，同时尽量减少因机械通气本身带来的潜在损伤，如呼吸机相关性肺损伤。从简单的定时压力/容量控制，到能够感知患者微弱吸气努力并给予同步比例支持的智能化模式，呼吸机技术的发展始终围绕着使机械通气更符合生理、更个体化的目标。

## 2.2 动脉血气分析

动脉血气分析是对动脉血液中氧气和二氧化碳分压、酸碱度及相关衍生指标进行定量测定的临床检验技术，它为评估人体呼吸功能与内环境稳态提供了至关重要的即时数据。在机械通气治疗中，动脉血气分析是指导通气参数设置、评估治疗效果、预防并发症及判断撤机时机的核心客观依据。

机械通管理主要关注动脉血气分析中的以下几组关键指标：

### (1) 氧合指标：

- 动脉血氧分压 ( $\text{PaO}_2$ )：作为直接衡量动脉血中物理溶解氧气压力的指标，能够直观反映肺换气功能的状态。在机械通气中， $\text{PaO}_2$  的水平是调整吸入氧浓度 (Fraction of inspired oxygen,  $\text{FiO}_2$ ) 和呼气末正压 (Positive End Expiratory Pressure, PEEP) 的首要依据。

- 动脉血氧饱和度 ( $\text{SaO}_2$ )：定义为血红蛋白与氧气相结合的百分比，与  $\text{PaO}_2$  通过氧解离曲线相关联，是监测氧输送效率的常用指标。

- 氧合指数 ( $\text{PaO}_2/\text{FiO}_2$ )：该比值在排除了  $\text{FiO}_2$  影响后，能更准确地反映肺本身的氧合效率，是诊断 ARDS 严重程度及评估其治疗效果的核心标准（如轻度：200-300 mmHg，中度：100-200 mmHg，重度： $<100$  mmHg）。

### (2) 通气与酸碱平衡指标：

- 动脉血二氧化碳分压 ( $\text{PaCO}_2$ ): 评估肺泡通气量的标准。当  $\text{PaCO}_2$  水平上升时, 通常表明肺泡通气量不足, 此时需通过调整潮气量等方式增加分钟通气量以改善通气状况; 反之,  $\text{PaCO}_2$  下降则提示存在通气过度。

- 酸碱度 (pH): 综合反映机体酸碱平衡状态。机械通气直接影响  $\text{PaCO}_2$ , 从而通过改变血液中碳酸浓度来快速影响 pH 值)。

- 碳酸氢根浓度 ( $\text{HCO}_3^-$ ) 与碱剩余 (BE): 主要反映代谢性因素对酸碱平衡的贡献, 有助于鉴别混合型酸碱失衡。

动脉血气分析结果直接驱动着机械通气的闭环管理, 具体场景包括如下:

(1) 初始参数设定与目标导向: 根据初始动脉血气结果 (如严重的低氧血症伴呼吸性酸中毒), 临床医生可快速制定通气策略。例如, 对 ARDS 患者, 目标是通过肺复张和最佳 PEEP 策略提高  $\text{PaO}_2/\text{FiO}_2$ , 同时采用小潮气量通气防止肺损伤, 允许一定程度的高碳酸血症以保护肺组织。

(2) 动态调整与反馈优化: 通气开始后, 需定期复查血气以评估对初始设置的反应。例如, 当  $\text{PaCO}_2$  下降过快引发呼吸性碱中毒时, 可考虑通过降低呼吸频率或潮气量来实施干预; 若氧合改善不理想, 则需考虑优化 PEEP 或调整  $\text{FiO}_2$ 。

(3) 撤机决策的支持: 在准备撤机前后, 动脉血气分析是判断患者能否维持有效自主呼吸的关键。成功的撤机试验要求患者在脱离呼吸机支持期间, 能保持基本稳定的  $\text{PaO}_2$ 、 $\text{PaCO}_2$  和 pH 值, 无显著的呼吸性酸中毒或低氧血症恶化。

因此动脉血气分析不仅是机械通气临床实践中不可或缺的监测工具, 其数据本身也是连接患者复杂生理状态与呼吸机工程参数之间的桥梁, 为高级别的智能预测与研究提供了核心的生理学变量。

## 2.3 数据集

### 2.3.1 MIMIC-III 数据集

MIMIC-III 数据集<sup>①</sup> (Medical Information Mart for Intensive Care III) 由麻省理工学院计算生理学实验室构建, 它是一个大型、免费的数据集, 由六万多名病人的健康数据信息构成<sup>[53]</sup>。这些病人来自贝斯以色列女执事医疗中心的 ICU, 记录时间为 2001 年至 2012 年。

该数据库整合了多维度的临床信息, 主要涵盖患者的人口学特征、病房内采集的生命体征数据、各类实验室检测结果, 以及诊疗措施、用药情况、护理记录、影像学报告和结局信息 (包括院内及出院后的死亡情况) 等内容。

该资源为多领域医学研究提供了重要支撑, 应用范围涉及流行病学分析、临床决策模型的构建与优化, 以及医疗信息化工具的研发等方面。其优势主要体现

<sup>①</sup>MIMIC-III 数据集可通过 <https://mimic.physionet.org/> 获取

在以下三个方面：

- 数据获取具有开放性，可为全球科研人员共享使用。
- 覆盖患者类型丰富，样本规模较大，尤其集中于重症监护人群。
- 包含大量连续时间序列信息，如检验指标、电子病历文本以及床旁监测数据等，为动态分析提供了良好基础。

该数据集共包括 26 张表格，可划分为四类，分别如下：

- 字典信息数据表。
- 患者及其入院情况的信息表。
- 在重症监护病房中收集的患者数据信息表。
- 医院记录系统收集的数据信息表。

MIMIC-III 数据集的发布标志着医疗数据科学领域的一个重要里程碑。它首次将大量去识别化的临床数据公开，为研究人员提供了前所未有的资源，以探索和开发新的医疗诊断和治疗模型。此外，MIMIC-III 的发布也促进了跨学科的合作，特别是在机器学习和人工智能领域，推动了这些技术在医疗保健中的应用。

当前，MIMIC-III 数据集已成为全球医疗研究的重要基石，被广泛应用于各种研究项目中，包括疾病预测、患者分层和治疗效果评估等。随着技术的进步，MIMIC-III 的数据也被用于开发和验证新的算法和模型，进一步提升了其在医疗决策支持系统中的应用价值。

### 2.3.2 eICU-CRD 数据集

eICU-CRD 数据集<sup>①</sup> (eICU Collaborative Research Database) 由飞利浦医疗与麻省理工学院计算生理学实验室联合发布，是一个大型、多中心的重症监护数据库<sup>[54]</sup>。该数据集收集了 2014 年至 2015 年间美国多家医院重症监护病房的患者临床数据，共包含超过 20 万次 ICU 住院记录，覆盖美国多个地区及不同类型医疗机构。

该数据库整合了 ICU 患者在住院期间产生的大量临床信息，主要包括患者人口学特征、生命体征、实验室检查结果、诊断信息、护理记录、药物使用情况、呼吸机治疗信息以及疾病评分等内容。同时，eICU-CRD 数据集还包含较为丰富的时间序列监测数据，可用于患者病情动态变化分析与临床决策建模。

相较于单中心 ICU 数据集，eICU-CRD 的主要特点在于其多中心数据来源，能够在一定程度上反映不同医院之间患者分布、治疗模式及临床实践的差异。因此，该数据集被广泛应用于临床风险预测、疾病分型、强化学习辅助决策以及医疗人工智能模型泛化能力评估等研究方向。

其优势主要体现在以下几个方面：

<sup>①</sup>eICU-CRD 数据集可通过 <https://physionet.org/content/eicu-crd/2.0/>获取

- 数据来源覆盖多个医疗中心，具有较好的患者异质性与临床多样性。
- 样本规模较大，包含丰富的 ICU 患者监测与治疗信息。
- 包含大量连续时间序列数据，适用于时序预测与动态决策建模任务。
- 支持对模型跨医院泛化能力与鲁棒性的研究与验证。

eICU-CRD 数据集的公开发布，为医疗人工智能领域提供了重要的数据资源支持，尤其在多中心临床研究及模型外部验证方面具有重要价值。

## 2.4 回归模型

多变量回归模型在健康科学研究中有着广泛的应用<sup>[55]</sup>。作为机器学习领域中用于探究自变量与因变量关联机制的重要方法，回归模型主要划分为线性回归与非线性回归两大类，主要用于预测和解释数据。回归分析的核心目的是通过建立数学模型来描述自变量和因变量之间的内在关系，通常使用均方误差 (Mean Square Error, MSE)、决定系数 (Coefficient of Determination,  $R^2$ ) 等作为评估指标。由于患者的状态和呼吸机通气时长不存在线性关系，所以本研究使用的是非线性回归模型。

### 2.4.1 Random Forest

随机森林 (Random Forest, RF) 是一种基于决策树的集成学习算法，其核心机制在于确保每一棵决策树的构建都依赖于独立采样的随机向量，且森林中所有树均服从相同的分布<sup>[56]</sup>，具体流程如图 2.4 所示。随机森林具有较高的回归精度，能够很好地容忍离群值和噪声，并且不会出现过拟合<sup>[57]</sup>，它已经成为数据挖掘领域和生物信息领域最热门的研究方法之一。

随机森林训练过程如下：

1. Bootstrap 采样：每棵树从原始训练数据集  $D = \{(x_i, y_i)\}_{i=1}^N$  进行有放回采样，生成样本集  $D_t$ 。

2. 随即特征选择：每次分裂节点时，从全部特征中随机选择  $m \ll d$  个特征（如果有  $d$  个总特征），计算某种分裂标准（如最小化 MSE）来选择最佳分裂点。

3. 每棵树单独训练：构建完全生长的回归树，不剪枝。

假设训练了一个随机森林回归模型，包含  $T$  棵回归树。对一个输入样本  $x$ ，每棵树  $h_t(x)$  给出一个实数预测值，则最终预测值为：

$$\hat{y} = \frac{1}{T} \sum_{t=1}^T h_t(x) \quad (2.1)$$

其中  $\hat{y}$  为模型对  $x$  的预测值， $T$  为森林中回归树的总数， $h_t(x)$  为第  $t$  棵树对  $x$  的预测值。

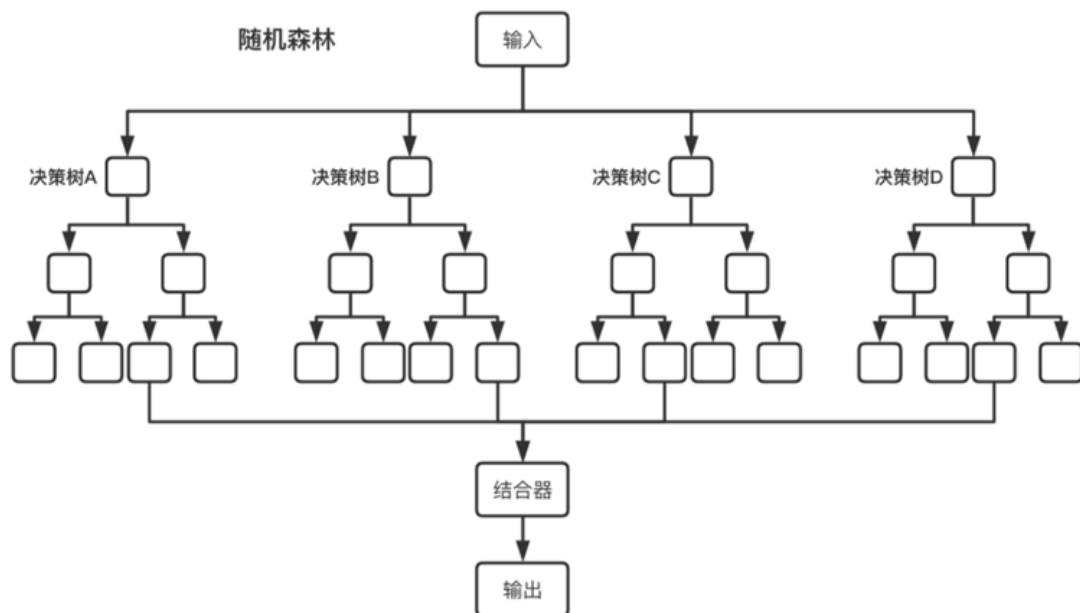


图 2.4 随机森林流程图

### 2.4.2 XGBoost

极限梯度提升 (eXtreme Gradient Boosting, XGBoost) 是一种建立在梯度提升思想之上的集成学习方法<sup>[58]</sup>, 方法通常以决策树作为基学习器, 通过逐步叠加多个弱模型来构建性能更优的预测模型。在处理结构化数据相关的回归与分类问题时, XGBoost 的泛化能力和预测效果均表现突出。

XGBoost 的核心是梯度提升决策树, 但进行了多项重要改进:

- 并行计算: 利用 CPU 多核进行并行处理。
- 正则化: 约束模型复杂度、抑制过拟合。
- 缺失值处理: 自动学习缺失值的处理方式。
- 剪枝策略: 更有效的树生长停止策略。

对于包含  $n$  个样本的数据集, XGBoost 的预测函数为:

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i), \quad f_k \in \mathcal{F} \quad (2.2)$$

其中  $K$  为树的数量,  $f_k$  为第  $k$  棵决策树,  $\mathcal{F}$  为所有可能的决策树空间。

采用加法训练方式, 从常数预测开始, 每次添加一棵树优化目标函数, 第  $t$  步的预测为:

$$\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + f_t(x_i), \quad t = 1, 2, \dots, T \quad (2.3)$$

XGBoost 的目标函数包含以下两部分:

$$\text{Obj}(\theta) = \sum_{i=1}^n L(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k) \quad (2.4)$$

其中第一部分是损失函数，用于评估模型预测值与实际值之间的偏离程度。第二部分是正则化项，用于约束模型复杂度、抑制过拟合，具体表达如下：

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 \quad (2.5)$$

其中  $T$  为叶子节点数， $w_j$  为叶子节点权重。

## 2.5 聚类模型

K-means 是一种无监督学习的聚类算法，广泛应用于各种领域的数据分析与分类场景，其主要目标是将数据对象划分为  $k$  个不同的簇，使得同一簇内部的样本在特征空间中具有较高相似度，而不同簇之间保持较明显的区分度<sup>[59]</sup>。K-means 的基本原理是通过优化目标函数，使各簇内样本到其对应中心点的平方距离之和达到最小，从而实现对数据结构的有效刻画，其工作流程如下：

1. 初始化：从数据集中随机选择  $k$  个点，将其设定为初始的聚类簇中心（即质心）。

2. 分配簇：依据各样本点与聚类中心之间的空间距离，逐一将其划归至距离最近的簇中。

3. 更新质心：对每个簇的质心进行重新计算，将其更新为该簇内全部样本点所处的平均位置。

4. 重复：重复步骤 2 和 3，直到簇中心收敛或满足预设的终止条件。

K-means 的目标函数是：

$$J = \sum_{i=1}^k \sum_{x \in C_i} \|x - \mu_i\|^2 \quad (2.6)$$

其中  $k$  是簇的数量， $C_i$  代表第  $i$  个簇， $\mu_i$  是第  $i$  个簇的质心， $x$  是簇  $C_i$  中的样本点。算法通过反复调整簇中心的位置和重新分配样本，最终使得簇内差异最小化，达到最佳聚类效果。K-means 聚类前后如图 2.5 所示。

## 2.6 神经网络

神经网络的研究最早可追溯至 1943 年，当时 McCulloch 和 Pitts 提出了第一个人工神经元模型——MP 模型<sup>[60]</sup>，这一开创性工作标志着人工神经网络的诞生。随后在 1958 年，Rosenblatt 发明了感知机，这是第一个真正意义上的神经网络模型，能够完成简单的线性分类任务<sup>[61]</sup>。1986 年，Rumelhart 等人<sup>[62]</sup>提出了反向传播算法，它成功攻克了多层神经网络难以有效训练的难题，进而推动了神经网络研究的进程。进入 21 世纪后，随着科技的进步，特别是 2012 年 AlexNet 在 ImageNet 竞赛中的突破性表现<sup>[63]</sup>，深度学习迎来了爆炸式发展。近年来，随

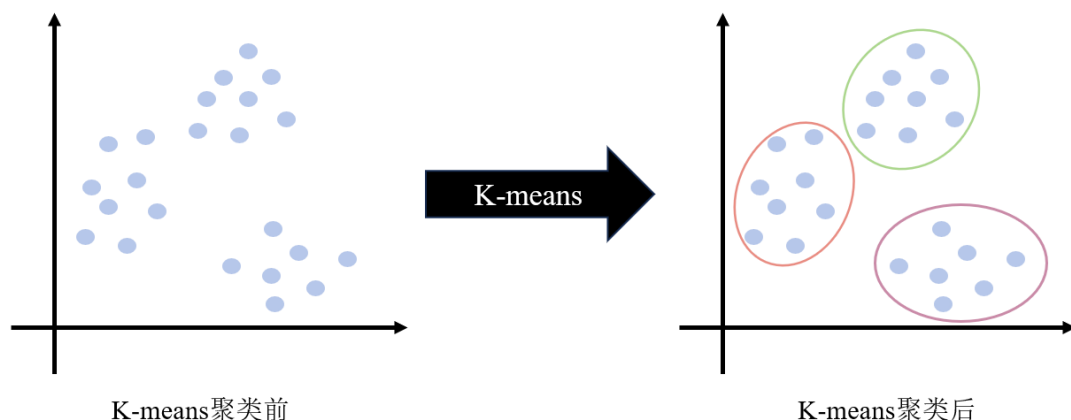


图 2.5 K-means 聚类过程图

随着 Transformer 架构、生成对抗网络、图神经网络等新型架构的不断涌现，神经网络已经发展成为人工智能领域最核心的技术之一，具体的结构图如图 2.6 所示。

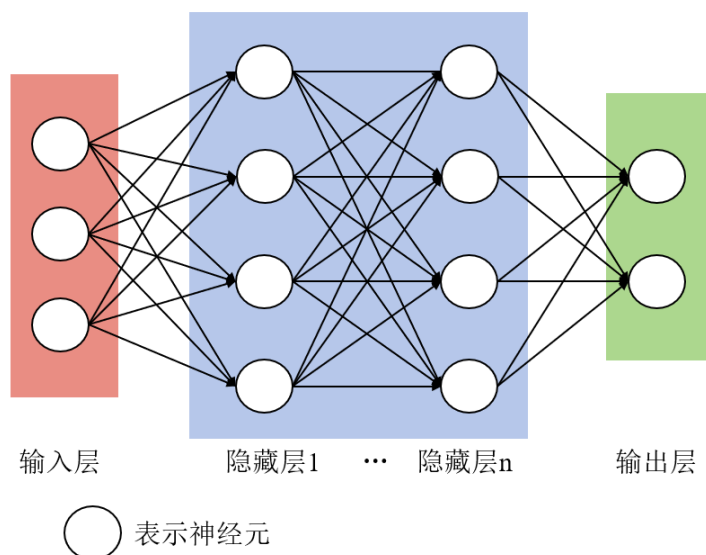


图 2.6 神经网络结构图

一个标准的人工神经元模型可由以下数学表达式完整描述：

$$y = f\left(\sum_{i=1}^n w_i x_i + b\right) \quad (2.7)$$

其中， $x_1, x_2, \dots, x_n$  表示神经元的输入信号，这些输入可以来自原始数据或前一层的输出； $w_1, w_2, \dots, w_n$  是对应的连接权重，反映了每个输入对神经元输出的影响程度； $b$  是偏置项，用于调整神经元的激活阈值； $f(\cdot)$  是激活函数，引入非线性变换能力； $y$  是神经元的最终输出。神经网络的可学习参数主要由上述所提的权重和偏置构成，这些参数在模型训练阶段经由优化算法持续迭代更新。

## 2.7 时序预测

### 2.7.1 LSTM 模型

目前的时序预测问题多基于深度学习解决。长短期记忆网络 (Long Short-Term Memory, LSTM) 是一种循环神经网络 (Recurrent Neural Network, RNN) 的变体<sup>[64]</sup>, 用于处理具有长期时间依赖关系的序列数据<sup>[65]</sup>。在多变量时间序列建模任务中, 通常将各观测指标视为独立但相关的输入序列, 模型通过刻画时间维度上的动态变化及变量间的相关结构, 对未来时刻的目标值进行预测。LSTM 在此类问题中具有显著优势, 其通过引入门控单元对信息进行筛选与更新, 从而缓解传统 RNN 在训练过程中易出现的梯度衰减与梯度不稳定现象, 而使模型对长时序依赖关系的建模与学习能力得到显著提升<sup>[66]</sup>, 该模型的记忆单元可以处理和存储先前的信息, 并根据需要更新和使用这些信息, 具体结构如图 2.7 所示。

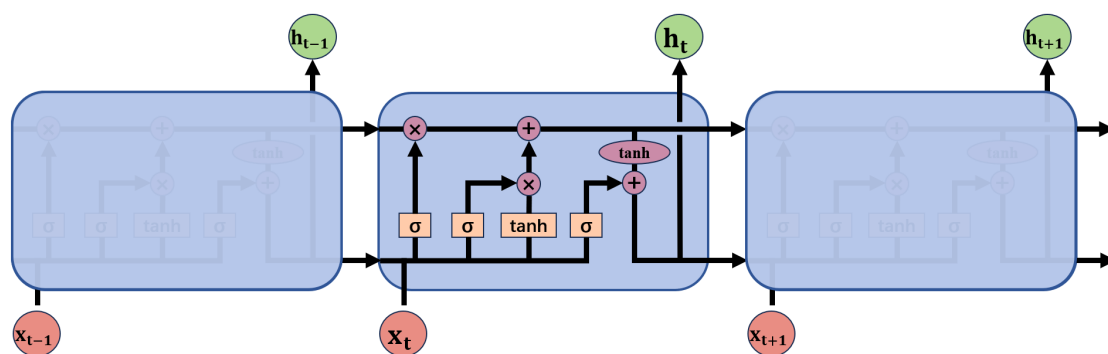


图 2.7 LSTM 结构图

### 2.7.2 基于 Transformer 的时序模型

Transformer 模型在时间序列预测领域的应用标志着建模范式的重要转变。该方法的核心创新在于摒弃了传统 RNN 的串行递归结构或 CNN 的局部归纳偏置, 转而采用基于自注意力机制 (Self-Attention) 的并行化架构。这一设计使其能够直接计算序列中任意两个时间点之间的关联权重, 从而实现对长期依赖关系的全局性建模, 有效克服了 RNN 类模型因梯度消失或爆炸而在长序列建模中面临的困境。其基本的多头自注意力机制可表述为:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (2.8)$$

其中, 对于时间序列输入  $X \in \mathbb{R}^{L \times d_{\text{model}}}$  ( $L$  为序列长度,  $d_{\text{model}}$  为模型维度), 通过线性变换得到查询 ( $Q$ )、键 ( $K$ )、值 ( $V$ ) 矩阵。将其扩展为多头注意力 (Multi-Head Attention) 可并行学习不同表示子空间的信息:

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h)W^O \quad (2.9)$$

$$\text{head}_i = \text{Attention}(XW_i^Q, XW_i^K, XW_i^V) \quad (2.10)$$

自注意力机制的置换不变性意味着其对输入序列的绝对或相对位置并不敏感，为弥补这一缺陷，模型必须显式地注入位置信息。常用方法包括：

(1) 正弦余弦位置编码 (Sinusoidal Positional Encoding) 使用预设的正弦和余弦函数生成不同频率的位置信号：

$$PE_{(pos,2i)} = \sin\left(\frac{pos}{10000^{2i/d_{model}}}\right), PE_{(pos,2i+1)} = \cos\left(\frac{pos}{10000^{2i/d_{model}}}\right) \quad (2.11)$$

其中  $pos$  为位置索引， $i$  为维度索引。该编码能自然地外推到比训练序列更长的位置。

(2) 可学习的位置嵌入 (Learnable Positional Embedding) 将每个位置索引映射为一个可训练的参数向量，此方法更具灵活性但外推能力可能受限。

在时间序列预测的具体实现中，通常采用编码器-解码器 (Encoder-Decoder) 架构。编码器负责对历史观测序列  $X_{1:T}$  进行深层表征学习，提取其蕴含的时序模式与动态特征。解码器则基于编码器的输出表征，结合已生成的部分未来序列，逐步预测未来多步值  $Y_{T+1:T+\tau}$ 。

为适配时间序列数据的特性，研究者还发展出多种 Transformer 变体。如 Informer<sup>[44]</sup>、Autoformer<sup>[67]</sup>、FEDformer<sup>[45]</sup> 等衍生模型针对时序特点优化了内存效率与长期预测能力。

## 2.8 强化学习

在强化学习的研究范式中，通常将待解决问题形式化为马尔可夫决策过程 (Markov Decision Process, MDP) 进行建模，它可以由形式为  $(S, A, P, R, \gamma)$  的元组表示，其中  $S$  是状态空间， $A$  是动作空间， $P$  是转移概率矩阵， $R$  和  $\gamma$  分别是奖励函数和折扣因子。策略定义为  $\pi : S \rightarrow A$  并训练以最大化 MDP 中预期的累积折扣奖励：

$$\max_{\pi} \mathbb{E} \left[ \sum_{t=0}^{\infty} \gamma^t R(s_t, \pi(a_t|s_t)) \right] \quad (2.12)$$

具体结构图如图 2.8 所示。

### 2.8.1 Q 学习和深度 Q 学习

Q 学习是主要的强化学习算法之一，它定义了一个 Q 值函数来表示预期的累积奖励：

$$Q^{\pi}(s, a) = \mathbb{E} \left[ \sum_{t=0}^{\infty} \gamma^t R(s_t, \pi(a_t|s_t)) \middle| s, a \right] \quad (2.13)$$

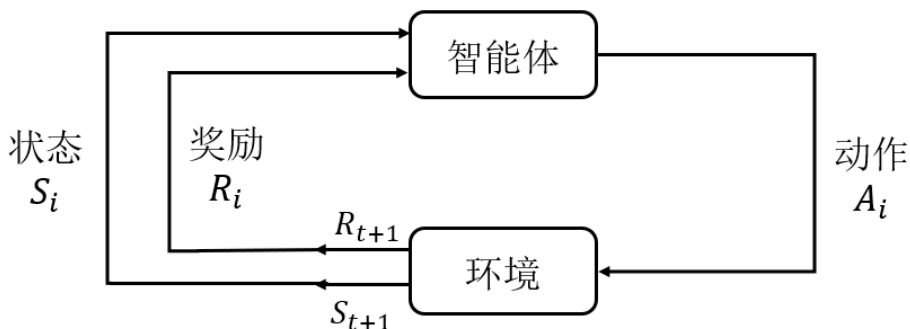


图 2.8 强化学习结构图

Q 学习通过最小化 Q 上的 Bellman 误差来训练 Q 值函数：

$$Q \leftarrow \arg \min_Q \mathbb{E} \left[ R(s, a) + \gamma \max_{a'} Q(s', a') - Q(s, a) \right] \quad (2.14)$$

深度 Q 网络 (Deep Q-Network, DQN) 是 Q 学习算法与深度神经网络相结合的重要成果。该方法的核心贡献在于引入函数逼近机制，有效解决了传统 Q 学习在处理高维状态空间时面临的维度灾难问题，即  $Q(s, a; \theta)$ ，其中  $\theta$  表示神经网络的参数。该算法通过最小化时序差分 (Temporal Difference, TD) 误差来优化网络参数，其目标函数定义为：

$$L(\theta) = \mathbb{E}(s, a, r, s') \sim \mathcal{D} \left[ \left( r + \gamma \max_{a'} Q(s', a'; \theta^-) - Q(s, a; \theta) \right)^2 \right] \quad (2.15)$$

DQN 采用了两个关键技术来提升训练稳定性：

- 经验回放 (Experience Replay)：智能体将其与环境交互产生的转移样本  $(s, a, r, s')$  存储于经验池  $\mathcal{D}$  中，训练时从中随机采样进行批量更新。这种机制打破了样本间的时间相关性，提高了数据效率并减少了训练方差。

- 目标网络 (Target Network)：使用一个独立的目标网络  $Q(s, a; \theta^-)$  来计算 TD 目标，其参数  $\theta^-$  定期从主网络复制而来。这种延迟更新策略有效缓解了因目标值频繁变化导致的学习不稳定问题。

## 2.8.2 离线强化学习算法

### 2.8.2.1 CQL 算法

离线强化学习算法具备从历史静态数据集中习得有效策略的能力，其显著特点在于学习过程中无需与环境进行实时交互，非常适用于临床环境。然而，在实践中，离线强化学习提出了一个主要的挑战，标准的离线强化学习方法可能会由于数据集和学习到的策略之间的分布转移导致 Q 值高估，在医疗保健环境中，这可能会转化为不安全的建议，使患者面临风险。保守 Q 学习 (Conservative Q-Learning, CQL) 通过构建一个保守的 Q 函数，旨在有效应对上述局限性问题，使策略在这个 Q 函数下界下的期望值为其真实值<sup>[46]</sup>。CQL 通过引入一个“保

守”项来惩罚那些与数据分布偏差较大的动作，从而避免这种高估现象，即它试图低估分布外 (Out-of-Distribution, OOD) 状态-动作对的 Q 值，以防止代理进入 OOD 状态。故 CQL 的优化目标变为：

$$Q \leftarrow \arg \min_Q \alpha \cdot \left[ \mathbb{E}_{s \sim \mathcal{D}, a \sim \pi(a|s)} Q(s, a) - \mathbb{E}_{s, a \sim \mathcal{D}} Q(s, a) \right] + \frac{1}{2} \left( R(s, a) + \gamma \max_{a'} Q(s', a') - Q(s, a) \right)^2 \quad (2.16)$$

其中  $\alpha$  是权重超参数，用于控制惩罚项的力度， $\pi$  是策略分布，表示从学习到的策略中采样的状态-动作对。CQL 通过最大化数据中的动作  $a \sim \mathcal{D}$  对应的 Q 值，并最小化学习策略  $\pi$  中高 Q 值的动作，从而抑制对那些在数据集中不常见动作的高估。

### 2.8.2.2 BCQ 算法

与 CQL 算法类似，Batch-Constrained Q-learning (BCQ) 算法也是为了应对离线强化学习中的分布偏移问题而提出的<sup>[48]</sup>。BCQ 主要通过限制策略对离线数据集中未见过的状态-动作对的依赖，以减少分布转移带来的风险。直接从静态数据集中学习策略时，BCQ 尝试通过约束策略的选择范围，使其更接近于已观察到的行为，从而避免因选择偏离数据分布的动作而做出不安全的决策。它通过引入生成模型，如变分自编码器 (Variational Auto-Encoders, VAE)，生成接近于数据分布的动作候选集合，再从中挑选动作以确保策略不会显著偏离历史数据：

$$p(a | s) \approx \text{VAE}(a|s) \quad (2.17)$$

$$a = \arg \max_{a' \in \mathcal{A}} Q(s, a') \cdot \mathbb{1}(a' \in \text{VAE}(a|s)) \quad (2.18)$$

这里的  $\mathbb{1}(a' \in \text{VAE}(a|s))$  是一个约束，确保只选择那些在生成模型中出现概率较高的动作，这样的设计减少了分布外动作带来的风险。

## 2.8.3 离线策略评估算法

### 2.8.3.1 FQE 评估算法

在线强化学习框架下，策略的性能评估通常依赖于与环境的实时交互过程。然而，当环境涉及真实患者的医疗场景时，以这种在线方式进行评估会带来巨大的风险。因此，本研究拟采用离线策略评估 (Off-Policy evaluation, OPE) 的方式对数据集进行策略评估。这些方法的性能最近在医疗环境中进行了评估，其中拟合 Q 评估 (Fitted Q Evaluation, FQE) 始终提供最准确的结果<sup>[68]</sup>。FQE<sup>[69]</sup> 将转换数据集  $D = \{s_t, a_t, s_{t+1}, r_t\}_{t=1}^n$  和策略  $\pi$  作为输入，并且在算法的每一步  $k$  处，使用  $D$  计算目标  $y_t = r_t + \gamma Q_{k-1}(s_{t+1}, \pi(s_{t+1}))$ ，进而求解

$Q_k = \operatorname{argmin}_{f \in F} \sum_{i=1}^n (f(s_i, a_i) - y_i)^2$ ，其中  $F$  是包含神经网络计算的所有函数的函数类。最终输出一个神经网络  $Q_\pi$ ，它估计策略  $\pi$  下任何状态-动作对  $(s, a)$  的值，然后通过取平均初始状态值来计算策略的性能，其中初始状态是前四小时的机械通气。由于医生策略有效地生成了数据集中的完整决策序列，因此可以通过取从该状态开始的完整决策序列的累积折扣奖励来计算每个初始状态的折扣回报。

根据当前策略和离线数据，FQE 利用拟合模型来更新  $Q$  值，以使其更好地逼近真实的  $Q$  函数。一旦  $Q$  函数收敛，便可通过  $Q$  值的期望来评估该策略的性能。即对于每个状态，取最优动作所对应的  $Q$  值，计算这些  $Q$  值的平均值，从而得到策略的整体表现。

### 2.8.3.2 DR 评估算法

尽管 FQE 能够通过拟合  $Q$  函数对目标策略进行稳定的离线评估，但由于其完全依赖函数逼近模型，因此当  $Q$  函数存在估计误差时，策略价值的评估结果可能产生一定偏差。为了进一步提高离线评估的可靠性，本研究在 FQE 的基础上引入 Doubly Robust (DR) 方法进行策略评估。DR 方法结合了基于模型的价值估计与基于重要性采样 (Importance Sampling, IS) 的无偏估计思想，在理论上能够同时降低估计偏差和方差，因此被广泛应用于离线强化学习中的策略评估问题<sup>[70-72]</sup>。

在离线数据集中，行为策略（即医生策略）记为  $\mu(a|s)$ ，目标策略记为  $\pi(a|s)$ 。DR 方法首先利用 FQE 得到目标策略  $\pi$  的动作价值函数估计  $Q_\pi(s, a)$ ，并进一步计算状态价值函数：

$$V_\pi(s) = \sum_a \pi(a|s) Q_\pi(s, a) \quad (2.19)$$

在此基础上，借助重要性权重机制，对行为策略与目标策略之间的分布偏差进行校正。对于一条轨迹  $\tau = s_t, a_t, r_{t=0}^T$ ，其逐步 DR 估计可表示为：

$$\hat{V}_{DR} = \sum_{t=0}^T \gamma^t [\rho_t (r_t - Q_\pi(s_t, a_t)) + V_\pi(s_t)] \quad (2.20)$$

其中  $\rho_t = \frac{\pi(a_t|s_t)}{\mu(a_t|s_t)}$  为重要性采样权重，用于评估在给定当前状态下，目标策略与行为策略所产生动作选择概率之间的偏离程度。该公式包含两个部分：第一项  $\rho_t (r_t - Q_\pi(s_t, a_t))$  用于利用真实奖励对  $Q$  函数估计误差进行修正；第二项  $V_\pi(s_t)$  则由 FQE 模型提供稳定的价值估计。当  $Q$  函数估计较准确时，DR 主要依赖模型估计；而当模型存在误差时，重要性采样项能够对估计进行补偿，从而实现双重稳健的性质。

在实际计算中，为了减少重要性采样权重带来的高方差问题，通常对权重进

行截断或归一化处理，以提高估计的稳定性。最终，通过对所有初始状态的 DR 估计值取平均，即可得到目标策略在离线数据集上的整体性能评估。

通过 FQE 与 DR 相结合，本研究能够在维持评估稳定性的基础上有效削减估计偏差，从而更加可靠地评估强化学习策略在机械通气参数推荐任务中的临床效果。

## 第3章 基于临床约束和集成学习的呼吸机启用决策方法

本章针对呼吸机启动时机的决策问题，提出了一种结合临床约束规则的集成模型，为患者使用呼吸机的时机提供了科学的决策依据。该模型将明确的插管指征编码为确定性规则对患者进行快速筛查，接着模型基于患者的生理状态分析出影响决策的重要特征。同时，模型设计了动态类别权重调整策略解决临床数据中严重的类别不平衡问题，有效提升了模型对少数类样本的识别能力。通过基学习器的叠加对大量数据进行学习，模型最终预测出患者需要通过呼吸机进行机械通气的概率，根据设定的阈值进而决定是否需要给患者安排呼吸机进行机械通气。

本章的组织结构安排如下：3.1节详细阐述数据集的预处理与特征工程的流程；3.2节阐述了模型的具体设计，包含分类模型，类别不平衡处理模型及临床约束模块；3.3节通过严谨的实验设计进行指标评估，全面验证所设计模型的有效性和优越性；3.4节对本章工作进行总结。

### 3.1 数据处理

本研究基于 2.3.1 节所述的 MIMIC-III 数据库进行数据处理<sup>[53]</sup>，通过标准化查询语言 (Structured Query Language, SQL) 将患者数据提取到 4 小时时间窗口的表中。对于每个患者，提取以下数据：人口统计学信息、生命体征、实验室指标、液体相关、是否使用呼吸机。处理后的部分数据示例如表 3.1 所示。选择了前 72 小时的机械通气数据，患者数据被分成状态、动作和奖励数组。对于数据插补，如果缺少不到 30% 的数据，则使用  $k=3$  的  $k$  最近邻 ( $k$ -Nearest Neighbors, KNN)<sup>[73]</sup> 插补，对于超过 70%，变量将从状态空间中被删除。

KNN 插补基于“相似患者具有相似临床指标”的假设，通过计算样本间的特征空间距离，寻找最相似的  $k$  个完整样本，并利用其对应特征值的加权平均估计缺失值。具体实施步骤如下：

1. 数据预处理与特征筛选。首先从原始数据中分离出需要进行 KNN 插补的特征子集。定义数据集  $\mathbf{X} \in \mathbb{R}^{n \times p}$ ，其中  $n$  为所有患者各时间点的样本数， $p$  为临床特征维度。筛选条件为：

$$\text{缺失率}(feature_j) = \frac{\text{缺失值数量}}{n} < 0.3 \quad (3.1)$$

同时，移除无需插补的元数据列，包括患者标识 (subject\_id、icustay\_id、hadm\_id)、时间戳 (start\_time、disctime) 以及人口学特征 (gender)，保留纯临床测量指标进行相似度计算。

2. 分块处理策略。为避免大规模医疗数据计算中的内存溢出问题,采用分块处理策略。将  $n$  个样本划分为  $m$  个连续数据块:

$$\mathcal{B}_i = \{X_{(i-1) \cdot S+1}, \dots, X_{\min(i \cdot S, n)}\}, i = 1, \dots, m \quad (3.2)$$

其中  $S = N \times 1000$  为预设块大小,  $N$  为可调参数。每个数据块独立进行 KNN 插补, 确保计算复杂度控制在可接受的范围内。

3. 距离计算与邻居选择。对于每个数据块  $\mathcal{B}_i$  中包含缺失值的样本  $\mathbf{x}_q$ , 基于双方均非缺失的特征子集计算其与块内所有完整样本  $\mathbf{x}_r$  的相似度距离。采用标准化欧氏距离度量:

$$d(\mathbf{x}_q, \mathbf{x}_r) = \sqrt{\sum_{j \in \mathcal{C}_{qr}} \left( \frac{x_{qj} - x_{rj}}{\sigma_j} \right)^2} \quad (3.3)$$

其中  $\mathcal{C}_{qr} = \{\text{特征 } j | x_{qj} \text{ 和 } x_{rj} \text{ 均非缺失值}\}$  为共享特征索引集,  $\sigma_j$  为第  $j$  个特征的标准差。随后, 按照距离升序排列, 选择前  $K$  个最近邻样本构成邻居集合  $\mathcal{N}_q = \mathbf{x}_{(1)}, \dots, \mathbf{x}_{(K)}$ ,  $K$  值通过交叉验证确定。

4. 缺失值估计。对于样本  $\mathbf{x}_q$  在第  $j$  个特征上的缺失值, 采用距离加权平均法进行估计。首先计算各邻居的权重:

$$w_k = \frac{1}{d(\mathbf{x}_q, \mathbf{x}_{(k)}) + \epsilon}, \quad k = 1, \dots, K \quad (3.4)$$

其中  $\epsilon = 10^{-6}$  为防止除零的极小常数。缺失值估计值为:

$$\hat{x}_{qj} = \frac{\sum_{k=1}^K w_k \cdot x_{(k)j}}{\sum_{k=1}^K w_k} \quad (3.5)$$

该加权策略赋予更相似样本更高的权重, 增强插补的局部适应性。对于分类特征, 则采用加权投票法确定最可能的类别。

5. 迭代优化。为提高插补稳定性, 实施迭代优化机制。以初始简单均值填补为基础, 执行多轮 KNN 插补:

①第  $t = 0$  轮: 用特征均值初始化所有缺失值 (基于非缺失值), 得到  $\mathbf{X}^{(0)}$ ;

②第  $t$  轮: 基于  $\mathbf{X}^{(t-1)}$  重新计算样本距离, 更新缺失值估计得到  $\mathbf{X}^{(t)}$ ;

③终止条件: 当连续两轮估计值变化小于阈值  $\tau$  或达到最大迭代次数  $T_{\max}$  时停止。

收敛准则为:

$$\frac{1}{|\mathcal{M}|} \sum_{(q,j) \in \mathcal{M}} |\hat{x}_{qj}^{(t)} - \hat{x}_{qj}^{(t-1)}| < \tau \quad (3.6)$$

其中  $\mathcal{M}$  为所有缺失位置集合。

6. 数据重构与整合。完成所有数据块的 KNN 插补后, 将填补后的临床特征与保留的元数据列重新整合:

$$\mathbf{X}_{\text{complete}} = [\mathbf{X}_{\text{meta}} \parallel \mathbf{X}_{\text{KNN-imputed}}] \quad (3.7)$$

其中  $\mathbf{X}_{\text{meta}}$  包含患者标识、时间戳等元数据， $\parallel$  表示列向拼接。

处理后得到患者样本数据量为 652382，每条数据包含 1 个标签值（是否上机）以及如下 18 个特征状态变量：

- 人口统计学信息：年龄、性别、体重、身高。
- 生命体征：序贯器官衰竭 (Sequential Organ Failure Assessment, SOFA) 评分、全身炎症反应综合征 (Systemic Inflammatory Response Syndrome, SIRS) 评分、格拉斯哥昏迷 (Glasgow Coma Scale, GCS) 评分、心率、呼吸频率、收缩压、舒张压、平均动脉压、休克指数、体温、血氧饱和度。
- 液体相关：尿量、静脉输液量、累计液体平衡量。

将根据上述方法得到的完整数据进行标准归一化，即使得数据具有零均值 ( $\text{mean} = 0$ ) 和单位标准差 ( $\text{standard deviation} = 1$ )，具体公式为：

$$z = \frac{X - \mu}{\sigma} \quad (3.8)$$

其中  $X$  是原始数据， $\mu$  是均值， $\sigma$  是标准差， $z$  是标准化后的数据。

表 3.1 部分初始数据示例表

icustay_id	subject_id	start_time	...	gender	...	spo2	sofa
200003	27513	8/3 5:40:00	...	M	...	95.75	0
200003	27513	8/3 9:40:00	...	M	...	95.5	7
...	...	...	...	...	...	...	...
200019	21789	7/8 21:35:00	...	F	...	100	7

### 3.2 模型设计

针对呼吸机启用时机的临床决策需求，本研究构建了一种基于患者生理状态的二分类预测模型，旨在通过量化评估患者的生理状态，为是否需要启用呼吸机提供数据驱动的决策支持。

该模型以患者实时采集的多维度生理状态特征作为输入，通过对这些特征进行标准化、缺失值插补与特征选择后，构建特征向量  $\mathbf{x} \in \mathbb{R}^d$ 。处理后的数据先基于临床约束模块进行筛选，过滤出规则无法确定的患者特征，接着输入至二分类机器学习模型中进行训练与决策。

模型输出为二元决策变量  $\hat{y} \in \{0, 1\}$ ，其中  $\hat{y} = 1$  表示模型判断患者当前需要启动呼吸机， $\hat{y} = 0$  则表示患者当前无需呼吸机支持。为提升模型的临床可解释性与决策可信度，本研究不仅输出分类结果，还同步输出属于正类（需启用呼吸机）的概率估计  $P(y = 1 | \mathbf{x})$ ，该概率值可视为患者呼吸衰竭风险的连续量化指标，有助于临床医生结合专业知识进行阈值调整与个体化决策。

在模型选择方面，本研究综合比较了多种分类算法，并通过相关指标评估各模型在平衡类别分布下的判别效能与稳健性。最终优选模型将集成特征重要性分析，识别影响机械通气决策的关键生理驱动因素，从而为临床提供为何在该时间点推荐启动呼吸机的合理解释。

该预测模型可嵌入实时监护系统，实现动态连续评估：每隔一定时间窗口对患者最新生理数据进行自动推理，输出当前时刻的上机推荐结果。本研究旨在通过该模型辅助临床医生早期识别潜在的呼吸功能恶化趋势，实现呼吸机介入时机的精准把握，最终达到临床患者状态的有效改善与医疗资源配置的科学优化。

基于患者生理状态的呼吸机启用决策模型总体结构图如图 3.1 所示。

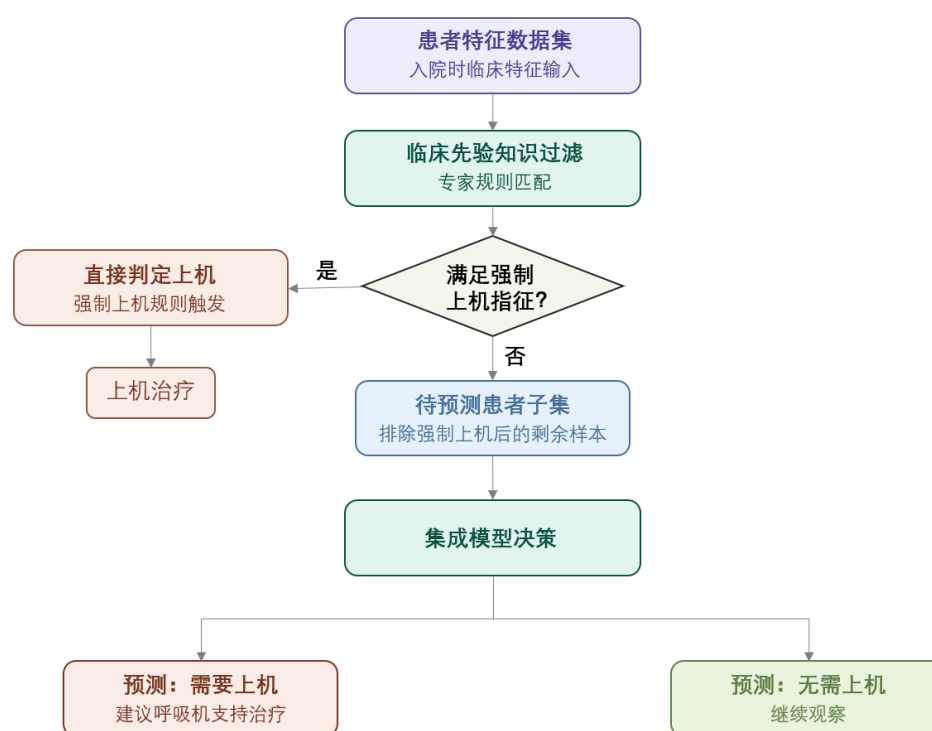


图 3.1 基于患者生理状态的呼吸机启用决策模型总体结构图

### 3.2.1 分类模块

目前分类模型主要包括集成学习模型、支持向量机及深度学习模型。支持向量机在小样本数据上表现稳定出色，而本研究所使用的数据量较大，故不予考虑使用。

集成学习模型处理结构化数据能力强，训练效率高，常表现最佳。Boosting 类集成模型作为一种集成学习算法，其核心理念在于将多个性能仅略优于随机猜测的弱学习器进行有效组合，最终构建出一个具备高精度的强学习器，具体流程如图 3.2 所示。其理论根基源于概率近似正确学习理论。该理论证明，只要存

在一个弱学习算法能够对问题的分布产生略微准确的假设，就可以通过 Boosting 技术将其提升至任意精度的强学习算法。Boosting 采用序列化、自适应的生成机制。其核心思想是：后续引入的基学习器专门致力于纠正其前驱模型所犯的预测错误。这种机制使得整个集成模型能够逐步聚焦于训练集中最难被正确学习的样本，从而实现模型性能的系统性提升。

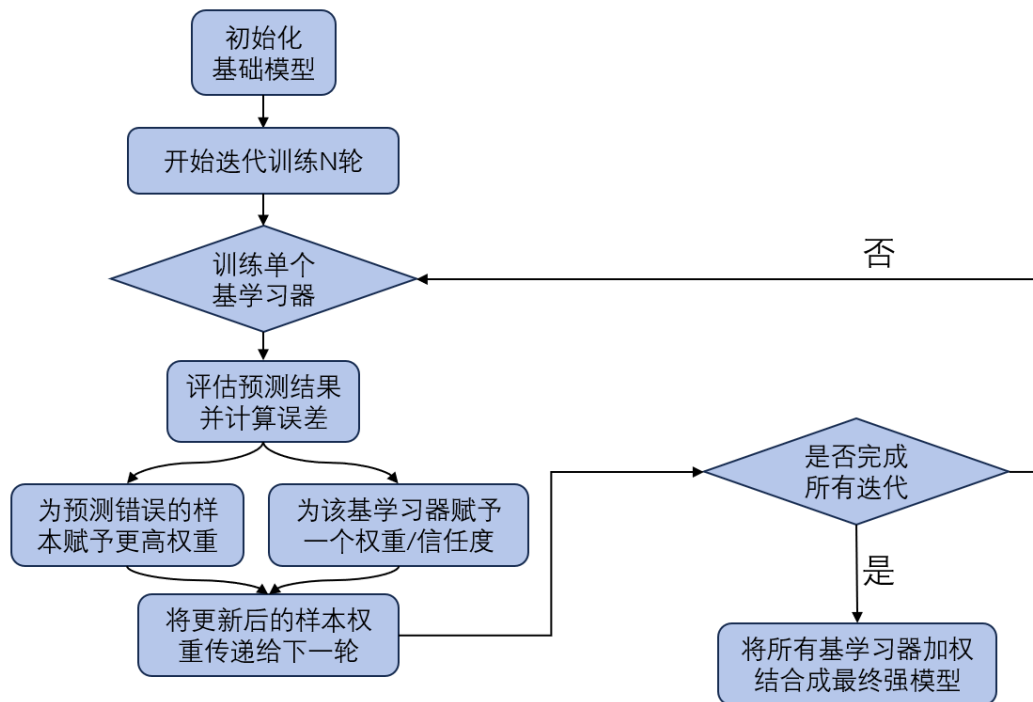


图 3.2 集成学习模型流程图

深度学习能捕捉数据中的复杂非线性关系。深度学习是一类受生物神经系统启发、基于深层人工神经网络结构的机器学习范式。其核心思想在于通过多层非线性处理单元的级联，对数据进行层层抽象，从而自动学习数据的分布式表示。与传统机器学习方法依赖人工特征工程不同，深度学习旨在端到端地从原始数据中直接学习具有判别力的高层特征。这一过程实现了表示学习，即模型自身在训练过程中，逐层地将原始输入转化为愈发抽象和任务相关的特征表示。本研究所采用的深度学习网络使用编码器式结构，由四个全连接隐藏层组成，每层均配备批量归一化、ReLU 激活函数和 Dropout 正则化。网络维度从初始的 256 个单元逐步递减至 32 个单元，最终通过 Sigmoid 函数输出 0 到 1 之间的概率值，具体的模型参数如表 3.2 所示。

### 3.2.2 类别不平衡处理模块

在临床预测任务中，如果正样本事件（如需要启用呼吸机）的发生率与负样本事件差距过大，会导致训练数据存在严重的类别不平衡，若直接使用原始数据

表 3.2 深度学习模型各层规模参数

层	类型	输入维度	输出维度	可训练参数量	设计目的
第 1 层	全连接 + BN + ReLU + Dropout	38	256	$38 \times 256 + 256 \times 2$	初始特征提取与降维
第 2 层	全连接 + BN + ReLU + Dropout	256	128	$256 \times 128 + 128 \times 2$	中级特征抽象
第 3 层	全连接 + BN + ReLU + Dropout	128	64	$128 \times 64 + 64 \times 2$	高级特征提取
第 4 层	全连接 + BN + ReLU + Dropout	64	32	$64 \times 32 + 32 \times 2$	最终特征压缩
输出层	全连接 + Sigmoid	32	1	$32 \times 1 + 1$	二分类概率输出

训练模型，分类器会因倾向于预测多数类而性能退化，表现为对关键少数类的召回率急剧下降。为消除此偏差，本研究设计并实现了动态类别权重调整模块，其核心在于权重修正，在训练阶段，通过特定机制引导其将更多注意力聚焦于少数类样本。

本模块采用基于样本数量的反向频率加权法。具体而言，对于包含  $N$  个样本的训练集，其中正样本数量为  $N^+$ ，负样本数量为  $N^-$ （满足  $N^+ + N^- = N$  且  $N^+ \gg N^-$ ），计算其类别权重向量  $w$  与集成模型专用权重  $\alpha$ 。计算过程如下：

1. 通用类别权重计算：采用 Scikit-learn 库的处理数据不平衡函数，并指定策略为 ‘balanced’。该策略会自动计算各类权重  $w_j$  为：

$$w_j = \frac{N}{C \times N_j} \quad (3.9)$$

其中， $C$  为类别总数（本研究中  $C = 2$ ）， $N_j$  为第  $j$  类样本数。此方法确保每个类别的总权重在训练中被平等对待。

2. 正样本缩放权重计算：为适配梯度提升树框架的相关参数，计算正样本的缩放权重  $\alpha$  为负、正样本数量之比：

$$\alpha = \frac{N^-}{N^+} \quad (3.10)$$

该参数直接影响模型的目标函数，使模型在每次梯度提升迭代中，将一个正样本分类错误的代价等同于将  $\alpha$  个负样本分类错误，从而显著提升模型对正类的敏感度。

在本研究中，正样本的数量为 441407，负样本的数量为 210975，正样本比例为 0.677。

通过此模块处理，模型优化目标从单纯追求整体准确率，转变为平衡地识别所有类别，尤其保障了对临床决策至关重要的机械通气需求检测情况，为构建高召回率的临床预警系统奠定了理论基础。

### 3.2.3 临床约束模块

单纯依赖机器学习的分类模块虽然能从历史数据中学习复杂模式，但存在临床可接受性不足的问题。若模型给出与公认急救指南相悖的建议，例如对已满足绝对插管指征的患者输出低风险概率，将严重削弱医生对系统的信任。

为解决上述问题，本研究设计一个临床先验规则约束模块，其核心思想是将明确、无争议的插管指征编码为确定性规则，在机器学习分类模块介入前对患者进行快速筛查。一旦患者满足任意一条绝对适应症规则，系统直接输出“需使用呼吸机”决策并终止模型预测；其余患者则进入后续机器学习流程。该模块相当于一个基于临床知识的安全防护网，确保模型永远不会违背基本急救准则。

本研究所用规则完全基于实时可获取的生命体征指标（心率 HR、呼吸频率 RR、收缩压 SBP、舒张压 DBP、平均动脉压 MAP、血氧饱和度 SpO<sub>2</sub>、体温 T、格拉斯哥昏迷评分 GCS）及其简单衍生量（休克指数 SI = HR/SBP、SOFA 评分值）。规则阈值及组合逻辑依据《急性呼吸窘迫综合征患者机械通气指南》制定<sup>[74]</sup>。

定义规则函数  $\mathcal{R}(\mathbf{x})$ ，输入为患者当前生命体征向量  $\mathbf{x}$ ，输出布尔值。若满足以下任一条件，则  $\mathcal{R}(\mathbf{x})=\text{true}$ ：

$$\left\{ \begin{array}{ll} \text{RR} \geq 35 \text{ 次/分} \quad \text{或} \quad \text{RR} \leq 8 \text{ 次/分} & (\text{呼吸频率严重异常}) \\ \text{SpO}_2 < 90\% \quad (\text{FiO}_2 \geq 0.5) & (\text{难治性低氧血症}) \\ \text{GCS} \leq 8 & (\text{气道保护能力丧失}) \\ \text{MAP} < 65 \text{ mmHg} \quad \text{且} \quad \text{休克指数} > 1.0 & (\text{失代偿性休克}) \\ \text{SOFA 评分} \geq 8 & (\text{严重多器官功能不全}) \end{array} \right.$$

需特别强调的是，本模块仅捕获具有明确插管指征的极端危重患者，其阈值设定极为保守。模块并不试图覆盖所有需插管患者，也不对模棱两可的临床情景进行干预，这正是后续分类模块的任务。通过将规则定位为高特异性筛查器，本研究避免了规则与模型之间的决策冲突，同时确保了系统的安全底线。

## 3.3 实验设计与结果分析

### 3.3.1 实验设置

本研究将基于集成学习模型和深度学习模型进行对比，集成学习模型包含以下：

**XGBoost:** Gradient Boosting 算法的扩展版本，以其高效的计算速度和良好的准确性而闻名。其采用的二阶梯度提升方法能够在提高预测准确度的同时控制模型的复杂度。此外，XGBoost 支持并行计算，这使得其在大数据集上具有显

著的优势。

**LightGBM<sup>[75]</sup>**：主要针对大规模数据集进行了优化。其采用的基于直方图的算法以及基于梯度的单边采样 (Gradient-based One-Side Sampling, GOSS) 技术, 使得 LightGBM 在保证模型精度的同时大幅提升了训练速度, 且内存占用更少。

**CatBoost<sup>[76]</sup>**：专注于解决类别型变量处理问题。其独特的对称树结构和高效的特征编码方法, 使得 CatBoost 在处理高维类别特征时具有显著的效果。同时, CatBoost 通过目标编码和随机排列等方式有效降低了过拟合风险。

深度学习模型的具体参数如表 3.2 所示。

### 3.3.2 评价指标

为全面评估模型在类别不平衡数据上的综合性能, 本研究采用接收者操作特征曲线与精确率-召回率曲线进行可视化分析与量化评估。受试者工作特征 (Receiver Operating Characteristic, ROC) 曲线主要用于评估分类模型在不同决策阈值条件下区分正负样本的综合能力, 能够从整体层面反映模型的判别性能。相比之下, 精确率-召回率 (Precision-Recall, PR) 曲线更加关注少数类别的识别效果, 尤其适用于类别分布不平衡的场景, 用于衡量模型在正类预测方面的准确程度与覆盖情况, 两者结合可提供对模型性能更深刻、更稳健的评估。

#### 3.3.2.1 ROC 曲线

ROC 曲线描绘了当模型的分类阈值从 0 到 1 变化时, 其真正例率 (True Positive Rate, TPR) 与假正例率 (False Positive Rate, FPR) 之间的权衡关系。TPR 与 FPR 的定义如下:

$$TPR = \frac{TP}{TP + FN}, \quad FPR = \frac{FP}{FP + TN} \quad (3.11)$$

TP、FP、FN 与 TN 分别对应真阳性、假阳性、假阴性及真阴性四种情形的样本统计量。基于这些基础指标绘制的曲线所围成的面积是 AUC-ROC 值, 它能够以单一标量形式反映模型整体排序性能。该度量取值 0.5 表明模型性能等同于随机猜测, 而趋近 1.0 则代表其具备了近乎理想的判别效果。本研究将对比不同模型的 ROC 曲线, 并报告其 AUC 值, 以评估模型在整体上区分“需要呼吸机”患者与“无需呼吸机”患者的能力。

#### 3.3.2.2 PR 曲线

当面对临床数据集中正负样本分布极不均衡的情形时, PR 曲线提供了比 ROC 曲线更具信息量的性能视图。PR 曲线展示了模型在不同阈值下的精确率 (Precision, 或称阳性预测值) 与召回率 (Recall, 即 TPR) 之间的权衡。其定义

如下：

$$Precision = \frac{TP}{TP + FP}, \quad Recall = \frac{TP}{TP + FN} \quad (3.12)$$

该曲线下面积是评估模型在不平衡数据集上性能的关键指标。在极度不平衡的情况下，即使一个模型的 AUC-ROC 值较高，其 AUC-PR 也可能很低，因为模型可能在保证高召回率的同时产生了大量假阳性，导致精确率急剧下降。高 AUC-PR 值意味着模型能够在高召回率下依然保持高精度，这对于临床决策至关重要，因为过高的假阳性率会引发临床警报疲劳。

### 3.3.3 实验结果

XGBoost、LightGBM、CatBoost 及深度学习模型的实验结果对比图如图 3.3 及图 3.4 所示，相关的具体数值如表 3.3 所示。可得 XGBoost 模型在呼吸机启用的决策问题中表现最佳。

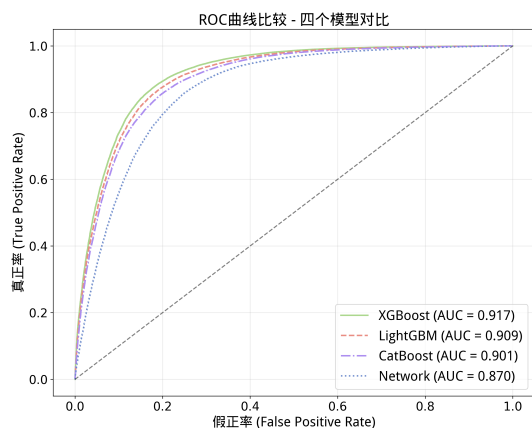


图 3.3 模型的 ROC 曲线对比图

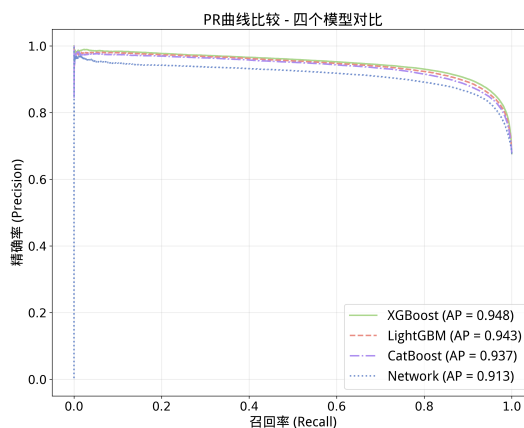


图 3.4 模型的 PR 曲线对比图

表 3.3 对比模型的 AUC-ROC 及 AUC-PR 值

模型	AUC-ROC	AUC-PR
XGBoost	0.9171	0.9483
CatBoost	0.9006	0.9372
LightGBM	0.9089	0.9428
神经网络	0.8702	0.9130

通过分析得到数据集的患者特征中，与 XGBoost 模型拟合高度相关的前十特征的重要性排序如图 3.5 所示，重要性从大到小分别是 GCS 评分、SOFA 评分、呼吸频率、血氧饱和度、体温、收缩压、体重、年龄、性别和累计液体平衡量，符合医学常理。

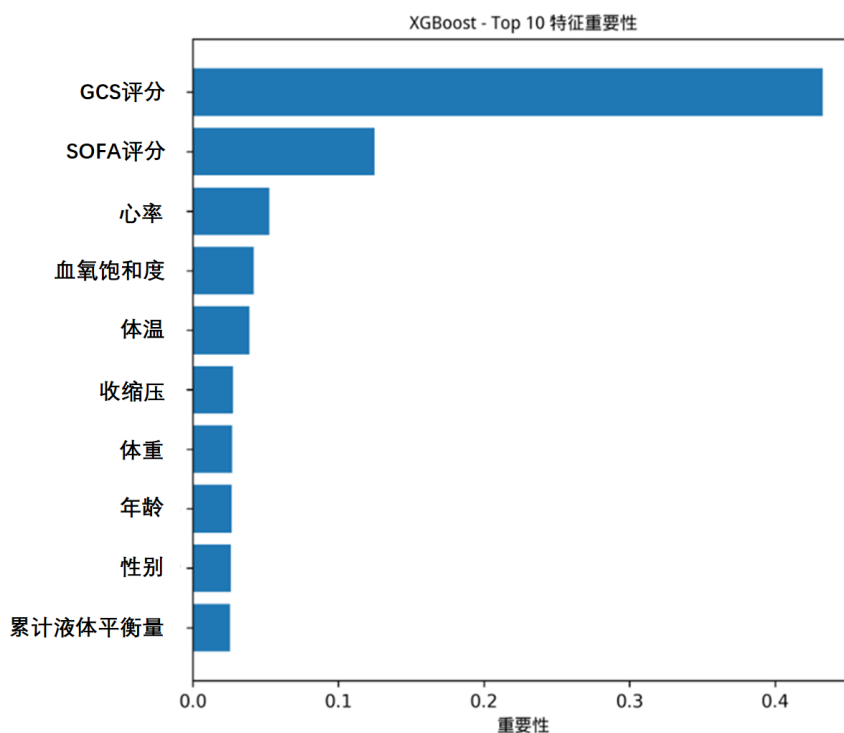


图 3.5 特征重要性排序

### 3.4 本章小结

本章聚焦于基于患者生理状态的呼吸机启用决策问题，开展了从数据清洗、模型构建到实验验证的系统性研究。

首先，详细阐述了基于 MIMIC-III 数据库的数据预处理流程，为模型训练提供了坚实的数据基础。其次，设计了一种临床先验规则约束模块，在机器学习分类决策模块介入前对患者进行快速筛查。接着，深入探讨了集成学习与深度学习两种主流建模方法。针对临床数据中严重的类别不平衡问题，设计了动态类别权重调整策略，显著增强了模型捕获少数类样本的能力。同时，确立了以 AUC-ROC 和 AUC-PR 为核心的多维度评估体系，确保了模型评估的全面性与客观性。最后，通过对比实验分析，验证了不同模型在呼吸机启用决策任务上的表现。在数据集 MIMIC-III 上的实验结果表明，XGBoost 模型在综合性能上优于其他集成模型及深度学习模型，ROC 值和 PR 值分别为 0.9171 和 0.9483，展现出最高的预测精度。本章的研究工作不仅筛选出了影响临床决策的关键特征，也成功构建了可靠的决策模型。

## 第 4 章 基于特征回归筛选和时序深度建模的呼吸机使用时长预测方法

本章针对患者的呼吸机使用时长预测的问题，提出了一种结合静态特征和动态特征的预测模型。该模型结合了特征回归筛选及时序深度建模的分步策略，构建了级联预测机制，通过回归模块识别影响机械通气时长的重要特征，通过改进的时序预测模块的多层次特征提取、残差学习和注意力机制实现对患者病情演变的全面建模，捕捉其动态规律，为呼吸机的使用时长提供了更精准的预测。

本章的组织结构安排如下：4.1节阐述数据集的预处理与特征工程的流程；4.2节阐述了模型的具体设计，包括特征回归模块及时序预测模块；4.3节通过严谨的对比实验、消融实验进行指标评估，全面验证预测模型的有效性和优越性；4.4节对本章工作进行总结。

### 4.1 数据处理

本章所用的数据处理方式如 3.1节所述，另外将患者的呼吸机使用时长作为回归模型的标签。

### 4.2 模型设计

针对患者呼吸机使用时长的预测任务，本研究设计了一种两阶段级联预测框架，该框架融合了静态生理特征分析与动态时序演化建模，旨在通过分阶段精细化预测提升最终输出的准确性与临床可解释性。

在第一阶段，本研究基于患者入院初期及监护过程中的静态与准静态生理指标（如年龄、合并症评分、初始生命体征、实验室检测结果等），构建回归模型预测初步的呼吸机使用时长。该阶段采用回归算法进行训练与比较，其目标是从患者的基础生理状态中提取与机械通气需求相关的全局模式，生成一个基准预测值  $\hat{y}^{(1)}$  作为后续时序预测的重要输入特征之一。

第二阶段，为捕捉患者病情在 ICU 监护期间的动态演变趋势，本研究将第一阶段输出的初步预测时长与患者按时间序列采集的纵向生理数据进行融合，构建面向时序预测的特征序列。具体而言，采用滑动时间窗策略，设定窗口长  $T$ ，对于每位患者构建如下形式的样本序列：

$$\mathcal{D}_{\text{seq}} = \left\{ \left( \mathbf{X}_{t:t+T-1}, \hat{\mathbf{y}}_{t:t+T-1}^{(1)} \right), y_{t+T-1} \right\}_{t=1}^{N_{\text{patients}}} \quad (4.1)$$

其中：

•  $\mathbf{X}_{t:t+T-1} \in \mathbb{R}^{T \times d}$  为连续  $T$  个时刻的原始生理特征矩阵（包括生命体征、实验室指标等）；

•  $\hat{\mathbf{y}}_{t:t+T-1}^{(1)} \in \mathbb{R}^T$  为对应时刻第一阶段模型输出的初步通气时长预测序列；

•  $y_{t+T-1} \in \mathbb{R}$  为窗口末端时刻（第  $t+T-1$  时刻）的真实机械通气时长，作为监督学习的目标变量。

在此基础上，本研究引入时序深度学习模型对该序列进行建模，其目标不再是简单拟合历史曲线，而是基于患者当前及历史状态对其当前时刻所需通气时长进行动态估计。

---

#### 算法 4.1 基于回归与时序建模的两阶段呼吸机使用时长预测算法

---

**Input:** 患者数据集  $\mathcal{D} = \{(X_i, y_i)\}_{i=1}^N$ ，时间窗口长度  $T$

**Output:** 训练完成的时序预测模型  $f_{seq}(\cdot)$  及最终机械通气时长预测  $\hat{y}$

```

1 阶段一：回归模型初步预测
2 提取患者静态特征及机械通气时长数据  $(X_i, y_i)$ ；
3 使用数据拟合回归模型  $f_{reg}(\cdot)$ ；
4 for 每个患者  $i = 1, 2, \dots, N$  do
5   | 计算初步预测通气时长： $\hat{y}_i^{(1)} = f_{reg}(X_i)$ ；
6 end
7 阶段二：时序特征构建与深度时序预测
8 for 每个患者  $i$  do
9   | 获取其 ICU 监测期间的纵向生理特征序列  $X_{t:t+T-1} \in \mathbb{R}^{T \times d}$ ；
10  | 将阶段一预测值扩展为时间序列： $\hat{\mathbf{y}}_{t:t+T-1}^{(1)}$ ；
11  | 构建融合特征序列： $\tilde{X}_{t:t+T-1} = (X_{t:t+T-1}, \hat{\mathbf{y}}_{t:t+T-1}^{(1)})$ ；
12 end
13 构建训练样本： $\mathcal{D}_{seq} = \{(\tilde{X}_{t:t+T-1}, y_{t+T-1})\}_{t=1}^{N_{patients}}$ ；
14 使用时序预测模型  $f_{seq}(\cdot)$  进行训练；
15 初始化时序模型参数  $\theta$ ；
16 for epoch = 1 to  $N_{epoch}$  do
17   | 从  $\mathcal{D}_{seq}$  采样小批量数据  $(\tilde{X}_{t:t+T-1}, y_{t+T-1})$ ；
18   | 进行前向传播  $\hat{y}_{t+T-1} = f_{seq}(\tilde{X}_{t:t+T-1}; \theta)$ ；
19   | 计算损失函数  $\mathcal{L}(\theta) = \frac{1}{B} \sum_{i=1}^B (\hat{y}_i - y_i)^2$ ；
20   | 计算梯度  $\nabla_{\theta} \mathcal{L}(\theta)$ ；
21   | 更新模型参数  $\theta \leftarrow \theta - \eta \nabla_{\theta} \mathcal{L}(\theta)$ ；
22 end
23 return  $f_{seq}(\cdot)$ ,  $\hat{y}$ ;
```

---

该级联框架的优势在于：第一阶段的回归模型能够从全局角度识别影响呼吸机使用时长的基础生理因素，提供稳定且可解释的基准预测；第二阶段的时序模型则专注于学习状态演变过程中的局部波动与临界变化，从而实现对患者实时治疗需求的动态响应。两阶段之间通过特征传递形成互补，既避免了单一模型可能存在的偏差-方差困境，也增强了预测结果在临床决策中的时效性与适应性。最终输出为时序模型在各时刻的预测值，可供临床用于实时评估患者的脱机准备状态与治疗调整时机。

基于特征回归筛选和时序深度建模的呼吸机使用时长两阶段级联预测模型结构如图 4.1 所示，具体流程见算法 4.1。

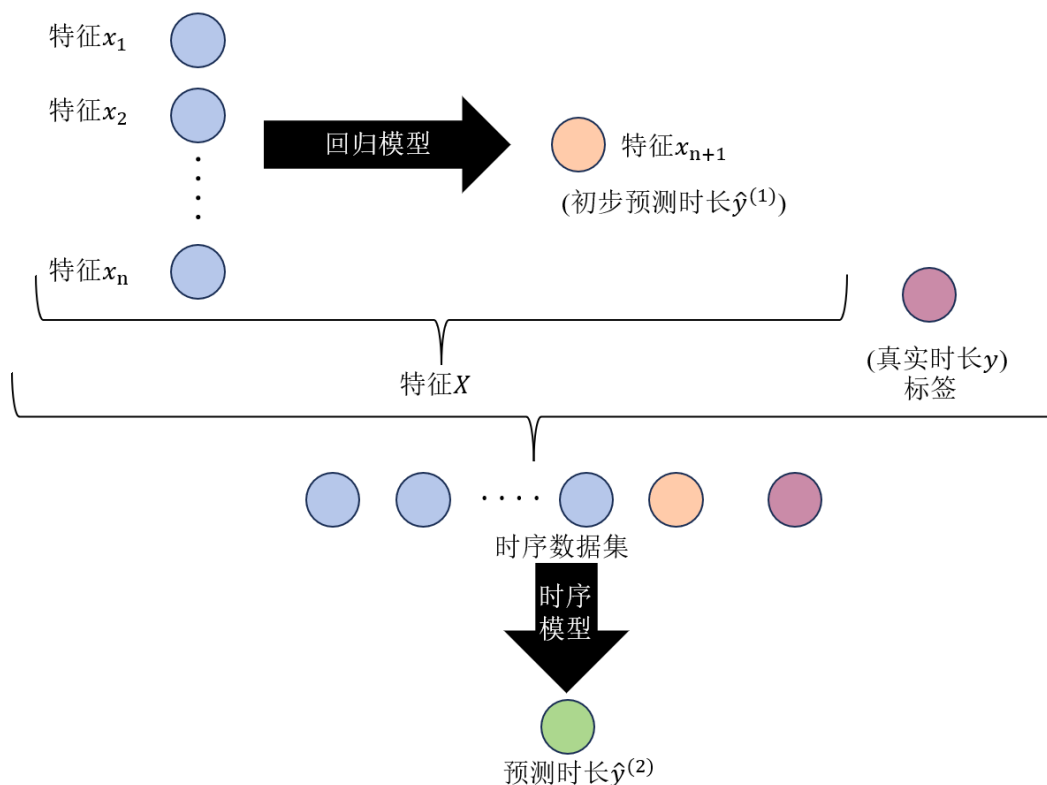


图 4.1 呼吸机的使用时长预测模型结构图

#### 4.2.1 特征回归模块

回归模型是一种用于描述变量间统计关联的经典数学模型。它基于概率统计理论，系统研究因变量（目标变量）与一个或多个自变量（预测变量）之间的量化关系。在机器学习与数据分析领域，回归模型被广泛应用于预测建模、时间序列分析、因果推断以及变量相关性探索等场景。

其核心思想是通过优化算法对观测数据进行拟合，寻找能最准确刻画变量间关系的函数形式，进而实现对未知数据或未来趋势的可靠预测。回归模型不仅提供预测功能，还能通过系数估计解释变量间的影响方向与强度，为决策提供统计依据。

从方法学角度看，回归分析涵盖多种技术路径：线性回归假设变量间呈线性关系，通过最小二乘法求解；广义线性模型扩展至指数族分布，可处理分类与计数数据；正则化回归通过引入惩罚项应对高维与共线性问题；而非参数回归则放松函数形式假设，适应复杂非线性模式。

在本研究中，回归模型可以深入挖掘患者的生理状态特征与预计机械通气时长的关系。本研究通过实验选择合适的回归模型，比较对象包括基本回归模型：线性 (Linear)、决策树 (Decision Tree)、支持向量机 (Support Vector Machine, SVM)、K 近邻 (k-Nearest Neighbors, KNN) 回归、最小绝对收缩和选择算子 (Least Absolute Shrinkage and Selection Operator, LASSO) 回归、弹性网络 (Elastic Net,

ENet) 回归, 以及集成方法模型: 随机森林 (Random Forest, RF)、自适应提升 (Adaptive Boosting, Adaboost)、梯度提升 (Gradient Boosting)、装袋算法 (Bootstrap aggregating, Bagging)、极端随机树 (Extra Tree)。

#### 4.2.2 时序预测模块

在重症监护场景中, 患者的生理状态呈现连续动态演变特征。仅依赖入院初期的静态快照数据 (如初始生命体征、实验室指标等), 难以有效捕捉病情进展中的关键趋势变化, 包括潜在的恶化拐点或治疗响应轨迹。这种动态性的缺失将直接制约风险预警的及时性与治疗决策的前瞻性。

为突破上述局限, 本研究设计并集成了时序预测模块, 其核心在于引入深度学习时序建模框架。我们通过引入先进的序列建模架构, 进而充分利用患者在 ICU 监护期间产生的纵向监测数据。这些时序数据构成患者状态演变的连续轨迹, 为模型学习病情动态规律提供信息基础。

传统时序模型通常旨在从宏观数据中挖掘周期、趋势等统计规律, 并基于已知影响因素预测未来观测值, 通常是多特征到与之相同的多特征的预测, 而本研究与时序预测的经典应用场景不同, 本研究的时序模块聚焦于患者状态感知与当前时刻的机械通气时长预测。具体而言, 模型以滑动时间窗方式输入患者当前时刻及历史多时刻的生理状态序列, 通过时序特征提取与状态编码, 最终输出对当前时刻所需机械通气时长的动态预测。这一设计使得预测能够实时响应患者状态变化, 当出现呼吸参数恶化或感染指标升高时, 模型可即时调整对通气时长需求的判断。

因此, 本研究提出了一种基于多尺度特征融合与注意力机制的残差 LSTM 模块 (命名为 MRA-LSTM)。该模块旨在克服传统 LSTM 在医疗时序数据分析中的局限性, 通过多层次特征提取、残差学习和注意力机制, 实现对患者病情演变的全面建模。模块核心组件设计如图4.2所示, 具体描述如下:

##### (1) 多尺度时序特征提取模块

考虑到医疗监测数据中既包含短期生理波动又蕴含长期病理趋势的特点, 本模块首先采用多尺度一维卷积神经网络并行提取不同时间尺度的局部时序模式:

- 短期特征提取层: 使用卷积核大小为 2 的一维卷积层, 捕捉相邻时间点间的瞬时变化模式 (如血压骤降、心率突增等短期异常信号)。
- 中期特征提取层: 使用卷积核大小为 3 的一维卷积层, 识别三个连续时间点构成的趋势性变化 (如呼吸频率的逐渐加快或减慢趋势)。

两个卷积层均采用 ReLU 激活函数和适当填充策略, 确保输出序列长度与输入保持一致。通过特征拼接将短期和中期特征融合, 形成一个包含多维多尺度特征的时序表征。

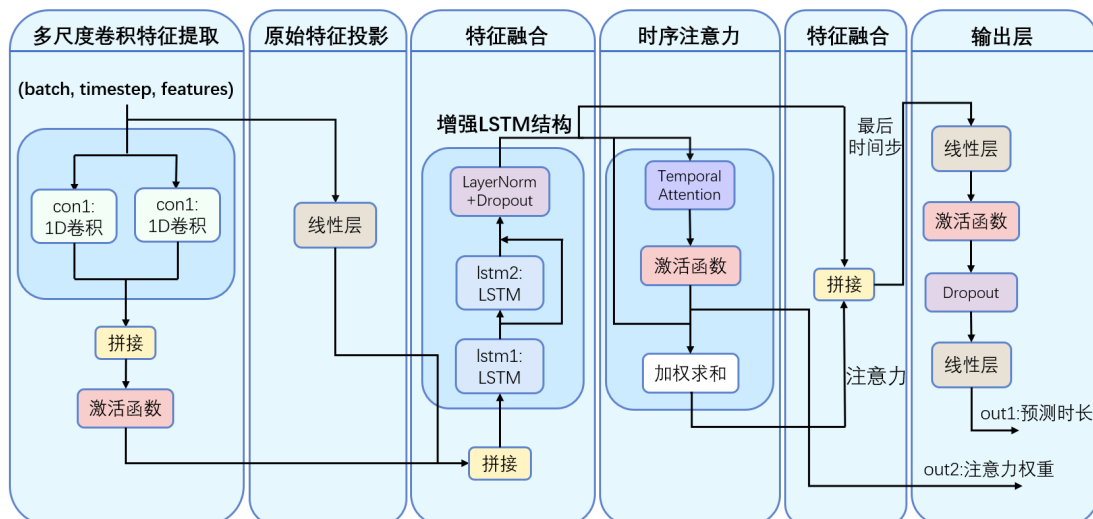


图 4.2 MRA-LSTM 模块结构图

### (2) 全局特征投影模块

为保留原始特征的完整信息并降低维度，同时引入非线性变换能力，本模块采用全连接层将 36 维原始输入特征投影至 64 维嵌入空间。该投影过程实现了通过非线性激活函数增强了特征的表达能力。

### (3) 特征融合与增强

将多尺度 CNN 提取的局部特征与线性投影后的全局特征在特征维度上进行拼接，形成 96 维的增强时序特征。这种设计实现了局部细节与全局上下文的有机结合，为后续的时序依赖建模提供了丰富的信息基础。

### (4) 残差 LSTM 时序建模模块

传统深度 LSTM 网络容易遭受梯度消失问题，影响深层特征的学习效果。为解决这一问题，本模块采用残差学习策略构建双层 LSTM 网络：

- 第一层 LSTM：接收融合后的增强特征，学习基础的时序依赖模式，输出初步的隐藏状态序列。
- 第二层 LSTM：以第一层输出为输入，学习更复杂的时序动态和长期依赖关系。
- 残差连接：将两层 LSTM 输出通过跳跃连接相加，再经过层归一化处理。这种设计确保即使深层网络未学习到有效信息，基础特征也能直接传递至后续层，能够对梯度消失现象起到显著抑制作用，同时推动模型收敛速度的提升。

### (5) 时序注意力机制

医疗时序数据中不同时间点对最终预测的贡献度存在差异，为动态识别关键临床时刻，本模块引入时序注意力机制：

- 注意力评分网络：采用两层全连接网络计算每个时间步的重要性分数，其中间层使用 Tanh 激活函数引入非线性。

- 注意力加权：借助 Softmax 函数对重要性分数进行归一化处理，从而生成总和为 1 的注意力权重分布。

- 上下文向量生成：基于注意力权重对 LSTM 隐藏状态进行加权聚合，最终生成的上下文向量能够有效凝练时序数据中的关键信息。

注意力机制使模型能够自动聚焦于对呼吸机上机时长预测最相关的关键时间点，如用药后的反应期、病情转折点等临床重要时刻。

#### (6) 双路径特征融合策略

为同时考虑整体时序趋势和最新生理状态，本模块采用双路径特征融合策略：

- 注意力路径：基于注意力加权的上下文向量，捕捉整个观测期间的整体趋势和关键变化。

- 最新状态路径：直接采用 LSTM 最后一个时间步的隐藏状态，反映患者当前的实时生理状况。

将两条路径的特征拼接后，通过多层感知机进行最终预测，这种融合方式模拟了临床医生既考虑病情演变趋势又重视当前状态的综合决策过程。

## 4.3 实验设计与结果分析

### 4.3.1 实验设置

#### 4.3.1.1 特征回归模块

利用多种回归方法，遍历拟合一遍数据，并对拟合优秀的回归方法进行交叉验证以防止过拟合，选择最优的回归方法进行后续的网络检索调参。

按照 8:1:1 的比例将数据集随机划分为训练集 (80%)，验证集 (10%) 和测试集 (10%)。为了保证实验的可复现性，随机种子固定为 0。

为了系统评估各类回归算法对研究数据的适应性与回归性能，本章构建了一个包含 12 种不同建模机制的回归模型库。该库覆盖了集成学习、线性与正则化模型以及其他非线性模型三大类别，旨在通过多角度、多策略的对比，揭示不同算法在呼吸机使用时长回归任务中的表现差异。

①集成学习模型部分：包含基于 Bagging 与 Boosting 两类策略的共 6 种算法。Bagging 类模型包括 RF、Bagging 与 Extra Tree，其通过 Bootstrap 采样构建多个基学习器并采用平均或投票机制降低方差、提升稳定性。Boosting 类模型则涵盖 XGBoost、Gradient Boosting 与 AdaBoost，这类模型以前向分步的方式逐棵构建决策树，后续生成的决策树主要承担对前序模型遗留残差进行修正的任务，从而逐步降低偏差、提升预测精度。集成模型通常能够有效处理高维特征间的复杂交互与非线性关系，且对数据中的噪声具有一定鲁棒性，因此在结构化数据的

预测任务中表现突出。

②线性与正则化模型：包括普通线性回归、Lasso 回归和 ENet 回归共 3 种算法。值得注意的是，对于 Lasso，本研究构建了包含 RobustScaler 的预处理管道，该缩放器使用中位数与四分位数范围进行标准化，从而降低极端值对线性模型参数估计的干扰。为考察变量间是否存在显著线性关联以及正则化对模型复杂度的控制效果，以线性回归为基准模型，用于揭示特征与目标之间是否存在全局线性趋势。Lasso 回归则在损失函数中引入  $L_1$  正则化项，通过对系数施加稀疏约束实现自动特征选择，有助于提升模型的解释性并缓解过拟合。ENet 回归融合了  $L_1$  和  $L_2$  正则化项，能在特征选择与组效应建模之间取得平衡，尤其适用于特征存在多重共线性的场景。

③其他非线性模型：为捕捉数据中可能存在的局部模式与非光滑响应关系，本部分引入 SVM、KNN 与 Decision Tree 共 3 种回归算法。SVM 以结构风险最小化为理论基础，借助核函数将原始数据映射至高维特征空间，并在该空间中求解能够实现最大分类间隔的最优超平面，适用于小规模数据中的非线性拟合。KNN 回归属于惰性学习算法，其预测结果为邻近样本目标值的加权平均，能够自然适应局部波动与复杂边界。决策树则以贪心策略递归划分特征空间，生成具有明确解释性的树状结构，为后续集成模型提供基学习器基础。

通过上述三类共 12 种回归模型的对比实验，本研究不仅可横向比较不同算法在相同评估准则下的性能优劣，还可从偏差 - 方差权衡、特征选择能力等多个维度，为临床预测模型的选择与优化提供实证依据。整体实验设计兼顾了方法的代表性、先进性与可解释性，旨在构建一个全面的算法评估体系。

#### 4.3.1.2 时序预测模块

确定回归模型后，将患者数据集的所有状态特征载入模型，得到预测的机械通气时长作为新的特征。构造的时序数据集具体如下：步长  $T$  为 5，特征为 3.1 节提取的患者特征以及预测机械通气时长，标签为最新状态下的真实机械通气时长。

深度学习在时间序列预测领域经历了显著的演进。早期的 LSTM 通过门控机制有效解决了长程依赖问题，成为该领域的基石<sup>[65]</sup>。随后，LSTNet 通过融合卷积神经网络 (Convolutional Neural Networks, CNN) 与跳跃连接，进一步增强了对局部特征和周期性模式的捕捉能力<sup>[77]</sup>。随着 Transformer 架构的兴起<sup>[78]</sup>，许多研究工作致力于对该框架进行改进。

为了全面验证本研究提出的预测模型在呼吸机时长预测任务上的有效性与先进性，实验环节选取了时序预测领域的多项主流算法作为基线参照，进行性能比对分析。这些基线模型涵盖了经典的深度神经网络、结合注意力机制的循环网

络、当前最先进的基于 Transformer 的架构以及一些著名且先进的算法：

为了评估 LSTM 模型的基础预测能力，本研究构建了一个标准的 LSTM 基线模型。该模型采用多层 LSTM 结构，通过引入门控单元来捕捉时序数据的长距离依赖，此种机制使模型具备了持续记忆及刷新历史状态的功能，有效处理临床监测数据中的时序特征。该基线用于验证传统 LSTM 在处理时长预测任务时的基本性能。

为了评估自注意力机制对预测性能的贡献，本研究构建了一个增强型的 LSTM 模型。该模型在传统 LSTM 的隐层输出之上引入了类似于 Transformer 模型所采用的自注意力机制，值得注意的是，该自注意力机制比本文所提模块的时序注意力机制复杂的多，区别如表 4.1 所示。自注意力机制能够动态地为序列中不同时间步的表示赋予相应权重。自动聚焦于对预测结果影响最大的关键时刻，从而缓解长序列中的信息遗忘问题。该基线用于验证单纯引入自注意力机制是否优于本研究提出的方法。

表 4.1 所提模块的时序注意力与 Transformer 自注意力对比

对比维度	时序注意力	Transformer 自注意力
核心机制	独立打分加权	两两关系建模
建模方式	重要性加权	依赖关系建模
参数规模	小	大
计算复杂度	$O(T)$	$O(T^2)$
数据需求	低（小样本友好）	高（依赖大数据）
过拟合风险	低	高
训练稳定性	高	较低
可解释性	强（权重直观）	一般（多头复杂）
抗噪声能力	强（平滑加权）	较弱（易放大噪声）
位置编码需求	不需要	需要
适用场景	医疗时序/分类任务	NLP/长序列建模

LSTNet 是一种经典的深度学习基线，它创新性地结合了 CNN 和 RNN。该模型借助 CNN 从多变量时序数据中挖掘变量间潜在的局部关联与短期依赖模式，并使用 RNN 捕捉长期演变趋势。此外，LSTNet 引入了跳跃连接以专门处理具有周期性模式的时间序列。选取该模型旨在对比 CNN 与 RNN 混合架构在医疗时序数据上的表现。

Informer<sup>[44]</sup>作为一种面向长序列时间序列预测任务优化的 Transformer 架构变体，针对标准模型中存在的计算复杂度过高问题，通过引入概率稀疏自注意力

机制实现了时间复杂度与内存占用的有效降低，并采用生成式解码器一步生成预测结果。选取 Informer 旨在评估高效 Transformer 架构在处理中长程医疗监测数据时的适用性。

FEDformer<sup>[45]</sup> 是目前时间序列预测领域的最先进模型之一。不同于 Informer 仅在时域进行建模，FEDformer 结合了季节性-趋势分解与频域分析。它通过傅里叶变换或小波变换在频域中通过随机采样低频模式来捕捉序列的全局特征。将该模型纳入对比，旨在以最严苛的标准验证本研究模型在捕捉全局趋势与局部波动方面的竞争力。

TimesNet<sup>[79]</sup> 是针对时间序列预测任务优化的新型深度学习架构，该模型的主要创新思路体现为将原始的一维时间序列重构为二维张量形式进行特征学习与建模分析。TimesNet 通过对原始时间序列进行一维离散傅里叶变换，识别出序列中最显著的周期性，并将时间点根据其相位重新排列为二维矩阵。这种转换使得模型能够利用成熟的二维视觉骨干网络同时捕捉时间序列的周期内和周期间的变化模式。纳入 TimesNet 进行对比研究，旨在验证将时间序列的周期先验知识显式建模的有效性，特别是对于具有明显周期性特征的医疗监测数据。

TSMixer<sup>[80]</sup> 是基于全连接层架构的时间序列预测模型，代表了当前轻量级时间序列建模的最新进展。该模型的核心思想是使用多层感知机交替处理时间维度和特征维度的信息，通过时间混合和特征混合两个并行的 MLP 模块捕获序列的时间相关性和特征相关性。TSMixer 摒弃了传统时序模型中复杂的注意力机制或卷积操作，以极简的架构实现了接近 Transformer 模型的效果。选择 TSMixer 作为对比基准，旨在评估不引入特征回归模型而单纯依靠此模型进行特征和时序处理是否能达到理想效果。

本研究实验软硬件的关键配置参数如表 4.2 所示。

表 4.2 实验环境配置参数

环境名称	型号	版本号或参数
操作系统	Ubuntu	20.04
GPU	NVIDIA Geforce RTX3090	24GB
Python	-	3.7.0
CUDA	-	10.1
Pytorch	-	1.13.1

### 4.3.2 评估指标

#### (1) $R^2$ 系数

本研究基于测试集计算回归模型对数据的拟合优度，即决定系数 (Coefficient of Determination,  $R^2$  Score)，该指标能够以数值形式呈现模型对于目标变量方差的阐释力度。其计算公式为：

$$R^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}} \quad (4.2)$$

$$SS_{\text{res}} = \sum_{i=1}^n (y_{\text{true}}^{(i)} - y_{\text{pred}}^{(i)})^2 \quad (4.3)$$

$$SS_{\text{tot}} = \sum_{i=1}^n (y_{\text{true}}^{(i)} - \bar{y}_{\text{true}})^2 \quad (4.4)$$

其中  $SS_{\text{res}}$  代表模型预测值和真实值之间的差异（模型没能解释的部分）， $SS_{\text{tot}}$  代表真实值与其平均值之间的差异（数据的总变异程度）， $\bar{y}_{\text{true}}$  是真实标签的平均值。

$R^2$  的具体数值含义如下：

- $R^2 = 1$ ：完美预测，模型完全拟合了数据。
- $R^2 = 0$ ：模型的预测能力等同于总是预测平均值（即模型没有学到任何特征与结果的关系）。
- $R^2 < 0$ ：模型非常糟糕，比直接猜平均值还要差。这通常发生在模型严重过拟合或者选错了模型的情况下。

故  $R^2$  越高（越接近 1），说明该回归模型对机械通气时长的预测越准确，能够解释更多的时长变化波动。

## (2) RMSE

在回归模型验证环节中，本研究采用 10 折交叉验证作为核心评估策略，旨在通过系统化的数据重采样与重复验证，对回归模型的泛化能力提供稳健且无偏的估计。该方法将训练集随机划分为 10 个互斥且规模近似的子集（称为“折”），依次将其中 1 折作为验证集，其余 9 折作为训练集，重复训练与验证 10 次，确保每个样本都参与一次验证。每次验证计算均方根误差 (Root Mean Square Error, RMSE) 为：

$$\text{RMSE}^{(k)} = \sqrt{\frac{1}{n_k} \sum_{i=1}^{n_k} (y_i^{(k)} - \hat{y}_i^{(k)})^2} \quad (4.5)$$

其中  $k = 1, 2, \dots, 10$  表示折的编号， $n_k$  表示第  $k$  折的样本数， $y_i^{(k)}$  与  $\hat{y}_i^{(k)}$  分别为该折中第  $i$  个样本的真实值与模型预测值。最终模型性能由 10 次验证的 RMSE 均

值与标准差分别表示为如下：

$$\overline{\text{RMSE}} = \frac{1}{10} \sum_{k=1}^{10} \text{RMSE}^{(k)} \quad (4.6)$$

$$\sigma_{\text{RMSE}} = \sqrt{\frac{1}{9} \sum_{k=1}^{10} \left( \text{RMSE}^{(k)} - \overline{\text{RMSE}} \right)^2} \quad (4.7)$$

这一做法不仅最大化利用了有限的的数据（每轮训练使用约 90% 的样本），还通过多次验证显著降低了因单次数据划分引入的随机偏差，从而更可靠地反映模型在未见数据上的预测稳定性与泛化能力。在医疗预测任务中，该方法尤为重要，它帮助识别模型在不同患者亚群上的表现波动，为后续临床部署提供统计可信的性能保证和误差范围估计。

为了全面、客观地量化预测模型在机械通气时长预测任务上的性能，本研究选取了两个广泛使用的统计学指标：平均绝对误差 (Mean Absolute Error, MAE) 和均方误差 (Mean Squared Error, MSE)。

### (3) MAE

MAE 通过计算预测值与真实值之间绝对差异的均值，反映了模型预测结果偏离真实时长的平均距离，具有良好的临床可解释性。在医学应用中，医生往往更关注预测值平均会产生多大的实际偏差，MAE 对此提供了最直接的度量。其计算公式如下：

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (4.8)$$

其中， $N$  表示测试集样本数量， $y_i$  为第  $i$  个患者的真实机械通气时长， $\hat{y}_i$  为模型的预测时长。

### (4) MSE

MSE 的计算方式表现为预测偏差平方和的算术平均，MSE 对误差进行了平方处理，这表明该指标对幅度较大的预测偏差施加了更为严厉的惩罚，这一点是与 MAE 存在差异的。在机械通气时长的预测中，极端的预测偏差（例如将需要通气 10 天的重症患者误判为 1 天）可能导致严重的医疗资源错配或撤机失败风险。引入 MSE 指标有助于评估模型在处理这些极端情况下的稳定性和鲁棒性。其计算公式如下：

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (4.9)$$

### 4.3.3 实验结果

#### 4.3.3.1 回归实验结果

实验采用遍历机制对4.3.1.1节提到的模型进行逐一评估，将处理过的训练集输入进回归模型进行训练拟合。

拟合结果如表 4.3所示。为了直观评估模型的拟合性能，本研究将测试集上的模型输出结果与实际观测值进行了曲线对比与可视化展示。考虑到可视化清晰度，本章重点展示了测试集中前 20 个样本的拟合情况，通过观察预测曲线（红色）与真实曲线（绿色）的重合度来判断模型的稳定性和偏差，如图 4.3所示。

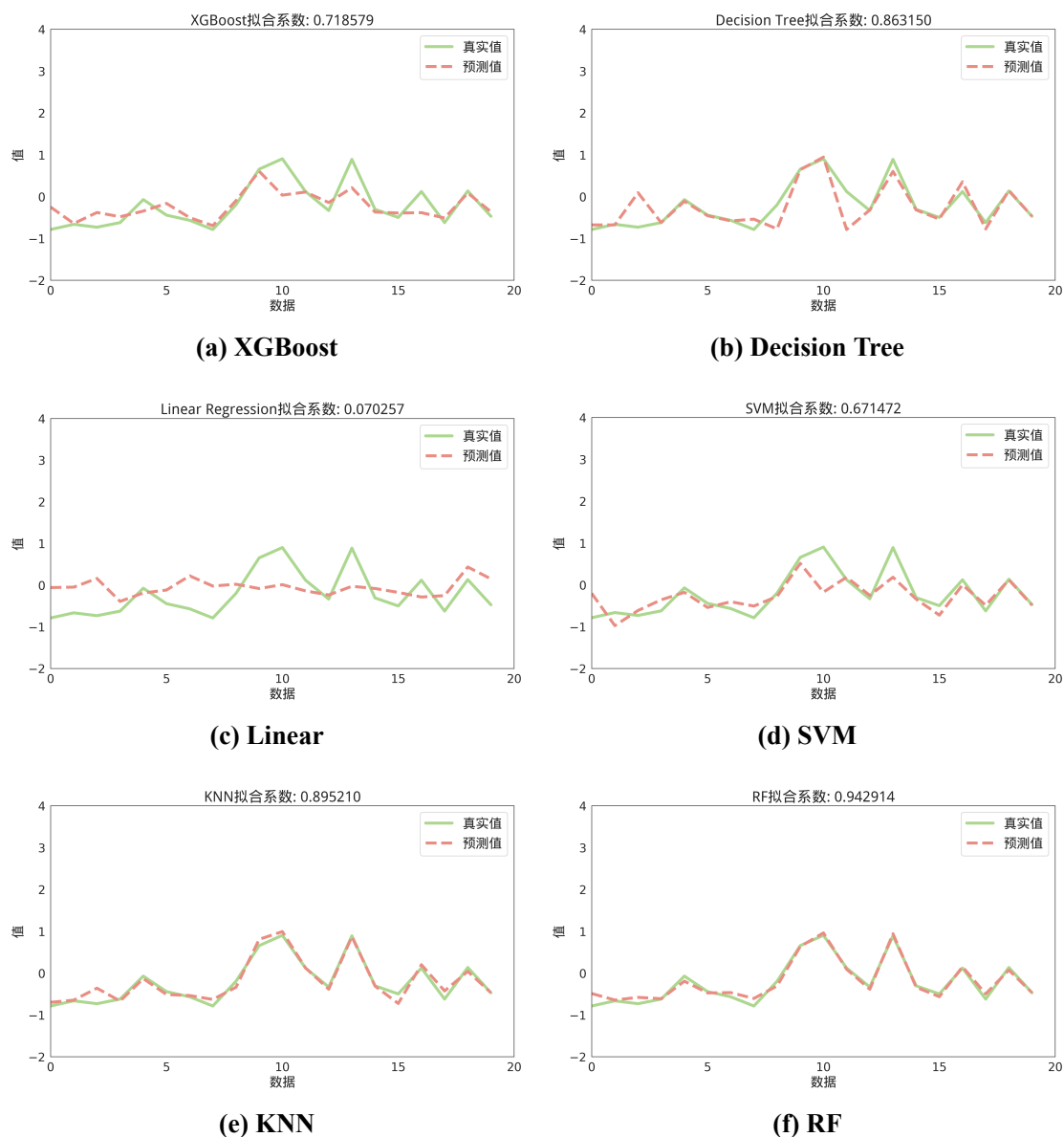


图 4.3 各回归模型的拟合情况图

根据图 4.3可知排行前 3 的表现优秀的模型分别为：RF、Bagging 和 KNN，将它们进行交叉验证，如图 4.4所示，结构表明 RF 回归模型在本实验中取得了

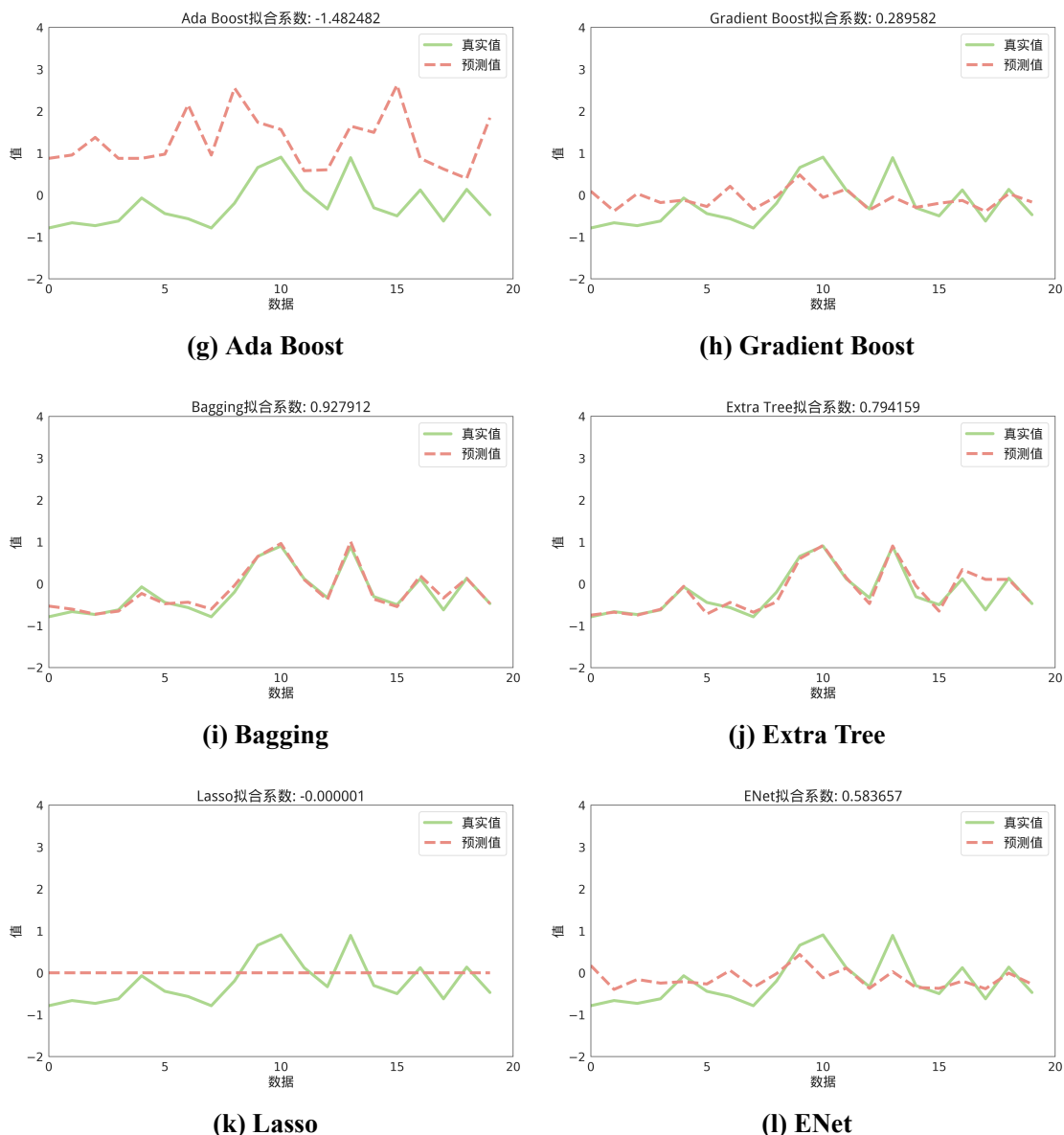


图 4.3 各回归模型的拟合情况图 (续)

最好的效果，通过后续的网格检索调参得到最优参数如下：决策树的数量为 140，随机数生成器的种子为 42。

#### 4.3.3.2 预测实验结果

由回归模型的对比结果可知，RF 回归模型是本实验中表现最优秀的拟合模型，本研究将其作为基准预测器。利用该模型对全体患者的每个记录时刻的生理状态特征进行拟合，生成对应的初步机械通气时长预测值。接着按照 4.2 节所示的方法构建时序数据集。

##### (1) 对比实验结果

本研究将基于上述所提到的时序预测架构进行实验对比，得出适合本时序数据集的时序预测模型。部分训练损失曲线如图 4.5 所示。

表 4.3 各回归模型的性能评分

模型	$R^2$ 系数	模型	$R^2$ 系数
XGBoost	0.718579	Decision Tree	0.863150
Linear	0.070257	SVM	0.671472
KNN	0.895210	<b>RF</b>	<b>0.942914</b>
Ada Boost	-1.482482	Gradient Boost	0.289582
Bagging	0.927912	Extra Tree	0.794159
Lasso	-0.000001	ENet	0.583657

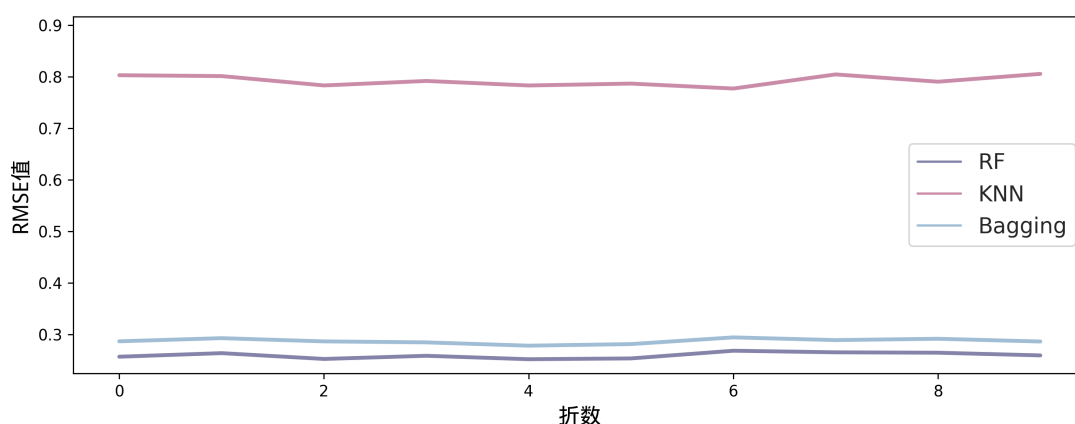


图 4.4 表现优秀的回归模型的交叉验证对比图

具体的对比结果如表 4.4 所示，RMRA-LSTM 模型（即 RF 回归模块结合 MRA-LSTM 时序预测模块）的 MSE 指标和 MAE 指标均取得了最小值，其中该模型对测试集的拟合情况如图 4.6 所示。在本研究的数据集中表现出了优秀的效果。这表明 RMRA-LSTM 在捕捉时序依赖性与模式方面具有更强的能力，有效提升了预测精度。

表现次优的是 R-LSTM 模型，其 MSE 与 MAE 分别为 0.0038 与 0.0361，但仍分别落后 RMRA-LSTM 模型 18.8% 与 11.4%。这表明原始的 LSTM 架构虽然是一种有效的时序建模架构，但本研究提出的结合了多种改进的 RMRA-LSTM 模型在特征利用或序列依赖性建模方面更具优势。

相比之下，其他结合了 Self-Attention 机制的复杂模型，如 R-Transformer、R-Informer 及 R-FEDformer，其预测误差 ( $MSE > 0.009$ ) 均显著高于 RMRA-LSTM，差距超过两倍。这一现象说明，标准的全局注意力机制在本任务的数据特性与预测目标下可能并非最有效的选择，模型需要更精细的时序归纳偏置或更高效的信息压缩机制。此外，RSA-LSTM (RF + LSTM + Self-Attention) 模型也未能达到预期，其误差指标较 RMRA-LSTM 模型高出约 78.1% (MSE) 与 36.4% (MAE)，进

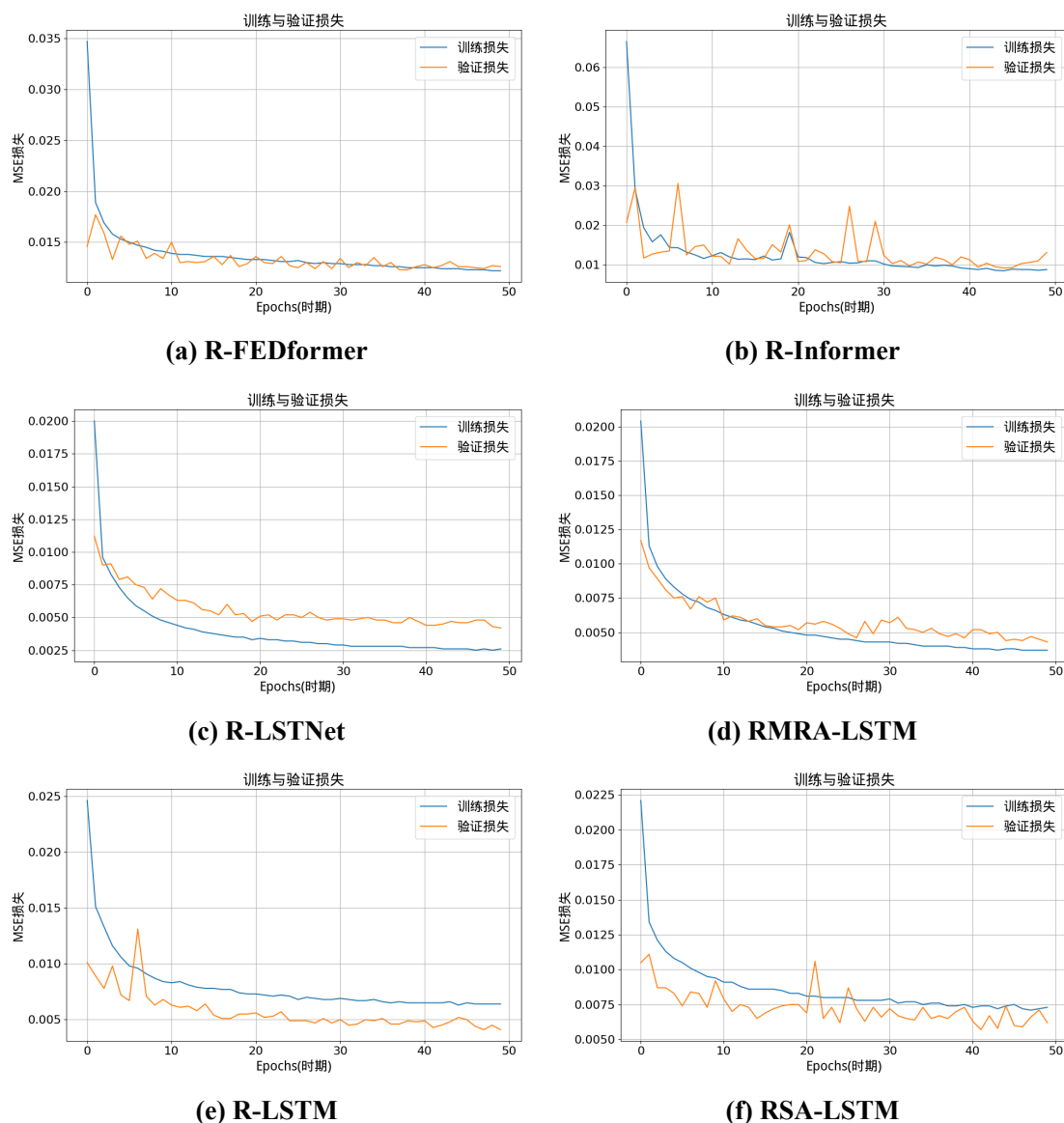


图 4.5 部分预测模型的训练损失曲线

一步印证了并非简单地叠加流行模块就能提升性能。R-TimesNet 模块和 TSMixer 模块虽然是目前很流行的算法，且在多种任务上表现优异，但是它们在本实验中均表现出较差的性能。

根据分析，LSTM 模块之于 Transformer 等变体模块的优势有如下几点：

- 数据集的时间步长  $T = 5$ ，序列长度较短，无法发挥全局注意力优势，而 LSTM 的门控机制更适合捕捉短期时序依赖。

- 数据集的状态维度不高，由于 Transformer 模块在训练阶段涉及海量参数的更新与学习，模型面临较高的过拟合风险，而 LSTM 模块的参数量远小于前者，参数效率更高，减少过拟合风险。

- 计算效率上 LSTM 更优，收敛更快。

- LSTM 的领域适配性更好，符合医疗数据的渐进变化特性。

综上所述，实验结果表明，本文提出的 RMRA-LSTM 模型在多项对比中均

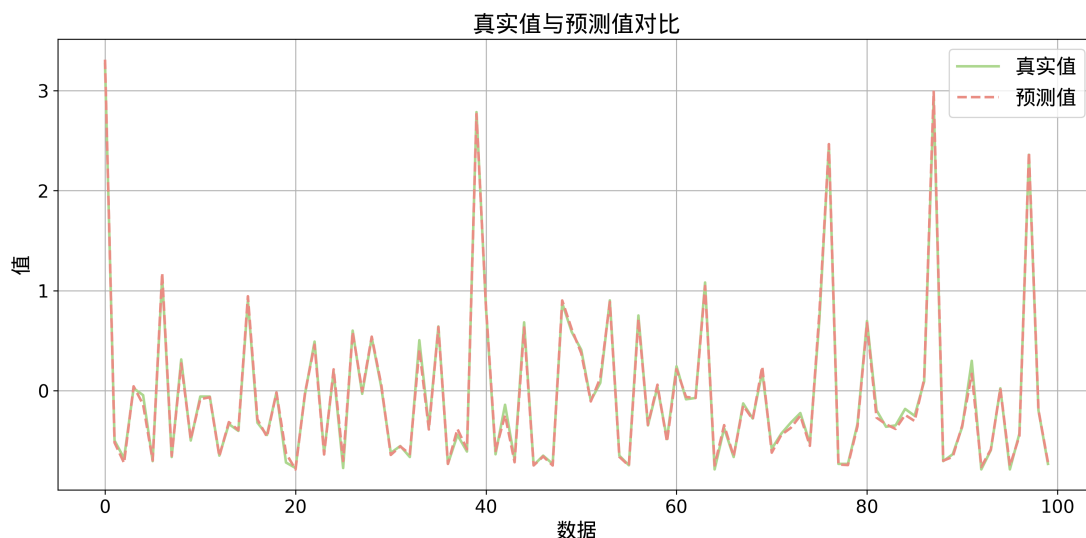


图 4.6 RMRA-LSTM 模型拟合情况图

由于测试集过大，只取前 100 个数据进行图像绘制

展现出领先的预测精度与稳定性，验证了其架构设计的有效性。此外，通过对实验结果的深入剖析，本研究进一步揭示了当前主流注意力机制模型在处理此类任务时可能存在的固有局限。

表 4.4 各预测模型在 MIMIC-III 数据集的实验结果

模型	MSE	MAE
<b>RMRA-LSTM(ours)</b>	<b>0.0032</b>	<b>0.0324</b>
R-LSTNet	0.0043	0.0373
R-LSTM	0.0038	0.0361
RSA-LSTM	0.0057	0.0442
R-Transformer	0.0098	0.0529
R-Informer	0.0110	0.0569
R-FEDformer	0.0121	0.0576
R-TimesNet	0.6899	0.5516
TSMixer	1.8814	0.9661

## (2) 消融实验结果

为检验本研究所设计模型的有效性，本节安排了相应的消融实验以进行验证分析，分别探究了基于 RF 模块和 MRA-LSTM 模块中多尺度 CNN、时序注意力机制与残差连接的性能贡献。本文补充多组消融实验，分别移除 RF 模块、多尺度 CNN 模块、时序注意力模块、残差连接，接着进行结果对比，结果如表 4.5 所示。

由表可知，本文完整模型（设置 E）取得  $MSE=0.0032$ 、 $MAE=0.0324$  的最优

性能，任一核心组件被移除后，预测误差均出现不同程度上升，证明各模块对最终性能均存在不可替代的贡献。具体分析如下：

移除 RF 回归模块（设置 A）性能下降最显著。移除 RF 初步预测模块后，MSE 升至 0.0340、MAE 升至 0.1217，性能大幅劣化。原因在于：RF 模块能够从高维临床特征中高效提取与呼吸机时长强相关的静态特征，并输出基准预测值，为后续时序建模提供稳定的先验信息；缺失该模块会导致后续网络难以捕捉关键影响因素，从而出现大幅预测偏差。

移除多尺度 CNN 模块（设置 B）性能下降。移除多尺度 CNN 后，MSE 由 0.0032 上升至 0.0043，MAE 上升至 0.0368。原因在于：多尺度 CNN 能够从不同时间尺度捕获短期波动与中期趋势，提取局部关键生理变化模式；缺失该模块会导致模型对时序局部细节感知不足，无法有效捕捉病情短时间内的波动规律。

移除时序注意力机制（设置 C）性能持续下降。移除时序注意力后，MSE 上升至 0.0048，MAE 上升至 0.0400。原因在于：时序注意力能够自动加权对时长预测更关键的时间步，聚焦病情转折等重要时刻；缺失该机制会使模型对时序信息的利用变得平均化，无法突出关键时段的决策价值。

移除残差连接（设置 D）性能小幅下降。移除残差连接后，MSE 上升至 0.0049，MAE 上升至 0.0396。原因在于：残差连接可缓解深层 LSTM 的梯度消失问题，保证长时序依赖信息有效传递；缺失该结构会导致网络在训练中难以收敛到最优状态，时序特征学习不充分。

综上，RF 模块、多尺度 CNN、时序注意力、残差连接四者相互配合、缺一不可，共同保障了模型对呼吸机使用时长的精准预测。

表 4.5 RMRA-LSTM 模型的消融实验结果

设置	RF	多尺度 CNN	注意力机制	残差连接	MSE	MAE
A	×	√	√	√	0.0340	0.1217
B	√	×	√	√	0.0043	0.0368
C	√	√	×	√	0.0048	0.0400
D	√	√	√	×	0.0049	0.0396
<b>E(ours)</b>	√	√	√	√	<b>0.0032</b>	<b>0.0324</b>

注：× 表示该模块不被包含，√ 表示该模块被包含。

#### 4.4 本章小结

本章主要围绕呼吸机使用时长的精准预测展开研究，旨在结合患者的静态生理特征与动态时序变化，构建高效的预测模型。

首先，在数据处理层面，本章构建了包含生理状态特征与时序特征的复合数据集。其次，在模型构建与选择层面，本章采用了结合特征回归筛选及时序深度建模的分步策略，将最优的特征回归模型的预测结果作为新特征融入时序输入，构建了级联预测机制。最后，通过严格的对比实验分析，在数据集 MIMIC-III 的结果表明本章提出的 RMRA-LSTM 模型在各项评估指标上均取得了最优效果，MSE 值和 MAE 值分别为 0.0032 和 0.0324。该模型不仅有效利用了 RF 在多维特征处理上的优势，更充分发挥了多尺度卷积网络在局部特征提取上的优势，通过残差连接和时序注意力机制充分挖掘了时序数据中长短时依赖关系的动态演化规律，从而验证了融合多尺度局部特征与全局时序建模在呼吸机使用时长预测任务中的有效性与优越性，为临床资源的合理规划与精准决策提供了支撑。

## 第 5 章 基于专家知识和强化学习的呼吸机参数推荐方法

本章针对呼吸机参数的推荐设定问题，首次提出了一种结合专家知识的参数推荐模型。该模型基于动脉血气分析参数指标通过聚类算法将患者划分为具有相似生理状态的子群，合理将医学先验知识嵌入数据结构，接着基于临床先验知识定义了离线强化学习相关的状态、动作，重构了融合动脉血气分析相关的奖励函数，引导模型兼顾患者的长期预后与短期生理稳定，使得离线强化学习模块的整体性能得到提升，最终构建的模型能够提供更优的参数推荐。

本章的组织结构安排如下：5.1节阐述数据集的预处理与特征工程的流程；5.2节阐述了模型的具体设计，包含聚类模块、RL 问题的定义以及离线强化学习模块；5.3节通过严谨的对比实验、消融实验进行指标评估，系统性验证该模型的有效性、安全性及相较于传统方法的性能提升；5.4节对本章工作进行总结。

### 5.1 数据处理

本章所用的数据处理方式如 3.1节所述。MIMIC-III 数据集的性别比例及年龄分布如图 5.1 所示。此外，本章的实验增加了 2.3.2 节所述的 eICU-CRD 数据集进一步验证模型的有效性，预处理方法类似 MIMIC-III 数据集。

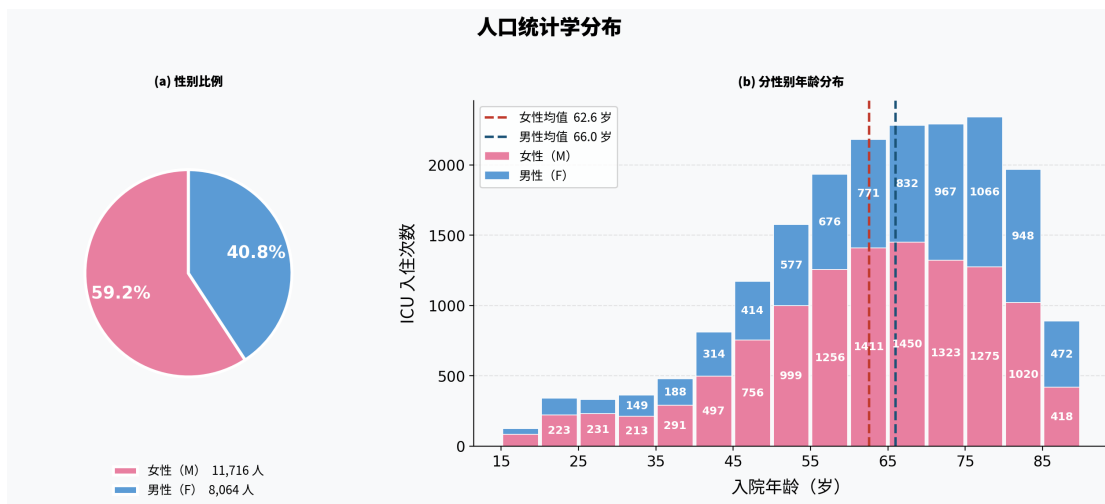


图 5.1 人口统计学分布

### 5.2 模型设计

基于专家知识和强化学习的呼吸机参数推荐模型总体结构图如图 5.2 所示。整体流程从分析患者的动脉血气 (Arterial Blood Gas, ABG) 指标开始，通过 K-means 算法将患者聚类为具有相似生理特征的不同亚组。这一聚类过程作为将

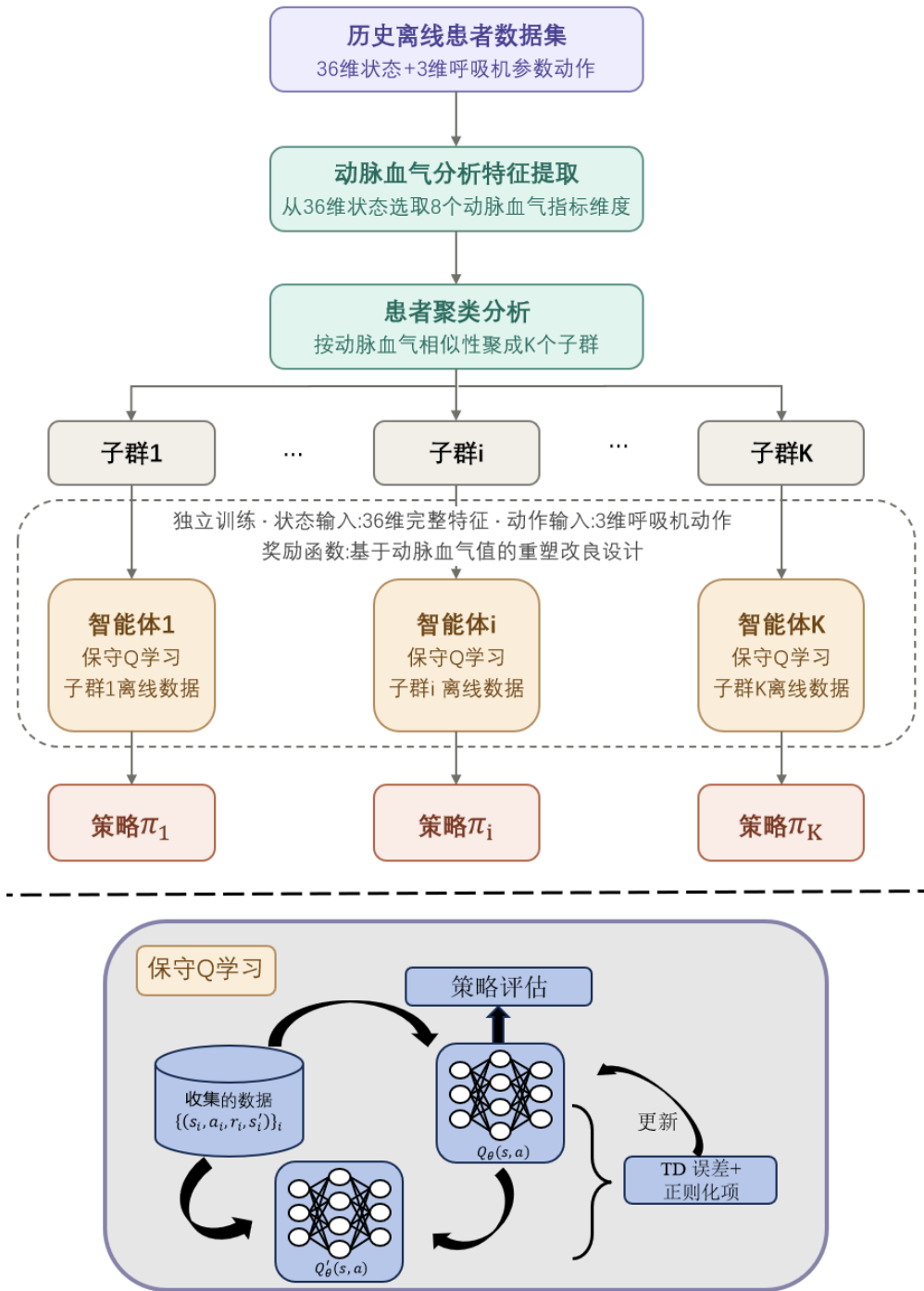


图 5.2 基于专家知识和强化学习的呼吸机参数推荐模型总体结构图

临床先验知识嵌入建模的机制。针对每个识别出的患者集群，使用保守 Q 学习 (Conservative Q-Learning, CQL) 算法分别训练离线强化学习模型，实现针对特定临床特征的个性化策略学习。每个亚模型均在 MIMIC-III 数据集上进行训练，这些数据包含患者状态、行动与重构的奖励。通过聚类融入领域知识，并基于离线强化学习进行训练，该模型旨在生成更贴合临床推理的个体化呼吸机参数推荐策略，最终促进更安全、更有效的机械通气治疗。具体流程见算法 5.1。

### 5.2.1 聚类模块

K-means 聚类是一种基于距离的算法，通过最小化类内平方误差将样本分配到不同的簇中<sup>[81]</sup>。聚类算法通过迭代调整聚类中心来确保分类结果的优化，并且选择合适的簇数以捕捉数据的内在结构。

通过 K-means 聚类基于 ABG 分析指标能够将患者划分为若干类，这些类代表了在 ABG 指标方面具有相似特征的患者群体。每一类群体的特征可以帮助进一步理解不同生理状态下的患者，并为后续的个性化治疗策略的生成提供有力支撑。因此，K-means 聚类在整个技术路线中起到了关键的桥梁作用，将传统的医学知识与先进的机器学习算法有机结合。

筛选出其中个人特征和有效的 ABG 分析参数作为患者的聚类特征，包括年龄，性别，体重，血氧饱和度 (SpO<sub>2</sub>)<sup>①</sup>，血液酸碱度 (pH)，动脉血二氧化碳分压 (PacO<sub>2</sub>)，碱剩余值 (Base Excess)，碳酸氢根值 (bicarbonate)。

---

#### 算法 5.1 基于专家知识和强化学习的呼吸机参数推荐训练算法

---

**Input:** 患者数据集  $\mathcal{D}$ ，动脉血气分析相关特征集合  $F_{abg}$ ，聚类数量  $k$   
**Output:** 每个患者类别对应的最优策略  $\{\pi_1, \pi_2, \dots, \pi_k\}$

- 1 从数据集  $\mathcal{D}$  中提取动脉血气分析相关特征子集:  $X_{abg} = \mathcal{D}[F_{abg}]$ ;
- 2 利用聚类算法对  $X_{abg}$  进行聚类，得到  $k$  个患者类别:  $C = \{C_1, C_2, \dots, C_k\}$ ;
- 3 **for**  $i = 1$  **to**  $k$  **do**
- 4     提取第  $i$  类的患者数据子集构建离线强化学习数据集  $D_i = \{(s_t, a_t, r_t, s_{t+1})\}$ ;
- 5     初始化 Q 网络参数  $\theta_i$ ，并初始化目标网络参数  $\theta_i^- = \theta_i$ ;
- 6     **for**  $epoch = 1$  **to**  $N_{epoch}$  **do**
- 7         从离线数据集  $D_i$  中采样小批量经验  $\{(s_t, a_t, r_t, s_{t+1})\}$ ;
- 8         计算目标 Q 值:  $y_t = r_t + \gamma \max_{a'} Q_{\theta_i^-}(s_{t+1}, a')$ ;
- 9         最小化均方误差损失更新 Q 网络:  $\mathcal{L}(\theta_i) = \mathbb{E}_{(s_t, a_t, r_t, s_{t+1}) \sim D_i} [(Q_{\theta_i}(s_t, a_t) - y_t)^2]$ ;
- 10         利用梯度下降更新参数:  $\theta_i \leftarrow \theta_i - \alpha \nabla_{\theta_i} \mathcal{L}(\theta_i)$ ;
- 11         周期性更新目标网络:  $\theta_i^- \leftarrow \theta_i$ ;
- 12     **end**
- 13     得到第  $i$  类患者的最优策略:  $\pi_i(s) = \arg \max_a Q_{\theta_i}(s, a)$ ;
- 14 **end**
- 15 **return**  $\{\pi_1, \pi_2, \dots, \pi_k\}$ ;

---

### 5.2.2 RL 问题定义

本研究的 MDP 的定义类似于 Peine<sup>[34]</sup>的工作，发作持续时间从患者的开始使用呼吸机进行机械通气的时间点到之后 72 小时。由于呼吸机参数的设定涉及到患者方方面面的状况，信息越多，模型能够学习得更全面，所以状态空间被扩充至 36 维。

#### (1) 状态空间

状态空间包含以下 36 个变量：

- 人口统计学信息：年龄、性别、体重、ICU 再入院情况、Elixhauser 评分。

<sup>①</sup>由于 SaO<sub>2</sub> 指标缺失，这里用 SpO<sub>2</sub> 代替

- 生命体征：SOFA 评分、SIRS 评分、GCS 评分、心率、收缩压、舒张压、平均动脉压、休克指数、体温、血氧饱和度。

- 实验室指标：钾、钠、氯、血糖、血尿素氮、肌酐、镁、二氧化碳、血红蛋白、白细胞计数、血小板计数、凝血酶原时间、国际标准化比值、pH 值、二氧化碳分压、碱剩余、碳酸氢盐。

- 液体相关：尿量、血管加压药使用情况、静脉输液量、累计液体平衡。

## (2) 动作空间

动作空间包含以下 3 个呼吸机设置：

- 理想重量调整的潮汐体积 (Volume of air in and out with each breath adjusted by ideal weight,  $V_t$ )

- 呼气末正压 (Positive End Expiratory Pressure, PEEP)

- 吸入氧气分数 (Fraction of inspired oxygen,  $FiO_2$ )

动作空间  $A$  是这 3 种设置的笛卡尔积，每个设置都可以取七个对应于范围的值。因此，本研究设置每个动作作为元组  $a = (v, o, p)$ ，其中  $v \in V_t, o \in FiO_2, p \in PEEP$ 。

## (3) 奖励函数

强化学习智能体的主要目标是使患者的 ABG 分析指标稳定在正常范围内且保持患者长期生存。这使函数设计不仅需要考虑到治疗期间的患者奖励，还应该防止长期影响的复杂性。对于奖励函数的设置，本研究中的奖励函数部分来源于 Kondrup 等人<sup>[35]</sup>的工作，其定义了一个终端奖励  $r(s_t, a_t, s_{t+1})$ ，当患者 90 天内去世，则取值 -1，否则取 +1。已知仅使用稀疏终端奖励会导致 RL 任务性能不佳，不利于提升样本的利用效率<sup>[82]</sup>，所以考虑将 APACHE-II 评分纳入奖励函数设计中。APACHE-II 是一种被广泛使用的疾病严重程度评分系统<sup>[83]</sup>，用于评估 ICU 中患者的病情，该评分系统通过综合多项生理指标、年龄以及患者的慢性健康状况来预测重症患者的死亡风险。考虑到 ABG 分析对于呼吸机调参的必要性，本研究着重提取了 ABG 分析的指标，即在对 APACHE-II 分数进行中间奖励塑造的基础上增加了 ABG 分析指标。APACHE-II 评分是根据相关生理指标计算的，按照预定范围进行打分，然后汇总，分数越大表明身体状态越糟糕。例如，关于体温：

- 体温在 36.0–38.4°C 范围内记 0 分。
- 体温在 38.5–38.9°C 或 34.0–35.9°C 范围内记 1 分。
- 体温在 32.0–33.9°C 范围内记 2 分。
- 体温在 39.0–40.9°C 或 30.0–31.9°C 范围内记 3 分。
- 体温  $\geq 41.0^\circ\text{C}$  或  $\leq 29.9^\circ\text{C}$  记 4 分。

需要指出的是，APACHE-II 的设计初衷是评估 ICU 患者的整体死亡风险，其

权重分配反映的是各指标对死亡率的预测能力，而非针对如呼吸机参数调整此类特定治疗任务的指导价值。

在呼吸机参数优化任务中，ABG 值具有不可替代的直接指导作用：PaCO<sub>2</sub> 直接受呼吸频率与潮气量调节，pH 反映通气与代谢的综合状态，SpO<sub>2</sub> 作为氧合的连续监测指标可替代 PaO<sub>2</sub> 反映肺换气功能，而碱剩余和碳酸氢根则有助于鉴别酸碱失衡的类型并指导呼吸机参数的精细调整。这些指标共同构成了临床医生调整呼吸机设置的直接依据。

基于上述任务特异性，本研究认为在呼吸机参数优化的奖励函数中，应当适度提升 ABG 指标的相对权重。若沿用原始权重，则相当于将 ABG 指标与其他非直接相关的指标等量齐观，无法体现其在调参中的核心地位；若权重过高，则可能导致智能体过度关注血气状态而忽略感染、肾功能等其他影响患者长期预后的重要因素。因此，本研究将 ABG 分析相关评分的权重设定为原始权重的 2 倍，这一取值显著突出了 ABG 值在调参决策中的指导作用，奖励函数具体设计如下：

$$r = \begin{cases} +1 & t+1 = l_i \text{ 且 } m_{t+1}^i = 1 \\ -1 & t+1 = l_i \text{ 且 } m_{t+1}^i = 0 \\ \frac{2(b_{t+1}^i - b_t^i) + (c_{t+1}^i - c_t^i)}{\max_A - \min_A} & \text{其它} \end{cases} \quad (5.1)$$

其中  $b_t^i$  是患者  $i$  在第  $t$  个时间步的与 ABG 指标相关的 APACHE-II 评分， $c_t^i$  是患者  $i$  在第  $t$  个时间步除与 ABG 指标外的相关 APACHE-II 评分， $m_t^i = 0$  表示患者  $i$  在第  $t$  个时间步已死亡，否则为 1， $l_i$  是患者  $i$  在 ICU 的住院时长， $\max_A, \min_A$  分别是为适应设计而修改的 APACHE-II 评分的最大值和最小值。

### 5.2.3 离线强化学习模块

本研究中创新性地引入了离线强化学习模块，旨在构建一个能够深度挖掘历史临床数据价值的智能呼吸机参数自主推荐系统。该模块旨在突破传统规则控制与在线学习在临床部署中的局限，即无需在与患者的实时交互中探索试错，而是通过从已有的历史诊疗轨迹数据中学习最优策略，实现安全、个性化且可解释的呼吸机参数推荐。具体的框架如图 5.3 所示。

该模块的引入，实现了从经验驱动向数据驱动诊疗模式的转变，具体体现在以下三个核心层面：

- **突破实时交互的安全瓶颈：**传统的在线强化学习依赖于试错机制，这在生命支持系统中存在巨大的伦理风险与安全隐患。本模块通过执行保守策略优化等离线学习算法，直接从大规模的历史诊疗轨迹中提取规律，规避了由于探索性操作导致的通气相关性肺损伤风险，确保了算法从实验室向临床端迁移的安全

性。

- 实现复杂时序下的精准决策：针对呼吸机参数调控中存在的高维非线性与动态演化特性，本模块利用深度神经网络对历史数据中的“患者状态-采取动作”链条进行建模。模块不仅能识别静态的生理指标，更能捕捉参数调整后的长时序反馈效应，从而针对不同表现的危重症患者，产出动态且具备个体差异化适应性的最佳机械通气策略。

- 强化临床可解释性与信任度：为了解决深度学习模型在临床应用中的黑盒质疑，本模块结合了离线策略评估技术。通过在部署前对学习到的策略进行验证评估，预估其奖励值。这种量化的风险评估机制，连同对关键呼吸机参数的显著性分析，赋予了系统更强的医学可解释性，为临床医生的辅助决策提供了有力支撑。

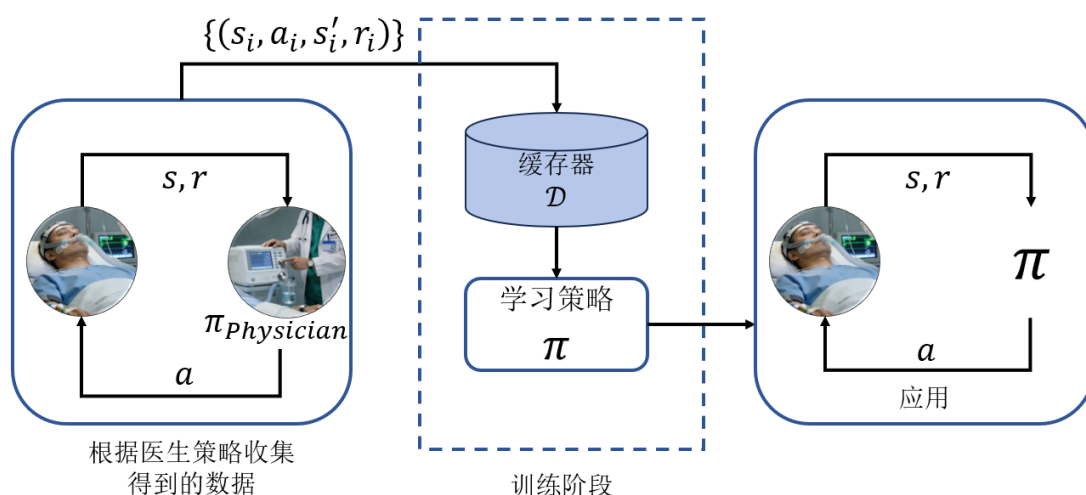


图 5.3 离线强化学习框架图

## 5.3 实验设计与结果分析

### 5.3.1 实验设置

#### 5.3.1.1 聚类模块

本研究将多种经典且具有代表性的聚类方法进行系统性比较，选取了覆盖不同聚类范式、具有广泛学术与应用背景的 5 类算法：K-means 聚类（基于划分）、高斯混合模型 (Gaussian Mixture Model, GMM)（基于概率分布）、凝聚层次聚类 (Agglomerative Hierarchical Clustering)（基于层次合并）、基于密度的空间聚类算法 (Density-Based Spatial Clustering of Applications with Noise, DBSCAN)（基于密度连通性）以及谱聚类 (Spectral Clustering)（基于图划分）。各方法简要介绍如下：

- **K-means** 聚类作为最广泛使用的划分式聚类方法，通过反复迭代实现样本分配：将各数据点划分至距离最近的簇中心，并以簇内所有样本与其对应中心之间平方距离之和的最小化作为优化准则。其假设各类簇呈凸球形分布，算法效率高且易于实现，但对初始中心点敏感，且对非球形簇或尺度不一致的数据效果受限。

- **GMM** 聚类假设数据由若干高斯分布混合生成，采用期望最大化算法进行参数估计，实现基于概率的软聚类。该方法能够捕捉不同协方差结构，适用于各类簇大小不一、重叠程度较高的场景，并为样本提供属于各簇的概率似然，增强了结果的可解释性。

- **Agglomerative Hierarchical** 聚类采用自底向上的聚合策略，在初始阶段将每一个数据样本均视为一个独立的簇，通过逐次合并最相似的簇来构建树状结构。其优势在于无需预先指定簇数，且可通过树状图直观展示不同粒度下的聚类结构，适合探索性分析，但计算复杂度较高，且合并决策不可逆。

- **DBSCAN** 基于样本空间分布的密度进行聚类，通过定义邻域半径与最小样本数来识别高密度区域，该方法不仅具备对任意形态簇的有效分割能力，还能自主检测出数据中的噪声点。该算法不依赖预设的簇数，对噪声与异常值具有较好的鲁棒性，但对密度不均匀的数据及参数设置较为敏感。

- **Spectral** 聚类是一种基于图论的聚类方法，它将聚类任务转化为图的划分问题进行处理。其核心流程为：首先依据样本间的相似性构建相似度矩阵，并在此基础上计算拉普拉斯矩阵的特征向量，接着在该空间中采用 **K-means** 等算法对样本进行聚类划分。该方法通过谱嵌入将数据映射到低维空间，能够处理非凸分布及高维数据，尤其在样本间相似度定义合理时表现出优良性能。

通过这 5 类经典方法对比，可从聚类稳定性、结构适应性、噪声鲁棒性以及在高维医疗特征空间中的可扩展性等多个维度，全面选择出最适合本研究的聚类模型。

### 5.3.1.2 离线强化学习模块

根据 Kondrup 等人<sup>[35]</sup>工作的经验进行参数网格搜索，确定最优的模型参数，然后执行 5 次分别运行 200 万步，并对结果进行平均。本研究实验软硬件的关键配置参数如表 5.1 所示。

#### (1) 对比实验

本研究设置了 7 种方法与 **K-CQL** 进行实验对比：医生策略、**DeepVent** 模型、**DDQN** 模型、**DTE-CQL** 模型、**K-BCQ** 模型、**TD3+BC** 模型和 **DTQL** 模型。具体介绍如下：

- **K-CQL** 模型：本研究提出的方法，在 **DeepVent** 框架基础上融入临床知识

表 5.1 实验环境配置参数

环境名称	型号	版本号或参数
操作系统	Ubuntu	20.04
GPU	NVIDIA Geforce RTX3090	24GB
Python	-	3.7.0
CUDA	-	10.1
Pytorch	-	1.13.1

约束，通过保守的 Q 值估计机制实现安全可靠的策略学习。

- 医生策略：直接使用 MIMIC-III 数据集  $(s_t, a_t, s_{t+1})$  中记录的真实临床决策序列作为基准策略，它代表了实际临床环境中医生群体的集体经验与决策模式。

- DeepVent 模型<sup>[35]</sup>：目前最先进的用于呼吸机调参的深度离线强化学习模型，核心框架是 CQL，本研究将其作为离线强化学习在呼吸机领域的应用基线。

- DDQN 模型：基于双重 Q 学习机制的传统深度 Q 学习模型，用于评估基础 Q 学习算法在呼吸机参数调整任务中的表现。

- DTE-CQL 模型<sup>[37]</sup>：基于 Transformer 编码器架构实现以帮助强化学习智能体更深入地了解患者的病理特征。本研究将其作为对比方法之一，用于验证 Transformer 结构在呼吸机参数优化任务中的协同效果。

- K-BCQ 模型：基于 BCQ 算法的改进版本，通过引入临床知识约束的动作空间限制机制，在保证安全性的同时从离线数据中学习策略。该模型的设置旨在对比现有的表现优秀的离线强化学习模型。

- TD3+BC 模型：在 TD3 算法的基础上仅增加了行为克隆正则化项，并引入了状态特征归一化。该方法通过一个超参数平衡 Q 值最大化与行为克隆的约束，在显著降低计算开销的同时，达到了与当时最先进方法相当的性能，本研究将其作为轻量级离线强化学习算法的代表。

- DTQL 模型：一种融合扩散模型与双策略框架的离线强化学习方法。该方法包含一个用于纯行为克隆的扩散策略和一个用于实际部署的单步策略，通过所提出的扩散信任区域损失将二者桥接，既保留了扩散策略的表达能力，又避免了迭代去噪采样带来的计算开销，本研究将其作为基于生成式模型的最新基线。

## (2) 消融实验

为了证明 K-CQL 模型中基于 ABG 的聚类 and 奖励函数塑造的有效性，本研究设置了消融实验，包含 DeepVent, DeepVent+, K-CQL- 和 K-CQL 的模型比较。DeepVent+ 模型通过引入基于 ABG 的奖励塑造扩展了原始的 DeepVent 模型，从而能够评估奖励重塑是否可以提高模型的性能。K-CQL-模型在 DeepVent 模型的基础上引入了聚类，但不包括与 ABG 相关的中间奖励，它被用作奖励函数的基

线，以评估加入与血气分析相关的中间奖励是否能够提高模型的性能。这三种模型与 K-CQL 模型的模块对比如表 5.2 所示。

表 5.2 DeepVent, DeepVent+, K-CQL- 和 K-CQL 的模块比较

模块	DEEPVENT	DEEPVENT+	K-CQL-	K-CQL
基于 ABG 的 K-means 聚类	×	×	√	√
基于 ABG 的中间奖励	×	√	×	√

注：× 表示该模块不被包含，√ 表示该模块被包含。

### 5.3.2 评估指标

#### 5.3.2.1 聚类效果评估指标

##### (1) 轮廓系数

轮廓系数 (Silhouette Coefficient) 是用于评估聚类质量的指标之一，尤其在 K-means 聚类中，轮廓系数能够帮助判断聚类结果的好坏<sup>[84]</sup>。高轮廓系数表明 K-means 在当前簇数下能够较好地聚类数据，使得数据点在簇内紧密，簇与簇之间分离明确。轮廓系数  $s(i)$  对于每个数据点  $i$  来说，定义如下：

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (5.2)$$

其中  $a(i)$  是数据点  $i$  与同一簇内所有其他样本的平均距离； $b(i)$  是数据点  $i$  与最近的其他簇的所有样本的平均距离。

##### (2) Calinski-Harabasz 指数

Calinski-Harabasz 指数的计算公式如下：

$$CH(k) = \frac{B}{W} \cdot \frac{n - k}{k - 1} \quad (5.3)$$

其中  $B$  为簇间离散度，其计算方式为各聚类中心相对于数据集全局中心的距离平方累加和， $W$  为簇内离散度，即表示所有簇的簇内距离和之和， $k$  为聚类数目， $n$  为总样本数。

##### (3) Davies-Bouldin 指数

Davies-Bouldin 指数的计算公式如下：

$$DB(k) = \frac{1}{k} \sum_{i,j=1}^k \max_{j \neq i} \left( \frac{\bar{d}_i + \bar{d}_j}{d_{ij}} \right) \quad (5.4)$$

其中  $\bar{d}_i$  为簇  $i$  内各个样本到簇中心的平均距离， $d_{ij}$  表示两个簇中心的距离。

Silhouette Coefficient 越接近 1（计算公式见式 5.2），Calinski-Harabasz 指数越高（计算公式见式 5.3），Davies-Bouldin 指数越低（计算公式见式 5.4），表明

簇内凝聚力更强、簇间分离度更大，说明聚类性能更优。结果表明 K-means 聚类算法实现了最好的效果。

### 5.3.2.2 离线强化学习模型评估指标

根据 2.8.3 节所介绍的，在本研究中同时采用 FQE 与 DR 两种离线策略评估方法，对离线强化学习策略进行多角度验证。当两种方法给出的评估结果在趋势上保持一致时，可以在一定程度上提高策略评估结果的可信度，从而为呼吸机参数优化策略的有效性提供更可靠的依据。具体的评估指标是平均初始状态值，该值越高代表应用该模型输出的策略所带来的奖励值越高，策略的表现越佳，模型的性能越好。

### 5.3.3 实验结果

#### 5.3.3.1 聚类实验结果

根据实验设置，本研究将 5.3.1.1 节所述的聚类模块进行实验对比，结果如表 5.3 所示，实验表明 K-means 聚类方法应用在本数据集中表现出了最好的效果。

表 5.3 聚类算法的实验对比结果

聚类算法	K-means	GMM	Agglomerative	DBSCAN	Spectral
Silhouette Score	<b>0.468</b>	0.389	0.373	-0.345	0.237
Calinski-Harabasz Score	<b>16637.743</b>	16582.070	15699.460	2090.266	7848.249
Davies-Bouldin Score	<b>0.889</b>	0.895	0.937	2.048	0.848

通过计算簇数为 2-10 下的平均轮廓系数，如图 5.4 所示可知当簇数为 3 时，轮廓系数的值最大，这表明在该簇数下聚类效果最好，因此选择 3 为最佳簇数。t-Distributed Stochastic Neighbor Embedding (t-SNE) 是一种数据降维和可视化的技术，主要用于将高维数据嵌入到低维空间（通常是二维或三维）中，以便能够直观地理解数据的结构，它的首要任务在于维持高维空间内样本点的局部邻域结构，力求使原本相似的样本在低维映射下仍保持彼此邻近的状态，聚类结果如 5.5 所示。

#### 5.3.3.2 离线强化学习实验结果

本研究首先使用 FQE 研究了 K-CQL 算法的性能，并与其他先进模型进行了比较，K-CQL 的性能是通过对分别在三个数据集上进行的深度离线强化学习训练结果取平均而得到的。结果表明 K-CQL 实现了优越的性能。然后在 K-CQL 上进行了消融实验，以证明基于 ABG 的 K-means 聚类和基于 ABG 的中间奖励

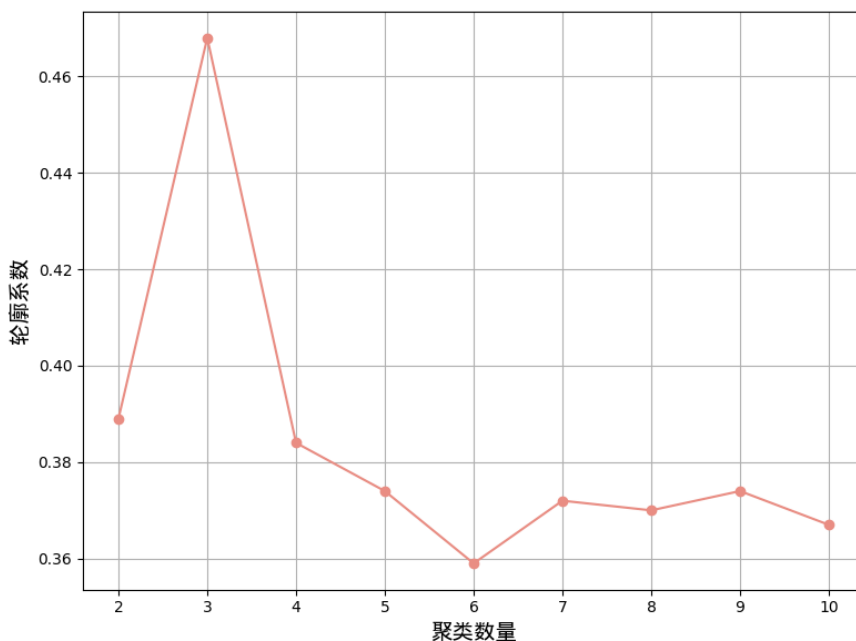


图 5.4 轮廓系数值随聚类数量变化的图

塑造的有效性。为了进一步提高离线评估的可靠性，本研究在 FQE 的基础上引入 DR 算法进行策略评估。随后，本研究分析了 K-CQL 输出策略建议的动作分布，结果显示所选动作符合临床建议。最后，本研究进一步在分布外 (Out-of-Distribution, OOD) 场景中评估了该模型，结果显示 K-CQL 在应用于 OOD 患者时仍保持高性能，使其在实际应用中更安全。

#### (1) 对比实验结果

最终确认了 CQL 的  $\gamma = 0.75$ 、 $\eta = 1e - 6$  和  $\alpha = 0.1$  的最佳值，该深度神经网络由两个隐藏层构成，每层均设置 256 个神经元并采用 ReLU 作为激活函数。

由表 5.4 可得，K-CQL 算法的初始状态值为 0.883，方差为 0.003，实现了最佳的性能，它的效果好于目前最先进的算法 DeepVent，训练过程中它的奖励值随步数的变化如图 5.6 所示。K-CQL 算法输出的策略性能高于医生策略，说明该模型能够通过训练自主学习在临床条件下对患者进行机械通气的最优参数。DTE-CQL 的评估值略大于医生策略，但并不是最优，表明 Transformer 的引入并非最优；DTQL 和 TD3+BC 的评估值为负，表明扩散模型的引入及加入简单的行为克隆项对该场景没有帮助；DDQN 算法的评估值过小，而 K-BCQ 算法的评估值为负，表明这两种算法都无法适应这一场景。由表 5.5 可知，在多中心 eICU-CRD 数据集上，各模型性能均较 MIMIC-III 数据集存在一定下降，表明跨中心数据分布差异会对离线强化学习策略学习带来挑战。但本文提出的 K-CQL 方法仍取得了最高的平均初始状态值，说明该方法在复杂临床数据分布下仍具有一定的稳定性与泛化能力。

#### (2) 消融实验结果

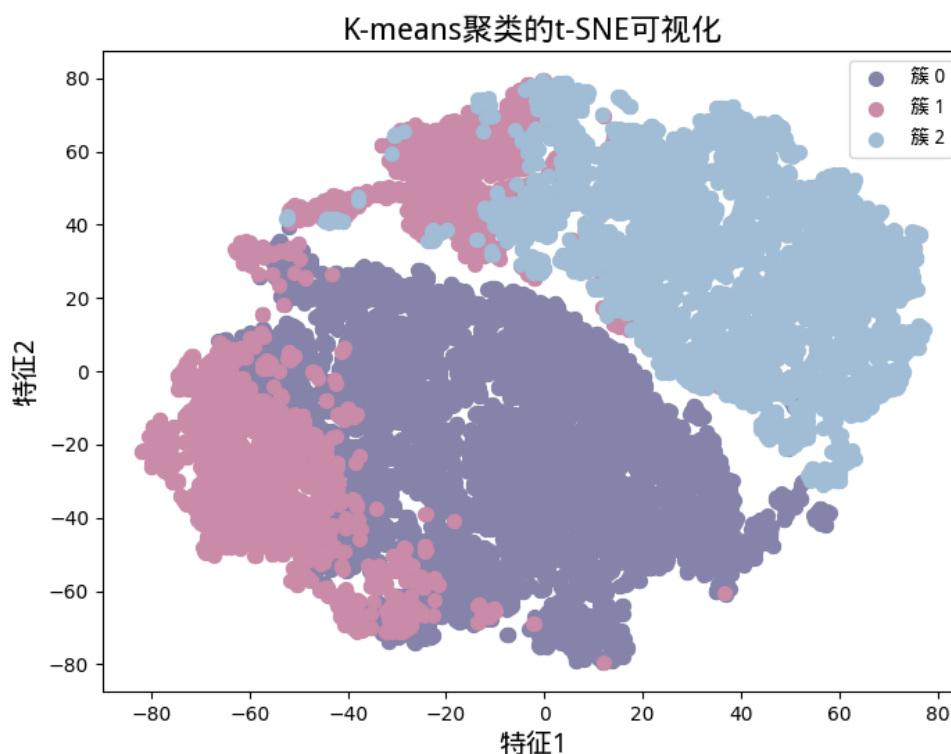


图 5.5 聚类结果图

将多维数据降维到二维数据以便进行可视化

表 5.4 K-CQL 对比实验中各模型在 MIMIC-III 数据集的平均初始状态值

模型	平均初始状态值	模型	平均初始状态值
PHYSICIAN	0.502±0.007	DEEPVENT	0.743±0.005
<b>K-CQL(ours)</b>	<b>0.883±0.003</b>	DDQN	0.050±0.014
K-BCQ	-9.272±0.015	DTE-CQL	0.566±0.004
DTQL	-9.672±0.003	TD3+BC	-13.051±0.002

消融实验中各模型包含的模块如表 5.2 所示，消融实验的结果如表 5.6 所示。结果表明 K-CQL 的性能目前是最先进的 DeepVent 的 1.10 倍。当引入与 ABG 分析相关的中间奖励时，这一倍数增加到 1.19。与 DeepVent 相比，无论是引入 K-means 还是引入基于 ABG 的中间奖励，都进一步提升了模型的性能。因此结果充分证明了基于 ABG 的 K-means 和基于 ABG 的中间奖励塑造的有效性，即引入专家知识确实提升了模型的性能。

### (3) DR 评估结果

基于 DR 算法进行策略评估的结果如表 5.7 所示。

由结果可知 K-CQL 模型的输出策略仍然达到了最好的性能。基于 DR 算法

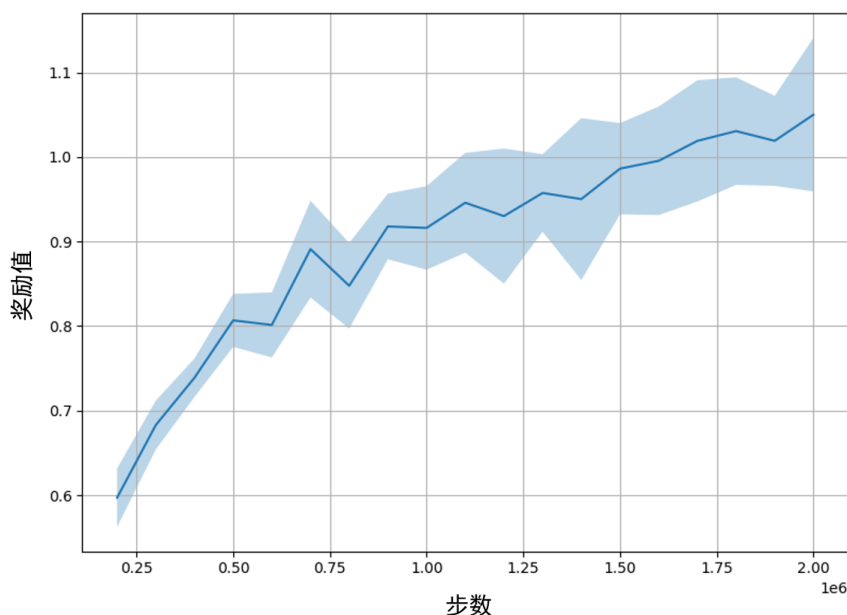


图 5.6 K-CQL 算法训练过程中奖励值随训练步数的变化曲线图

表 5.5 K-CQL 对比实验中各模型在 eICU-CRD 数据集的平均初始状态值

模型	平均初始状态值	模型	平均初始状态值
PHYSICIAN	0.599±0.005	DEEPVENT	0.666±0.008
<b>K-CQL(ours)</b>	<b>0.748±0.006</b>	DDQN	0.023±0.002
K-BCQ	-8.734±0.021	DTE-CQL	0.244±0.024
DTQL	0.208±0.012	TD3+BC	-12.406±0.011

表 5.6 K-CQL 消融实验中各模型的平均初始状态值

模型	平均初始状态值	模型	平均初始状态值
DEEPVENT	0.743±0.005	DEEPVENT+	0.765±0.004
<b>K-CQL(ours)</b>	<b>0.883±0.003</b>	K-CQL-	0.814±0.001

进行评估的结果值均大于基于 FQE 算法进行评估的结果值，这是因为 FQE 方法通过拟合 Q 函数来估计目标策略的价值，其性能高度依赖于函数逼近模型的准确性。当模型容量有限或数据分布存在偏差时，所学习到的  $Q_{\pi}(s, a)$  往往会产生一定的保守估计，从而导致策略价值被低估。

相比之下，DR 方法在利用 Q 函数估计的基础上引入了重要性采样项，通过真实奖励信号对模型估计误差进行校正。当  $Q_{\pi}(s, a)$  存在低估偏差时，DR 中的修正项  $\rho_t(r_t - Q_{\pi}(s_t, a_t))$  可以利用实际观测到的奖励对价值函数进行补偿，从而得到更接近真实策略价值的估计结果，因此会存在 DR 的估计值高于 FQE 的现

表 5.7 基于 DR 算法评估模型的输出策略的平均初始状态值

模型	平均初始状态值	模型	平均初始状态值
DEEPVENT	1.289±0.003	<b>K-CQL(ours)</b>	<b>1.341± 0.015</b>
DDQN	0.597±0.010	K-BCQ	-7.748±0.028
DTQL	-9.339±0.015	TD3+BC	-14.748±0.009

象。

#### (4) 模型输出的动作分布

接下来,本研究评估了 K-CQL 相对于 DeepVent、K-BCQ、DDQN 及医生策略的动作分布。PEEP 经常被设定为 5 cmH<sub>2</sub>O,但它可以根据生理参数的变化进行个性化设置<sup>[85]</sup>,由图 5.7 可知 K-CQL 在 0-5 cmH<sub>2</sub>O 范围内做出大多数建议。高 PEEP 设置与更高的气压伤和气胸发生率显著相关<sup>[86]</sup>,因此应避免。在 FiO<sub>2</sub> 方面, K-CQL 再次遵循临床护理标准。更具体地说, K-CQL 在与数据集中的医生相同的范围内选择了动作,即在 35-50% 和大于 55% 的范围内提出了许多建议。最后,对于理想的 V<sub>t</sub>,最优值通常在 4-8 ml/kg 范围内<sup>[87]</sup>。K-CQL 在 5-10ml/kg 内做出了大部分选择。总体而言,可以观察到,与医生的策略相比, K-CQL 能够为患者提供安全建议,输出的策略满足临床标准。

#### (5)OOD 情况下的 K-CQL 验证

BCQ 和 CQL 是两种常见的比较流行的用于离线强化学习的算法,它们都旨在从离线数据中学习策略,减少了 Q 值高估的问题。Q 值高估是离线强化学习中的一个常见问题,在临床环境中,这可能会导致产生对患者有危险的参数建议。因此,本文研究了 K-BCQ 和 K-CQL 提出的建议是否能够对抗分布外 (Out-of-Distribution, OOD) 状态-动作对的高估。与 Kondrup<sup>[35]</sup>的工作类似,本研究生成了患者的 OOD 数据集来探索强化学习中高估的问题。异常值患者被定义为在分布的顶部和底部 1% 内的机械通气开始时的状态特征的人,大约 25% 的患者被归类为异常值。本研究分别计算数据集分布内 (In-Distribution, ID) 和分布外 (Out-of-Distribution, OOD) 上训练的 FQE 估计 K-CQL 和 K-BCQ 的平均初始 Q 值。

由于本研究的数据集中没有中间奖励的轨迹的最大回报设置为 1,并且在数据集上训练 FQE,因此高于 1 的值应被视为高估。由图 5.8 观察到 K-BCQ 在 ID 和 OOD 设置中存在高估 Q 值的情况,此外, K-BCQ 的高估在 OOD 设置中会加剧,由此可得 BCQ 算法在该场景并不能完全消除高估的情况。CQL 比 BCQ 更加直接且系统地处理 Q 值高估问题, CQL 的核心思想是主动抑制高估,因此避免了这些问题。K-CQL 的平均初始状态值估计在两种设置中都保持在 1 的高估阈值以下,并且 OOD 几乎没有变化,这表明模型在两种设置中的稳定性。

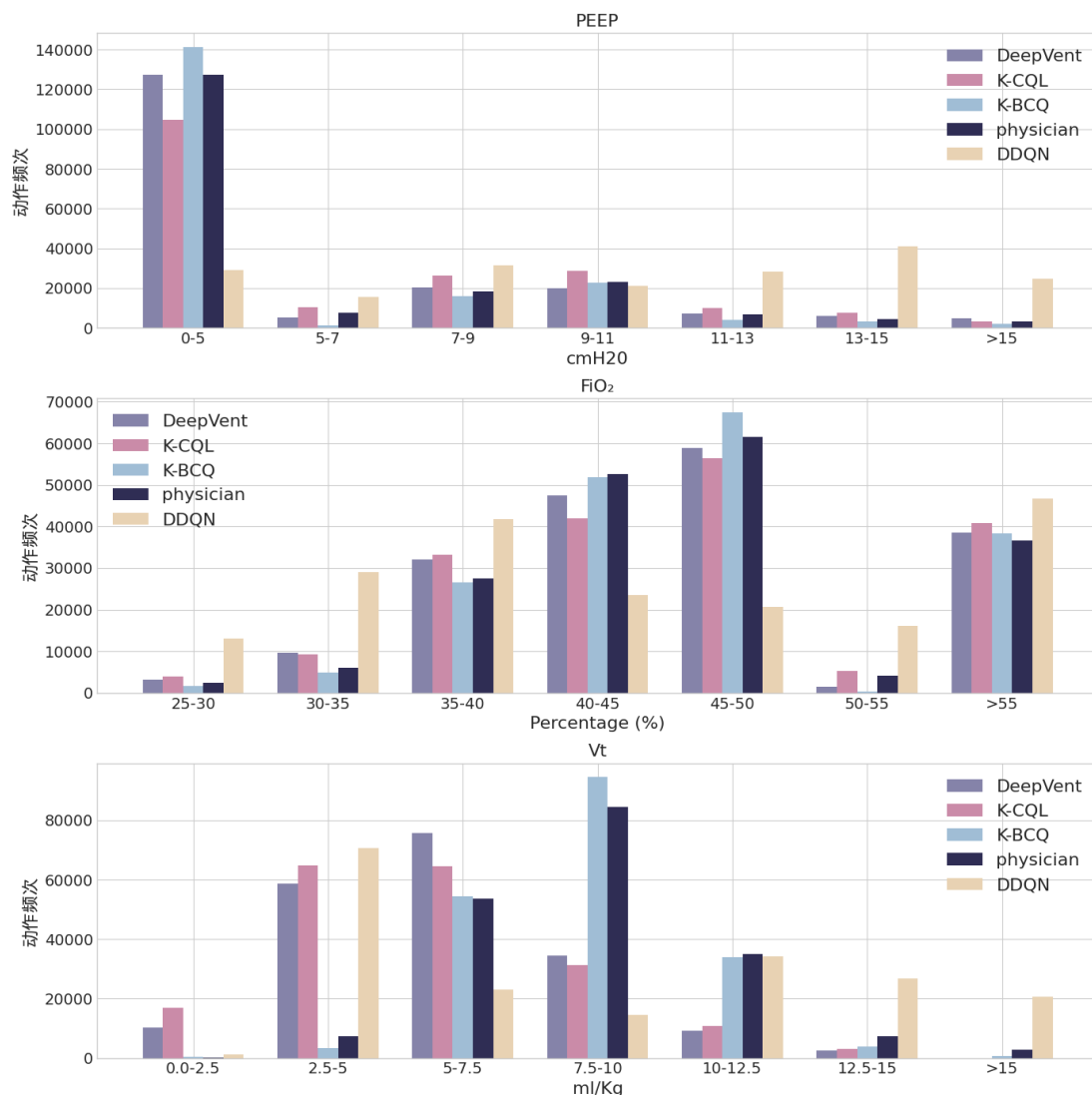


图 5.7 各模型输出的策略的机械通气动作分布图

横轴表示离散化的动作范围，纵轴表示在测试集中每个动作被选择的频率。

## 5.4 本章小结

本章聚焦于机械通气治疗中的呼吸机参数推荐问题，提出了一种融合专家知识与强化学习的 K-CQL 模型，旨在为临床提供安全且高效的辅助决策支持。

首先，本章构建了融合医学知识的聚类模型。通过 K-means 算法，利用动脉血气分析参数指标将患者划分为具有相似生理状态的三个子群，成功将医学先验知识嵌入数据结构。其次，针对强化学习中传统稀疏奖励难以训练的问题，设计了包含长期生存奖励、APACHE-II 评分及动脉血气分析指标相关的奖励函数，引导模型兼顾患者的长期预后与短期生理稳定。在此基础上，构建了结合聚类与保守 Q 学习的 K-CQL 模型。最后，基于 FQE 与 DR 两种离线策略评估方法对比了多种基线模型，验证了模型的有效性和优越性。在数据集 MIMIC-III 及 eICU-CRD 数据集集中的实验结果表明：

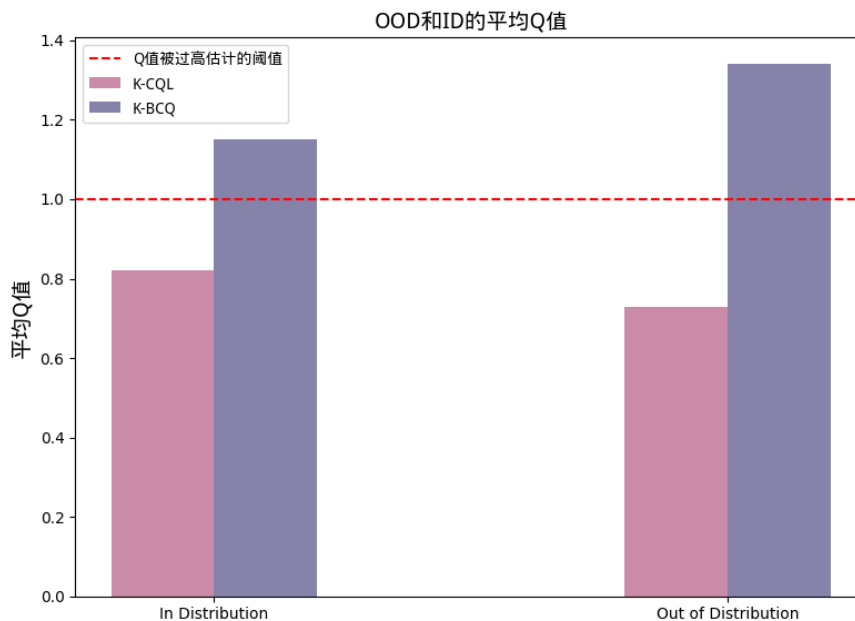


图 5.8 K-BCQ 和 K-CQL 在 ID 和 OOD 状态下的初始状态估计值

平均初始 Q 值超过 1 被认为是 Q 值高估。

(1) 性能提升: K-CQL 在期望回报上优于医生策略及现有表现最优秀的 Deep-Vent 模型, 且消融实验证实了引入动脉血气分析相关的预分类以及中间奖励的有效性。

(2) 临床安全性: 模型推荐的呼吸机参数 (PEEP,  $FiO_2$ , Vt) 分布符合临床安全标准, 避免了可能导致气压伤或气胸的高风险操作。

(3) 鲁棒性: 在 OOD 数据实验中, K-CQL 有效抑制了 Q 值高估现象, 表现出更强的稳定性与可靠性。

综上所述, 本章提出的 K-CQL 模型不仅在算法性能上实现了突破, 更在临床安全性与鲁棒性上展现了应用潜力。

## 第6章 总结与展望

### 6.1 研究总结

本论文针对机械通气治疗中呼吸机启用时机决策、使用时长预测、个性化参数推荐三大核心问题，开展数据驱动与临床知识融合的方法研究，分别构建三类模型并完成实验验证，形成覆盖呼吸机全流程的辅助决策体系，主要工作总结如下：

(1) 针对呼吸机启用时机难以精准判断、现有模型缺乏临床硬约束的问题，提出了一种融合临床约束规则的集成决策模型。在数据层面，对 MIMIC-III 重症数据集进行多维度生理特征提取与缺失值预处理；在模型层面，结合临床约束、类别不平衡处理模块及集成模块搭建模型，预测患者需要使用机械通气的概率，进而判断患者是否需要使用呼吸机。所提模型在公开数据集 MIMIC-III 中进行实验验证，ROC 值为 0.9171，PR 值为 0.9483，结果表明该模型优于现有的模型。

(2) 针对呼吸机使用时长预测精度不足、时序特征利用不充分的问题，提出了一种融合特征回归筛选和时序深度建模的分步增强预测模型 RMRA-LSTM。在数据层面，构建包含患者入院基线与动态时序指标的数据集；在模型层面，通过随机森林回归模块完成关键特征筛选与初步时长预测，接着设计多尺度 CNN、残差 LSTM 与时序注意力机制搭建时序预测模型，精准捕捉病情动态演变规律。在公开数据集 MIMIC-III 中验证模型效果，所提模型的 MSE 值为 0.0032，MAE 值为 0.0324，实验表明该模型优于现有的模型。

(3) 针对呼吸机推荐参数缺乏临床知识、安全性与适用性不足的问题，提出一种嵌入专家知识的离线强化学习参数推荐模型 K-CQL。在数据层面，提取动脉血气分析指标；在模型层面，通过 K-means 聚类实现患者亚群划分，将临床先验嵌入数据结构，设计动脉血气分析指标与 APACHE-II 评分相结合的复合奖励函数，兼顾短期稳定与长期预后。结合深度离线强化学习模块，搭建了增强型智能参数推荐模型，进而将临床医生的专业知识与数据驱动算法相结合，生成有效的呼吸机参数设置策略供医生作为参考。在公开数据集 MIMIC-III 及 eICU-CRD 数据集集中的实验表明该模型生成的策略的平均状态初始值为 0.883，优于现有的模型。

### 6.2 研究展望

本文在呼吸机辅助决策方向取得阶段性成果，但仍存在可拓展空间，未来研究方向如下：

(1) 提升模型对多模态数据的利用能力。在现有数据维度基础上，引入多模态信息融合与高精度时序补全技术，整合高频监护、呼吸机波形、超声影像等多元数据，增强模型对病情复杂变化的捕捉能力，提升呼吸机参数调节的时效性与预警准确性。

(2) 提升模型在小样本与异质性数据下的泛化能力。在现有数据驱动模型基础上，引入迁移学习、小样本学习，对小规模本地数据、罕见病患者及多品牌呼吸机异构数据进行自适应学习，实现快速微调适配，提升模型在不同医院、不同设备、不同患者群体中的稳定性与通用性。

(3) 推进模型与临床信息系统的深度集成。现有研究已验证模型的算法有效性，但如何将其嵌入真实医疗环境仍是关键挑战。未来可探索将三类模型封装为标准化决策支持模块，通过与电子病历系统、床旁监护仪及呼吸机设备对接，实现实时数据采集与模型推理的闭环联动，并设计面向临床医生的可视化决策看板，以直观方式呈现推荐结果及其特征依据，从而在明确辅助而非替代医生决策的边界前提下，将研究成果切实转化为临床实践价值。

## 参考文献

- [1] ZHANG B, QIU X, TAN X. Balancing therapeutic effect and safety in ventilator parameter recommendation: an offline reinforcement learning approach[J]. *Engineering Applications of Artificial Intelligence*, 2024, 131: 107784.
- [2] PHAM T, BROCHARD L J, SLUTSKY A S. Mechanical ventilation: state of the art[J]. *Mayo Clinic Proceedings*, 2017, 92(9): 1382-1400.
- [3] ZEIN H, BARATLOO A, NEGIDA A, et al. Ventilator weaning and spontaneous breathing trials; an educational review[J]. *Emergency*, 2016, 4(2): 65.
- [4] CHASTRE J, FAGON J Y. Ventilator-associated pneumonia[J]. *American Journal of Respiratory and Critical Care Medicine*, 2002, 165(7): 867-903.
- [5] JOSEPH N M, SISTLA S, DUTTA T K, et al. Ventilator-associated pneumonia: a review[J]. *European Journal of Internal Medicine*, 2010, 21(5): 360-368.
- [6] 刘永萍, 单艳华, 李云. 脓毒性休克机械通气患者并发呼吸机相关性肺炎的影响因素分析及预测模型构建[J]. *中国医科大学学报*, 2024, 53(5): 473-477.
- [7] MARTI J, HALL P, HAMILTON P, et al. One-year resource utilisation, costs and quality of life in patients with acute respiratory distress syndrome (ARDS): secondary analysis of a randomised controlled trial[J]. *Journal of Intensive Care*, 2016, 4(1): 56.
- [8] FIGUEROA-CASAS J B, DWIVEDI A K, CONNERY S M, et al. Predictive models of prolonged mechanical ventilation yield moderate accuracy[J]. *Journal of Critical Care*, 2015, 30(3): 502-505.
- [9] GIANG C, CALVERT J, RAHMANI K, et al. Predicting ventilator-associated pneumonia with machine learning[J]. *Medicine*, 2021, 100(23): e26246.
- [10] ZHU Y, ZHANG J, WANG G, et al. Machine learning prediction models for mechanically ventilated patients: analyses of the MIMIC-III database[J]. *Frontiers in Medicine*, 2021, 8: 662340.
- [11] ASHRAFIN, ABDOLLAHI A, ALAEI K, et al. Enhanced prediction of ventilator-associated pneumonia in patients with traumatic brain injury using advanced machine learning techniques [J]. *Scientific Reports*, 2025, 15(1): 11363.
- [12] WU Z, LIU Y, XU J, et al. A ventilator-associated pneumonia prediction model in patients with acute respiratory distress syndrome[J]. *Clinical Infectious Diseases*, 2020, 71(Supplement\_4): S400-S408.
- [13] AUSTIN P C, TU J V, HO J E, et al. Using methods from the data-mining and machine-

- learning literature for disease classification and prediction: a case study examining classification of heart failure subtypes[J]. *Journal of Clinical Epidemiology*, 2013, 66(4): 398-407.
- [14] CHERIFA M, PIRRACCHIO R. What every intensivist should know about big data and targeted machine learning in the intensive care unit[J]. *Revista Brasileira de Terapia Intensiva*, 2019, 31: 444-446.
- [15] GUTIERREZ G. Artificial intelligence in the intensive care unit[J]. *Critical Care*, 2020, 24(1): 101.
- [16] TROCHE G, MOINE P. Is the duration of mechanical ventilation predictable?[J]. *Chest*, 1997, 112(3): 745-751.
- [17] SAYED M, RIAÑO D, VILLAR J. Novel criteria to classify ARDS severity using a machine learning approach[J]. *Critical Care*, 2021, 25(1): 150.
- [18] LIU C F, HUNG C M, KO S C, et al. An artificial intelligence system to predict the optimal timing for mechanical ventilation weaning for intensive care unit patients: a two-stage prediction approach[J]. *Frontiers in Medicine*, 2022, 9: 935366.
- [19] KUO H J, CHIU H W, LEE C N, et al. Improvement in the prediction of ventilator weaning outcomes by an artificial neural network in a medical ICU[J]. *Respiratory Care*, 2015, 60(11): 1560-1569.
- [20] RAZA A, RUSTAM F, SIDDIQUI H U R, et al. Ventilator pressure prediction employing voting regressor with time series data of patient breaths[J]. *Health Informatics Journal*, 2025, 31(1): 14604582241295912.
- [21] PARRECO J, HIDALGO A, PARKS J J, et al. Using artificial intelligence to predict prolonged mechanical ventilation and tracheostomy placement[J]. *Journal of Surgical Research*, 2018, 228: 179-187.
- [22] WANG Z, ZHANG L, HUANG T, et al. Developing an explainable machine learning model to predict the mechanical ventilation duration of patients with ARDS in intensive care units [J]. *Heart & Lung*, 2023, 58: 74-81.
- [23] SAYED M, RIAÑO D, VILLAR J. Predicting duration of mechanical ventilation in acute respiratory distress syndrome using supervised machine learning[J]. *Journal of Clinical Medicine*, 2021, 10(17): 3824.
- [24] SCHWAGER E, LIU X, NABIAN M, et al. Machine learning prediction of the total duration of invasive and non-invasive ventilation during ICU stay[J]. *PLOS Digital Health*, 2023, 2(9): e0000289.
- [25] VILLAR J, GONZÁLEZ-MARTÍN J M, FERNÁNDEZ C, et al. Predicting the length of mechanical ventilation in acute respiratory disease syndrome using machine learning: the PIONEER study[J]. *Journal of Clinical Medicine*, 2024, 13(6): 1811.

- [26] SUO D, AGARWAL N, XIA W, et al. Machine learning for mechanical ventilation control [A/OL]. arXiv (2021). <https://arxiv.org/abs/2102.06779>.
- [27] SIU R, ABBAS J J, FULLER D D, et al. Autonomous control of ventilation through closed-loop adaptive respiratory pacing[J]. *Scientific Reports*, 2020, 10(1): 21903.
- [28] AI W, SURESH V, ROOP P S. Development of closed-loop modelling framework for adaptive respiratory pacemakers[J]. *Computers in Biology and Medicine*, 2022, 141: 105136.
- [29] LI-WEI H L, ADAMS R P, MAYAUD L, et al. A physiological time series dynamics-based approach to patient monitoring and outcome prediction[J]. *IEEE Journal of Biomedical and Health Informatics*, 2014, 19(3): 1068-1076.
- [30] AKBULUT F P, AKKUR E, AKAN A, et al. A decision support system to determine optimal ventilator settings[J]. *BMC Medical Informatics and Decision Making*, 2014, 14(1): 3.
- [31] ORUGANTI VENKATA S S, KOENIG A, PIDAPARTI R M. Mechanical ventilator parameter estimation for lung health through machine learning[J]. *Bioengineering*, 2021, 8(5): 60.
- [32] PENG H, DU B, LIU M, et al. Dynamic graph convolutional network for long-term traffic flow prediction with reinforcement learning[J]. *Information Sciences*, 2021, 578: 401-416.
- [33] SUTTON R S, BARTO A G. Reinforcement learning: an introduction[M]. London, England: MIT press Cambridge, 2014: 1-15.
- [34] PEINE A, HALLAWA A, BICKENBACH J, et al. Development and validation of a reinforcement learning algorithm to dynamically optimize mechanical ventilation in critical care [J]. *NPJ Digital Medicine*, 2021, 4(1): 32.
- [35] KONDRUP F, JIRALERSPONG T, LAU E, et al. Towards safe mechanical ventilation treatment using deep offline reinforcement learning[C]//*Proceedings of the AAAI Conference on Artificial Intelligence: Vol. 37*. Menlo Park, CA: AAAI, 2023: 15696-15702.
- [36] CHEN S, QIU X, TAN X, et al. A model-based hybrid soft actor-critic deep reinforcement learning algorithm for optimal ventilator settings[J]. *Information Sciences*, 2022, 611: 47-64.
- [37] YUAN Y, SHI J, YANG J, et al. Conservative q-learning for mechanical ventilation treatment using diagnose transformer-encoder[C]//*2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. Piscataway,NJ: IEEE, 2023: 2346-2351.
- [38] NI J, HUANG Z, CHENG J, et al. An effective recommendation model based on deep representation learning[J]. *Information Sciences*, 2021, 542: 324-342.
- [39] 王永言. 呼吸机智能通气关键技术研究[D]. 山东大学, 2023.
- [40] MA X, WANG M, LIN S, et al. Knowledge and data-driven prediction of organ failure in critical care patients[J]. *Health Information Science and Systems*, 2023, 11(1): 7.
- [41] 孟祥福, 石皓源. 基于 Transformer 模型的时序数据预测方法综述[J]. *计算机科学与探索*, 2025, 19(01): 45-64.

- [42] 康梦轩, 宋俊平, 范鹏飞, 等. 基于深度学习的网络流量预测研究综述[J]. 计算机工程与应用, 2021, 57(10): 1-9.
- [43] DEY R, SALEM F M. Gate-variants of gated recurrent unit (GRU) neural networks[C]//2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS). Boston, MA: IEEE, 2017: 1597-1600.
- [44] ZHOU H, ZHANG S, PENG J, et al. Informer: beyond efficient transformer for long sequence time-series forecasting[C]//Proceedings of the AAAI Conference on Artificial Intelligence: Vol. 35. Menlo Park, CA: AAAI, 2021: 11106-11115.
- [45] ZHOU T, MA Z, WEN Q, et al. Fedformer: frequency enhanced decomposed transformer for long-term series forecasting[C]//Proceedings of Machine Learning Research: Vol. 162 Proceedings of the 39th International Conference on Machine Learning. New York: PMLR, 2022: 27268-27286.
- [46] KUMAR A, ZHOU A, TUCKER G, et al. Conservative q-learning for offline reinforcement learning[C]//Advances in Neural Information Processing Systems (NeurIPS): Vol. 33. Red Hook, NY: Curran Associates, Inc, 2020: 1179-1191.
- [47] KOSTRIKOV I, NAIR A, LEVINE S. Offline reinforcement learning with implicit q-learning [A/OL]. arXiv (2021). <https://arxiv.org/abs/2110.06169>.
- [48] FUJIMOTO S, MEGER D, PRECUP D. Off-policy deep reinforcement learning without exploration[C]//Proceedings of Machine Learning Research: Vol. 97 Proceedings of the 36th International Conference on Machine Learning. New York: PMLR, 2019: 2052-2062.
- [49] FUJIMOTO S, GU S S. A minimalist approach to offline reinforcement learning[C]//Advances in Neural Information Processing Systems (NeurIPS): Vol. 34. Red Hook, NY: Curran Associates, Inc, 2021: 20132-20145.
- [50] CHEN T, WANG Z, ZHOU M. Diffusion policies creating a trust region for offline reinforcement learning[C]//Advances in Neural Information Processing Systems (NeurIPS): Vol. 37. Red Hook, NY: Curran Associates, Inc., 2024: 50098-50125.
- [51] CHEN L, LU K, RAJESWARAN A, et al. Decision transformer: reinforcement learning via sequence modeling[C]//Advances in Neural Information Processing Systems (NeurIPS): Vol. 34. Red Hook, NY: Curran Associates, Inc, 2021: 15084-15097.
- [52] 朱思明. 生理学复习提纲[M]. 北京: 科学出版社, 2001: 87-96.
- [53] JOHNSON A E, POLLARD T J, SHEN L, et al. MIMIC-III, a freely accessible critical care database[J]. Scientific Data, 2016, 3(1): 1-9.
- [54] POLLARD T J, JOHNSON A E, RAFFA J D, et al. The eICU collaborative research database, a freely available multi-center database for critical care research[J]. Scientific data, 2018, 5(1): 180178.

- [55] HARRELL F E, et al. Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis[M]. New York: Springer, 2001.
- [56] BREIMAN L. Random forests[J]. Machine Learning, 2001, 45(1): 5-32.
- [57] LIU Y, WANG Y, ZHANG J. New machine learning algorithm: random forest[C]//International Conference on Information Computing and Applications. Berlin: Springer, 2012: 246-252.
- [58] CHEN T, GUESTRIN C. XGBoost: a scalable tree boosting system[C]//KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York, NY: Association for Computing Machinery, 2016: 785-794.
- [59] MACQUEEN J. Some methods for classification and analysis of multivariate observations [C]//Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability. 1965: 281-297.
- [60] MCCULLOCH W S, PITTS W. A logical calculus of the ideas immanent in nervous activity [J]. The Bulletin of Mathematical Biophysics, 1943, 5(4): 115-133.
- [61] ROSENBLATT F. The perceptron: a probabilistic model for information storage and organization in the brain.[J]. Psychological Review, 1958, 65(6): 386.
- [62] RUMELHART D E, HINTON G E, WILLIAMS R J. Learning representations by back-propagating errors[J]. Nature, 1986, 323(6088): 533-536.
- [63] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[C]//Advances in Neural Information Processing Systems (NeurIPS): Vol. 25. Red Hook, NY: Curran Associates, Inc, 2012.
- [64] SHERSTINSKY A. Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network[J]. Physica D: Nonlinear Phenomena, 2020, 404: 132306.
- [65] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
- [66] GREFF K, SRIVASTAVA R K, KOUTNÍK J, et al. LSTM: a search space odyssey[J]. IEEE Transactions on Neural Networks and Learning Systems, 2016, 28(10): 2222-2232.
- [67] WU H, XU J, WANG J, et al. Autoformer: decomposition transformers with auto-correlation for long-term series forecasting[C]//Advances in Neural Information Processing Systems (NeurIPS): Vol. 34. Red Hook, NY: Curran Associates, Inc, 2021: 22419-22430.
- [68] TANG S, WIENS J. Model selection for offline reinforcement learning: practical considerations for healthcare settings[C]//Proceedings of Machine Learning Research: Vol. 149 Proceedings of the 6th Machine Learning for Healthcare Conference. New York: PMLR, 2021: 2-35.
- [69] LE H, VOLOSHIN C, YUE Y. Batch policy learning under constraints[C]//Proceedings of

- Machine Learning Research: Vol. 97 Proceedings of the 36th International Conference on Machine Learning. New York: PMLR, 2019: 3703-3712.
- [70] THOMAS P, BRUNSKILL E. Data-efficient off-policy policy evaluation for reinforcement learning[C]//Proceedings of Machine Learning Research: Vol. 48 Proceedings of The 33rd International Conference on Machine Learning. New York: PMLR, 2016: 2139-2148.
- [71] JIANG N, LI L. Doubly robust off-policy value evaluation for reinforcement learning[C]//Proceedings of Machine Learning Research: Proceedings of the 33rd International Conference on Machine Learning. New York: PMLR, 2016: 652-661.
- [72] FARAJTABAR M, CHOW Y, GHAVAMZADEH M. More robust doubly robust off-policy evaluation[C]//Proceedings of Machine Learning Research: Proceedings of the 35th International Conference on Machine Learning. New York: PMLR, 2018: 1447-1456.
- [73] COVER T, HART P. Nearest neighbor pattern classification[J]. IEEE Transactions on Information Theory, 1967, 13(1): 21-27.
- [74] 雍伟哲. 急性呼吸窘迫综合征患者机械通气指南 (试行)[J]. 中华医学信息导报, 2016(4): 5-5.
- [75] KE G, MENG Q, FINLEY T, et al. Lightgbm: a highly efficient gradient boosting decision tree[C]//Advances in Neural Information Processing Systems (NeurIPS): Vol. 30. Red Hook, NY: Curran Associates, Inc, 2017.
- [76] OSTROUMOVA L, GUSEV G, VOROBEV A, et al. CatBoost: unbiased boosting with categorical features[C]//Advances in Neural Information Processing Systems (NeurIPS). Red Hook, NY: Curran Associates, Inc, 2017.
- [77] LAI G, CHANG W C, YANG Y, et al. Modeling long-and short-term temporal patterns with deep neural networks[C]//SIGIR '18: Proceedings of the 41st International ACM SIGIR Conference on Research and Development in Information Retrieval. New York, NY: Association for Computing Machinery, 2018: 95-104.
- [78] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Advances in Neural Information Processing Systems (NeurIPS): Vol. 30. Red Hook, NY: Curran Associates, Inc, 2017: 5998-6008.
- [79] WU H, HU T, LIU Y, et al. TimesNet: temporal 2D-variation modeling for general time series analysis[C]//International Conference on Learning Representations. Washington, D.C.: ICLR, 2023.
- [80] EKAMBARAM V, JATI A, NGUYEN N, et al. Tsmixer: lightweight mlp-mixer model for multivariate time series forecasting[C]//KDD '23: Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. New York, NY: Association for Computing Machinery, 2023: 459-469.

- [81] SINAGA K P, YANG M S. Unsupervised K-means clustering algorithm[J]. IEEE Access, 2020, 8: 80716-80727.
- [82] 杨惟轶, 白辰甲, 蔡超, 赵英男, 刘鹏. 深度强化学习中稀疏奖励问题研究综述[J]. 计算机科学, 2020, 47(3): 182-191.
- [83] KNAUS W A, DRAPER E A, WAGNER D P, et al. APACHE II: a severity of disease classification system[J]. Critical Care Medicine, 1985, 13(10): 818-829.
- [84] ROUSSEEUW P J. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis[J]. Journal of Computational and Applied Mathematics, 1987, 20: 53-65.
- [85] NIEMAN G F, SATALIN J, ANDREWS P, et al. Personalizing mechanical ventilation according to physiologic parameters to stabilize alveoli and minimize ventilator induced lung injury (VILI)[J]. Intensive Care Medicine Experimental, 2017, 5(1): 8.
- [86] ZHOU J, LIN Z, DENG X, et al. Optimal positive end expiratory pressure levels in ventilated patients without acute respiratory distress syndrome: a bayesian network meta-analysis and systematic review of randomized controlled trials[J]. Frontiers in Medicine, 2021, 8: 730018.
- [87] KILICKAYA O, GAJIC O. Initial ventilator settings for critically ill patients[J]. Critical Care, 2013, 17(2): 123.

## 致谢

我经常感叹，时间好快，三年的研究生生涯就此落下帷幕。在一个很普通的和往常一样过着寝室-食堂-实验室三点一线的日子，我写到了这里，心中思绪万千，有太多要感谢的人。

当一个故事走到终点，往往会想起最初的时候。首先要感谢我的研究生导师赵云波老师，还记得第一次见面是在 22 年 5 月您约的一个腾讯会议，那是夏令营推荐前期的一个相互了解的过程，整个会议聊得很开心，结束后和旁边的朋友报喜，感觉保研中科大十拿九稳。正式进入科大读研后，正是有您的指导和帮助，我成功发了一篇还算不错的期刊，包括后期学位论文的开题报告及终稿的把关，都离不开您的悉心教导。

其次感谢我的父母，感谢你们的支持和鼓励，我被一步步托举着走到了今天。读研期间有无数个崩溃的时刻，你们总是能在我无助伤心时承接我的情绪，对我耐心开导，为我兜底。

接着感谢给予过我帮助的师兄师姐们，感谢我的朋友们，在科大的三年，正是因为有你们的陪伴，日子才不显得那么孤单。

感谢所有遇见的人和事，不管是好的还是坏的，我要么得到收获，要么学到教训。

最后感谢自己，感谢你一次次克服困难，走到了这里。

谨以此献给我的研究生三年生涯，献给自己一往无前的勇气。未来也请勇敢自信地向前走下去。

## 在读期间取得的科研成果

### 已发表论文:

- [1] **Jiaying Xi**, Shaojie Dong, Haonquan Zhou, Yunbo Zhao. K-CQL: An Arterial Blood Gas Analysis-based Deep Offline Reinforcement Learning Algorithm for Mechanical Ventilation Treatment[J]. Health Information Science and Systems, 2025, 13(1):48. (对应正文第五章)

### 已接收论文:

- [1] **Jiaying Xi**, Yunbo Zhao, Haoquan Zhou, Siyuan Guo and Chenwei Xu. RLMV: An Algorithm for Predicting the Duration of Mechanical Ventilation in Patients[C]//2026 15th Asian Control Conference (ASCC). IEEE, 2026. (对应正文第四章)

### 已公开专利:

- [1] **席嘉滢**, 汤敏, 周浩泉, 赵云波. 基于机器学习的呼吸机上机时间预测方法、系统、电子设备及存储介质 [P]. 安徽省: CN202510195051.2, 2025-06-10.

### 参与的科研项目:

- [1] 基于强化学习的呼吸机参数智能精准控制技术研究, 合肥综合性国家科学中心先导医学与前沿技术项目, 2023-2027.