

Beijing Forest Studio
北京理工大学信息系统及安全对抗实验中心



不规则多元时间序列预测研究

博士研究生 李新帅

2024年02月25日

- **总结反思**
 - PPT部分内容制作不严谨
 - 讲解部分内容过于基础，应详略得当
 - 算法讲解部分应更多的关注改进算法和思路
- **相关内容**
 - 2023.06.11 李新帅：《基于Transformer的时间序列分析》
 - 2023.09.10 齐首华：《时序知识图谱推理方法》

- 预期收获
- 题目内涵解析
- 研究背景与意义
- 研究历史与现状
- 知识基础
- 算法原理
 - Raindrop
 - Warpformer
- 特点总结与工作展望
- 参考文献

- 预期收获
 - 1. 了解不规则多元时间序列预测任务的基本概念
 - 2. 理解图神经网络技术在序列预测任务中的基本应用
 - 3. 理解序列内、序列间关系在时间序列任务中的建模方法

- 研究目标

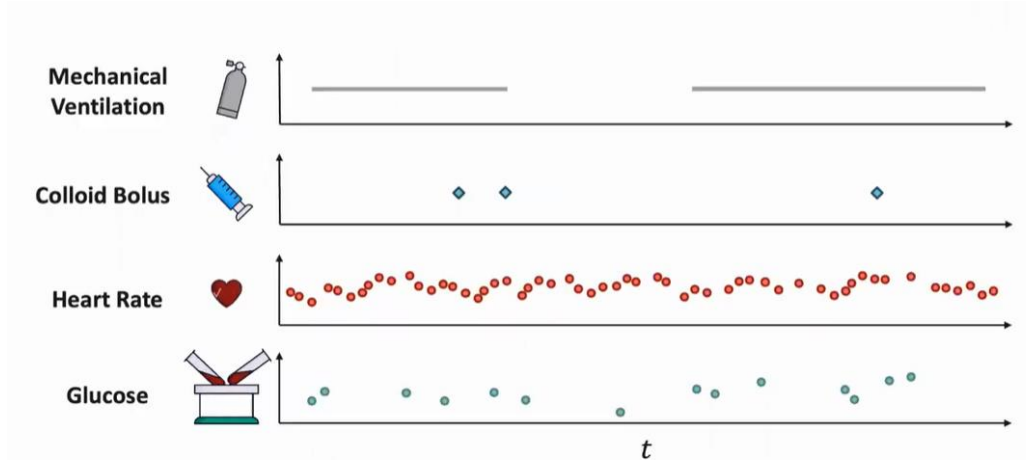
- 以**不规则多元时间序列**为研究对象，面向时间序列**趋势预测**任务
- 结合深度学习、网络嵌入、注意力机制、时间序列建模等技术
- 实现不规则多元时间序列**预测准确率**的显著提升

- 内涵解析

- **多元时间序列**：包含多个一元时间序列，在不同时间点观测到的**多个变量**之间存在**关联关系**的数据序列
- **不规则多元时间序列**：包含多个不规则的一元时间序列，序列内各特征数据的记录时间戳**不规则间隔**，且同一记录时间戳下可能仅记录**部分维度**的特征数据
- **时间序列预测**：利用获得的数据按时间顺序排成序列，分析其**变化方向和程度**，从而对未来若干时期可能达到的水平进行推测。

- 研究背景

- 广泛应用于科学、医学、金融、工程和工业等多种领域
- 不规则多元时间序列存在**采样率不均**、**缺失率高**等问题使研究面临众多挑战



- 研究意义

- 通过分析历史序列数据的变化方向或程度，对未来若干时期可能达到的水平进行推测，是能源管理、经济以及疾病传播分析的**关键方法**
- 对不规则多元时间序列的研究更好地模拟和预测真实世界中不规律、不连续的数据，有助于提高预测的**准确性和可靠性**，具有重要的**实用价值**

David等人应用**LSTM算法**建模价格历史以及技术分析指标预测股价的未来趋势，使得模型具有长期记忆的能力，有效捕获长距离信息。

冯等人提出的RSR模型是第一个考虑将关系图建模应用于时间预测领域的。它通过维基百科等知识库获取序列之间的关系，构建**图拓扑结构**。

周等人提出使用稀疏自注意力机制来降低时间复杂度和内存占用。它可以有效地处理**超长输入序列**，并且更适合长序列预测任务。

张提出了RAINDROP模型，其指定了一种消息传递策略，沿**传感器图**的边缘交换神经信息，并处理传感器读数不对齐的情况。



2017

2019

2021

2022

2018

2020

2021

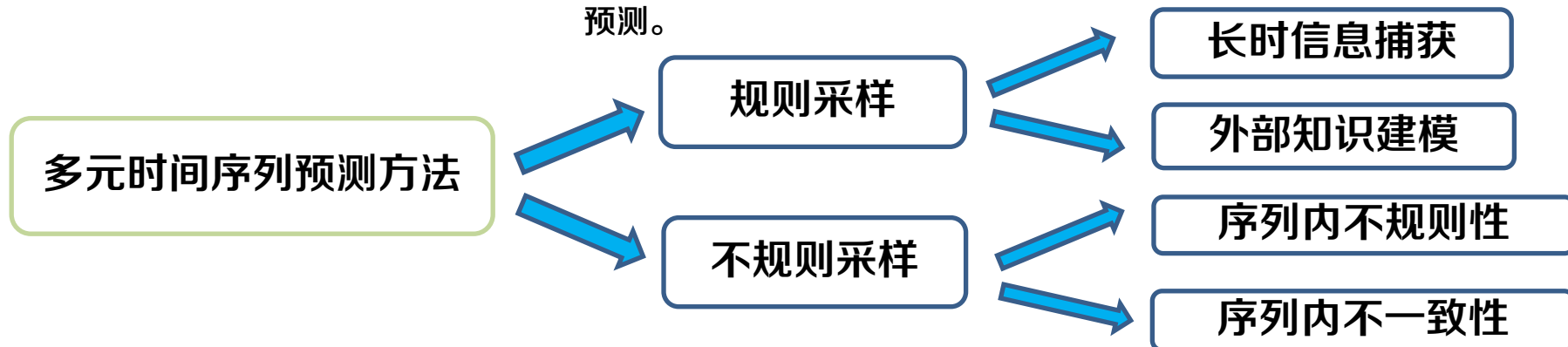
2023

Rubanova等人把RNNS推广到由**常微分方程(ODES)**定义的连续时间隐式动力学，这是一个称之为ODE-RNNS的模型，有效解决不规则时间序列的预测问题。

Ramit Sawhney等人提出了MAN-SF模型，该模型建立在RSR模型的基础上，考虑将社会新闻作为影响股票走势的**外部知识**进行预测。

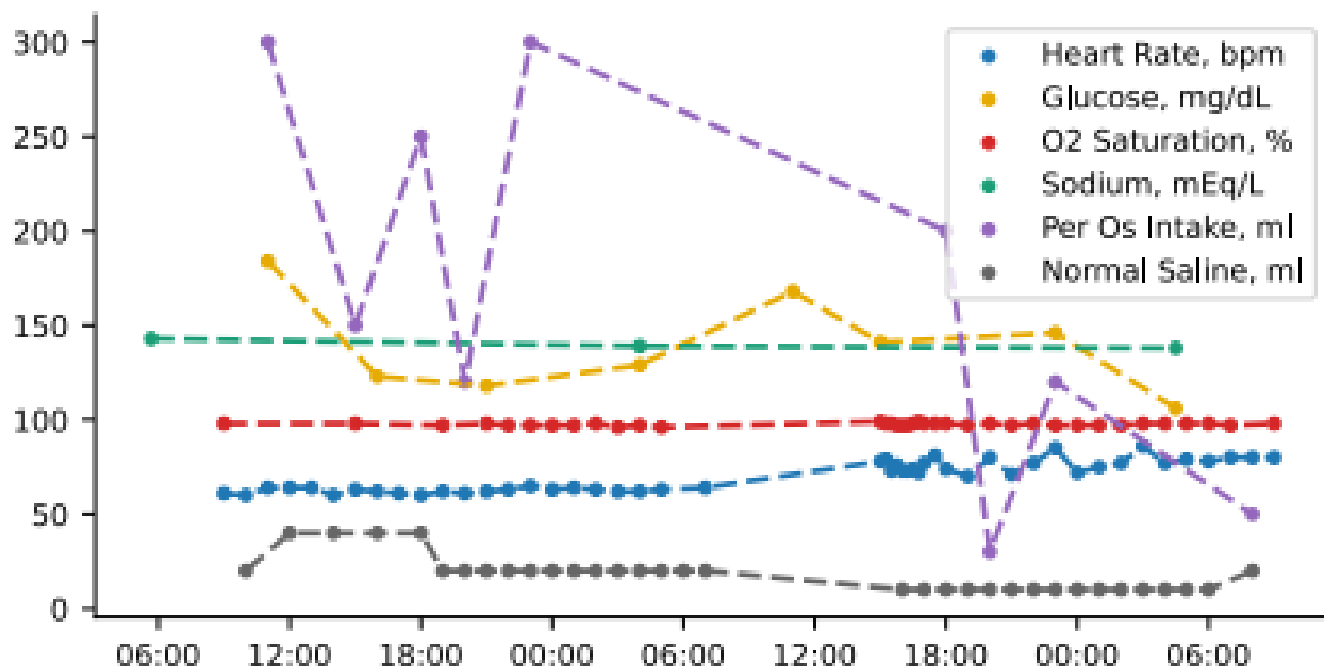
Shukla提出利用多时间注意力网络学习连续时间值的嵌入，产生包含**可变量**观测值的时间序列。

张等人充分考虑**序列内的不规则性和序列间的不一致性**。可以将不同的不规则采样序列按照特定的粒度自适应地统一到标准位置。



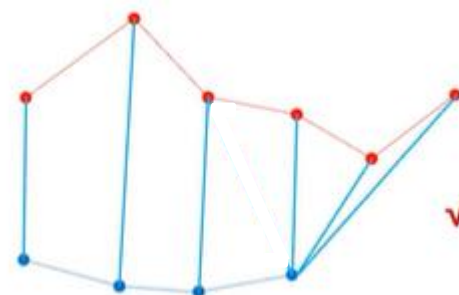
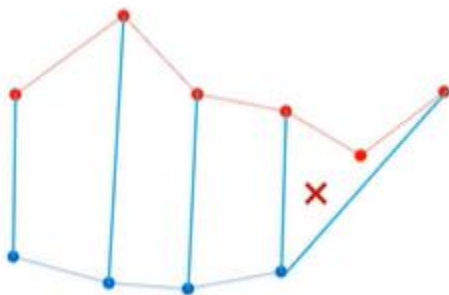
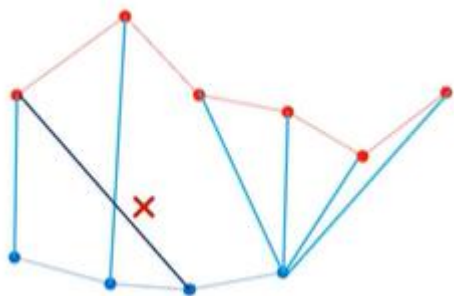
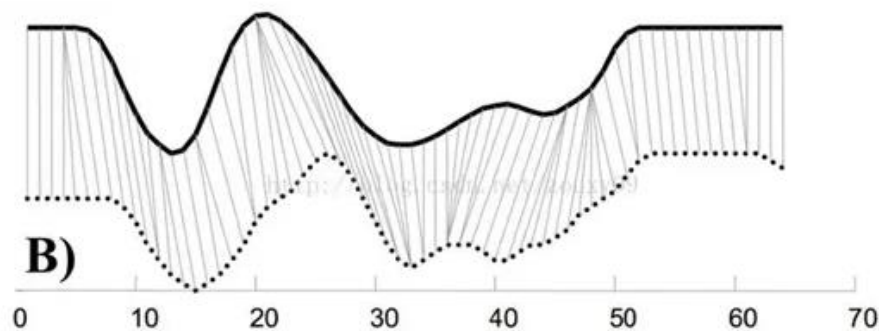
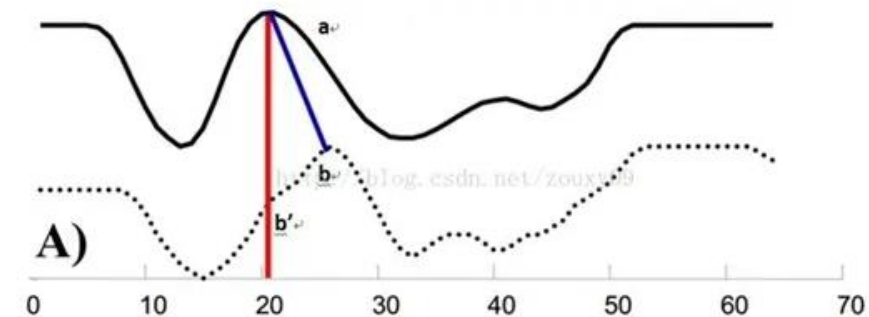
- 不规则多元时间序列预测方法

- 序列内的**不规则性**: 对单个不规则序列进行填充、预测等操作时, 对序列内的不规则性统一建模, 而忽略对变量差异性的**适应性学习**;
- 序列间的**差异性**: 对时间共现的变量建模, 针对**孤立时间点**无法聚合邻居影响, 导致图建模方法失效。



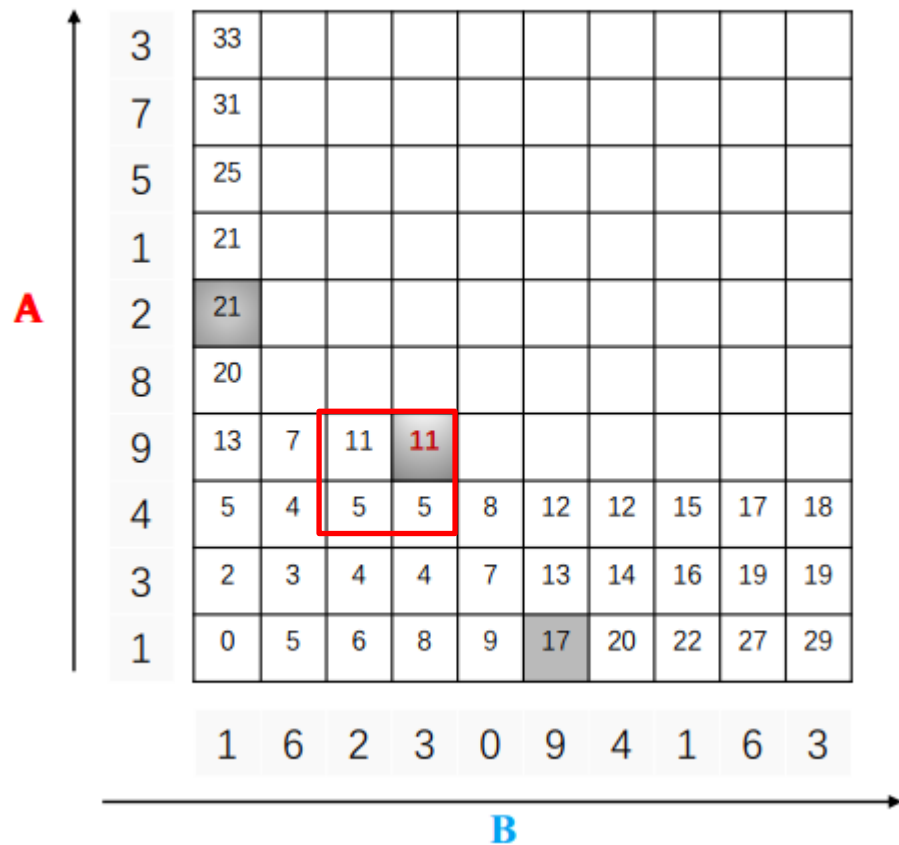
动态时间规整

- **Dynamic Time Warping, DTW**
 - 解决传统的**欧几里得距离**无法有效地求的两个时间序列之间的**距离**
 - 在时间轴上进行局部放缩
- **要求**
 - 单向对应，不能回头
 - 逐项对应，不能有空
 - 对应之后，距离最近

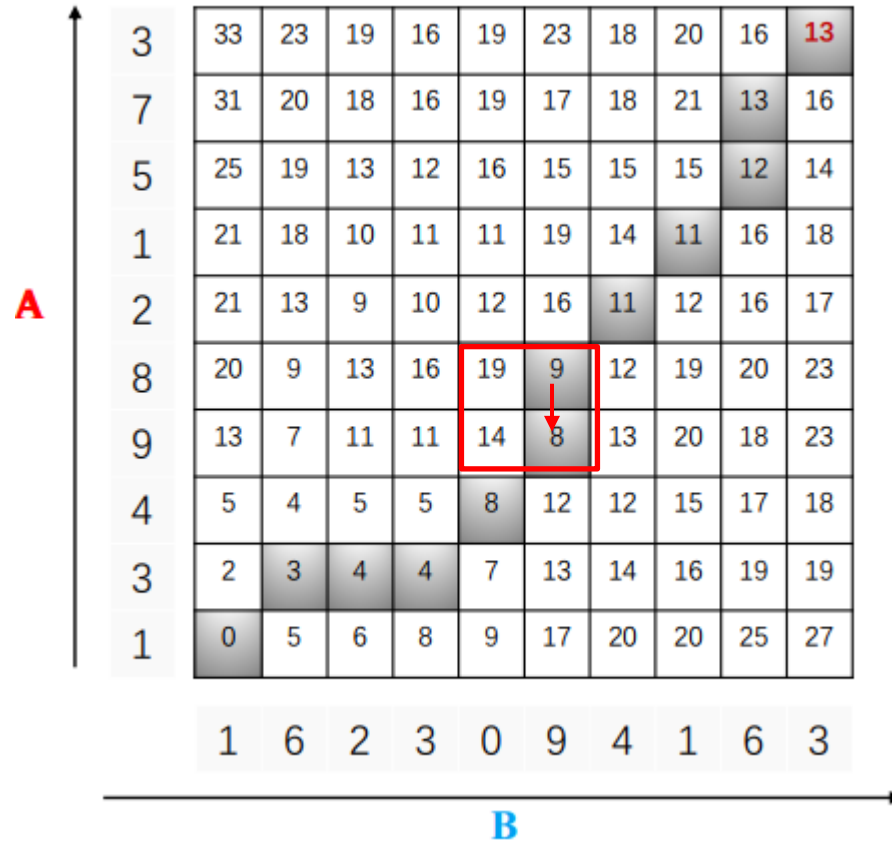


- 寻找最优路径

累计距离矩阵(D)



累计距离矩阵(D)





Graph-Guided Network for Irregularly Sampled Multivariate Time Series

T	目标	处理不规则的观测值、变化的时间间隔、任意数量的观测值，实现趋势分类
I	输入	P19（34维特征）、P12（36维特征）、PAM（17维特征）数据集
P	处理	1.传感器依赖图构建 2.图神经网络聚合目标传感器嵌入 3.使用时间注意力权重聚合所有传感器嵌入 4.训练下游分类器任务
O	输出	患病可能性、活动类型（2分类、8分类）

P	问题	少数研究考虑不规则多元时间序列的关系结构，而那些研究在捕获传感器间连接方面的能力有限
C	条件	不规则多元时间序列数据
D	难点	计算注意力权重时通过共享参数来利用不同样本之间的相似性 通过时间注意力考虑序列传感器观测的重要性
L	水平	ICLR 2022 CCF A

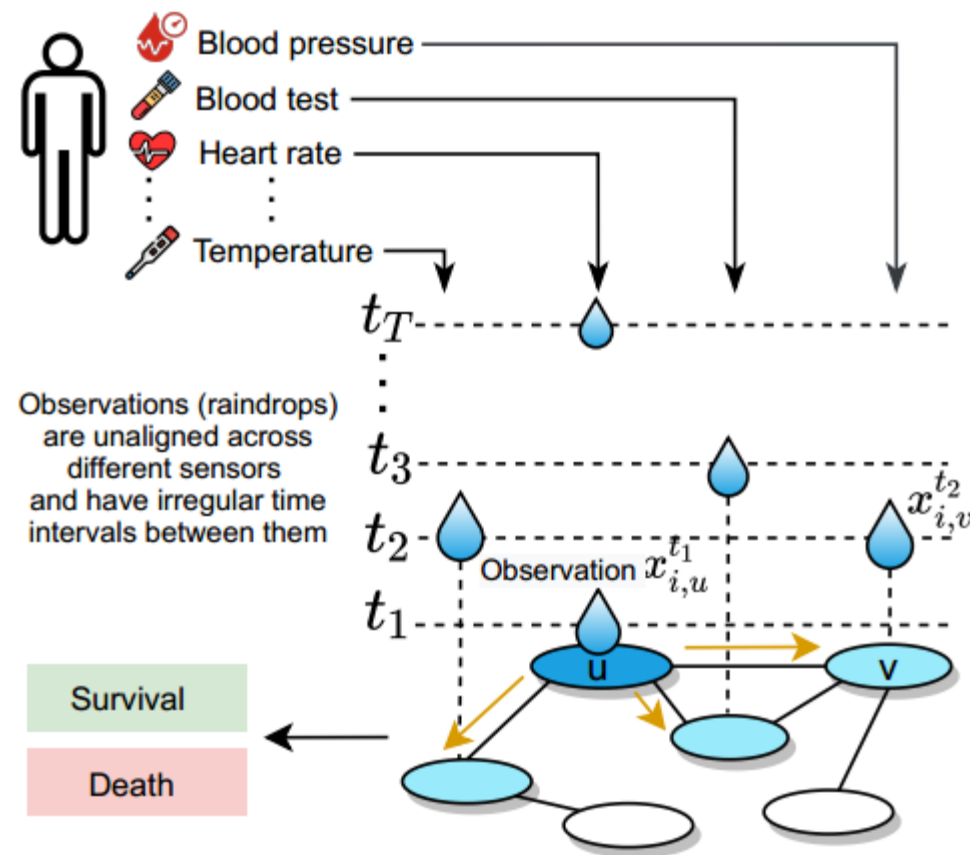
传感器依赖图

- 建立有向加权图 $G_i = (\mathcal{V}, \mathcal{E}_i)$ 对于每个样本 S_i
- 将从 u 到 v 的边表示为三元组 $(u; e_{i,uv}; v)$
- 其中 $e_{i,uv} \in [0,1]$ 表示样本 S_i 中传感器 u 和 v 之间的关系强度。

所有图初始化为全连通图

- 即对于任意 u, v 和 S_i , $e_{i,uv} = 1$

$$e_{i,uv}^{(l)} = \frac{e_{i,uv}^{(l-1)}}{|\mathcal{J}_{i,u}|} \sum_{t \in \mathcal{J}_{i,u}} \alpha_{i,uv}^{(l),t}$$



- 初始化嵌入表达

- 对于样本 \mathcal{S}_i 中的传感器 u 在 t 时刻的初始化嵌入表示: $x_{i,u}^t$

- 高维线性映射

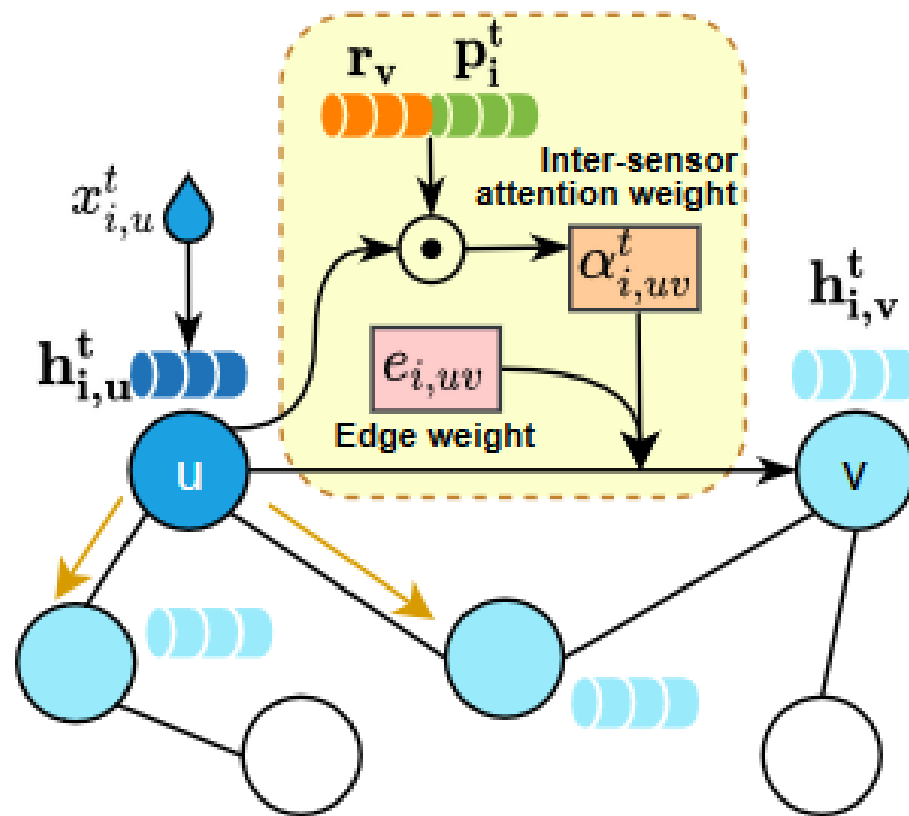
$$h_{i,u}^t = \sigma(x_{i,u}^t, \mathcal{R}_u)$$

- 传递消息

$$\alpha_{i,uv}^t = \sigma(h_{i,u}^t D[r_v \parallel p_i^t]^T)$$

$$h_{i,v}^t = \sigma(h_{i,u}^t w_u w_v^T \alpha_{i,uv}^t e_{i,uv})$$

a Sample \mathcal{S}_i records the value $x_{i,u}^t$ of sensor u at time t



- 传感器时间权重计算（自注意力机制）

$$\beta_{i,v} = \text{softmax} \left(\frac{Q_{i,v} K_{i,v}^T}{\sqrt{d_k}} s \right)$$

$$Q_{i,v} = H_{i,v} W_Q$$

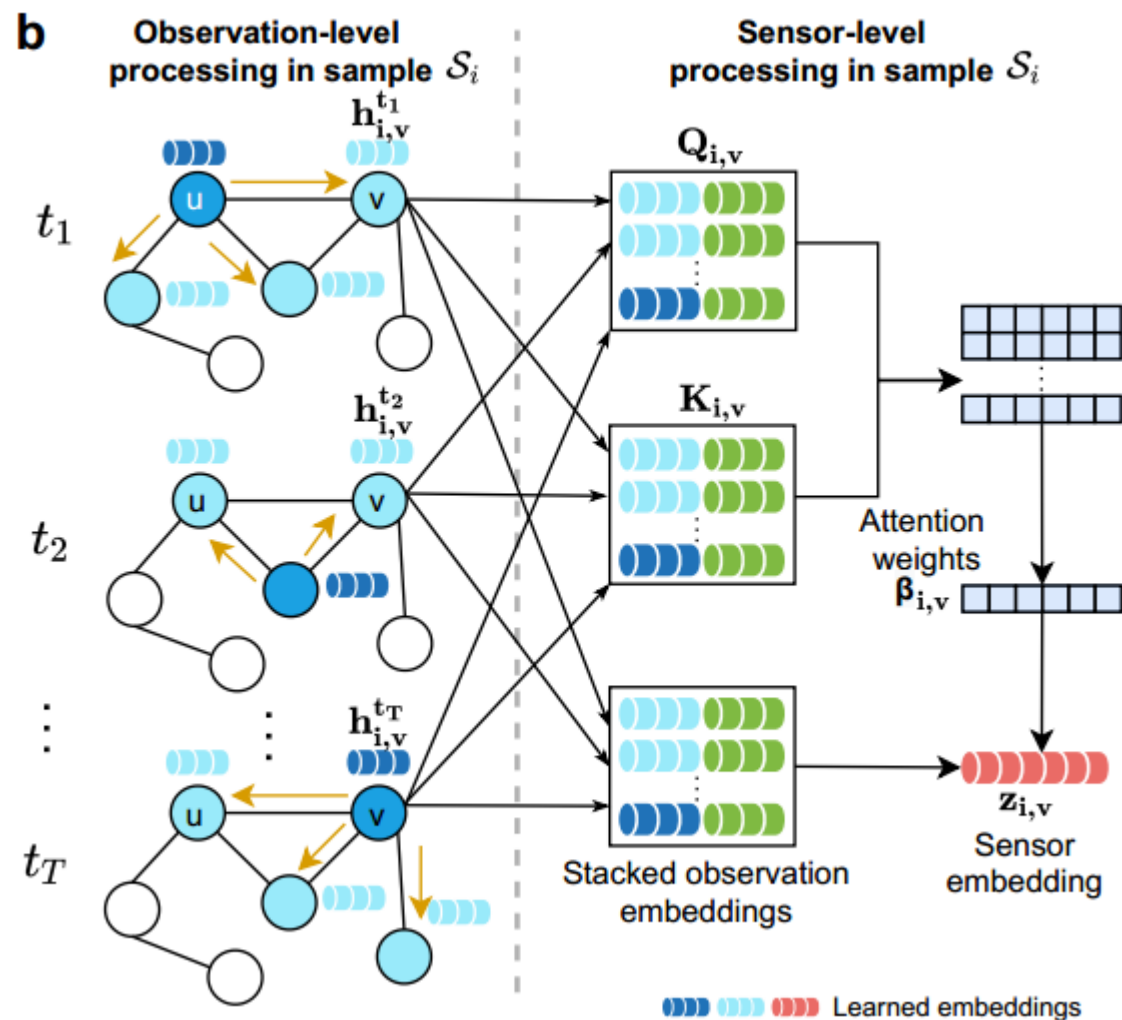
$$K_{i,v} = H_{i,v} W_K$$

$$H_{i,v} = h_{i,v}^t \parallel p_i^t$$

- 生成传感器嵌入

$$z_{i,v} = \sum_{t \in \mathcal{T}_{i,v}} (\beta_{i,v}^t [h_{i,v}^t \parallel p_i^t] W)$$

$$z_i = g(z_{i,v} | v = 1, 2, \dots, M)$$



- 数据集：
 - P19: PhysioNet Sepsis Early Prediction challenge 2019。包含38803例患者，每个患者由34个不规则采样传感器监测，包括8个生命体征和26个实验室值。每个样本都有一个包含年龄、性别等6个元素的静态向量。数据集高度不平衡，只有约4%的阳性样本。
 - P12: PhysioNet Mortality Prediction Challenge 2012。包括11988名患者（样本），每个患者包含36个传感器（不包括体重）的多元时间序列，收集于ICU前48小时，每个样本都有一个包含年龄、性别等9个元素的静态向量。数据集高度不平衡，有约93%的阳性样本。
 - PAM: PAMAP2 Physical Activity Monitoring。用3个惯性测量单位测量9名受试者的日常生活活动。包含5333个感觉信号片段（样本），每个样本由17个传感器测量，包含600个连续观测值。（多分类）

• 对比方法

- Transformer、Trans-mean (平均观测值代替缺失值)
- GRU-D
- SeFT
- IP-Net、mTAND
- DGM²-O、MTGNN

• 评价指标

- **AUROC**:关注的是模型的分类能力，模型的真正例率(True Positive Rate, 召回率)与假正例率(False Positive Rate)之间的权衡。
- **AUPRC**:关注精度(Precision)和召回率(Recall)之间的权衡。它对于在**正类别样本较少**的情况下更具敏感性。
- Precision、Recall、F1、Accuracy (PAM数据集)

• 实验结果

- Raindrop在三个基准数据集上获得了最好的性能，这表明它具有很强的时间序列分类性能。

Methods	P19		P12		PAM			
	AUROC	AUPRC	AUROC	AUPRC	Accuracy	Precision	Recall	F1 score
Transformer	83.2 ± 1.3	47.6 ± 3.8	65.1 ± 5.6	95.7 ± 1.6	83.5 ± 1.5	84.8 ± 1.5	86.0 ± 1.2	85.0 ± 1.3
Trans-mean	84.1 ± 1.7	47.4 ± 1.4	66.8 ± 4.2	95.9 ± 1.1	83.7 ± 2.3	84.9 ± 2.6	86.4 ± 2.1	85.1 ± 2.4
GRU-D	83.9 ± 1.7	46.9 ± 2.1	67.2 ± 3.6	95.9 ± 2.1	83.3 ± 1.6	84.6 ± 1.2	85.2 ± 1.6	84.8 ± 1.2
SeFT	78.7 ± 2.4	31.1 ± 2.8	66.8 ± 0.8	96.2 ± 0.2	67.1 ± 2.2	70.0 ± 2.4	68.2 ± 1.5	68.5 ± 1.8
mTAND	80.4 ± 1.3	32.4 ± 1.8	65.3 ± 1.7	96.5 ± 1.2	74.6 ± 4.3	74.3 ± 4.0	79.5 ± 2.8	76.8 ± 3.4
IP-Net	84.6 ± 1.3	38.1 ± 3.7	72.5 ± 2.4	96.7 ± 0.3	74.3 ± 3.8	75.6 ± 2.1	77.9 ± 2.2	76.6 ± 2.8
DGM ² -O	86.7 ± 3.4	44.7 ± 11.7	71.2 ± 2.5	96.9 ± 0.4	82.4 ± 2.3	85.2 ± 1.2	83.9 ± 2.3	84.3 ± 1.8
MTGNN	81.9 ± 6.2	39.9 ± 8.9	67.5 ± 3.1	96.4 ± 0.7	83.4 ± 1.9	85.2 ± 1.7	86.1 ± 1.9	85.9 ± 2.4
RAINDROP	87.0 ± 2.3	51.8 ± 5.5	72.1 ± 1.3	97.0 ± 0.4	88.5 ± 1.5	89.9 ± 1.5	89.9 ± 0.6	89.8 ± 1.0

• 对比方法

- Transformer、Trans-mean (平均观测值代替缺失值)
- GRU-D
- SeFT
- mTAND

• 实验设置

- 删除概率：10%、20%、30%、40%、50%
- 数据集：PAM
- 评价指标：Precision、Recall、F1、Accuracy

实验结果

- 当遗漏的传感器数量从10%到50%时，在20个设置中的18个设置中，Raindrop达到了最佳性能。
- 随着丢失数据量的增加，Raindrop产生了更大的性能改进。

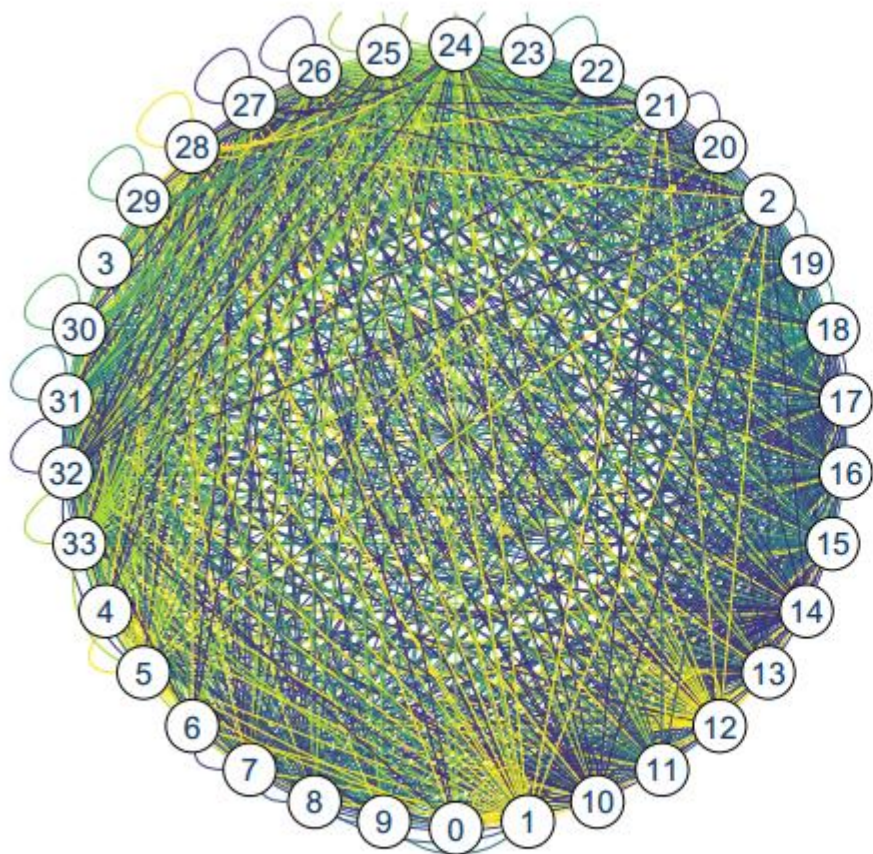
Missing sensor ratio	Methods	PAM (Setting 2: leave-fixed-sensors-out)			
		Accuracy	Precision	Recall	F1 score
10%	Transformer	60.3 ± 2.4	57.8 ± 9.3	59.8 ± 5.4	57.2 ± 8.0
	Trans-mean	60.4 ± 11.2	61.8 ± 14.9	60.2 ± 13.8	58.0 ± 15.2
	GRU-D	65.4 ± 1.7	72.6 ± 2.6	64.3 ± 5.3	63.6 ± 0.4
	SeFT	58.9 ± 2.3	62.5 ± 1.8	59.6 ± 2.6	59.6 ± 2.6
	mTAND	58.8 ± 2.7	59.5 ± 5.3	64.4 ± 2.9	61.8 ± 4.1
	RAINDROP	77.2 ± 2.1	82.3 ± 1.1	78.4 ± 1.9	75.2 ± 3.1
20%	Transformer	63.1 ± 7.6	71.1 ± 7.1	62.2 ± 8.2	63.2 ± 8.7
	Trans-mean	61.2 ± 3.0	74.2 ± 1.8	63.5 ± 4.4	64.1 ± 4.1
	GRU-D	64.6 ± 1.8	73.3 ± 3.6	63.5 ± 4.6	64.8 ± 3.6
	SeFT	35.7 ± 0.5	42.1 ± 4.8	38.1 ± 1.3	35.0 ± 2.2
	mTAND	33.2 ± 5.0	36.9 ± 3.7	37.7 ± 3.7	37.3 ± 3.4
	RAINDROP	66.5 ± 4.0	72.0 ± 3.9	67.9 ± 5.8	65.1 ± 7.0
30%	Transformer	31.6 ± 10.0	26.4 ± 9.7	24.0 ± 10.0	19.0 ± 12.8
	Trans-mean	42.5 ± 8.6	45.3 ± 9.6	37.0 ± 7.9	33.9 ± 8.2
	GRU-D	45.1 ± 2.9	51.7 ± 6.2	42.1 ± 6.6	47.2 ± 3.9
	SeFT	32.7 ± 2.3	27.9 ± 2.4	34.5 ± 3.0	28.0 ± 1.4
	mTAND	27.5 ± 4.5	31.2 ± 7.3	30.6 ± 4.0	30.8 ± 5.6
	RAINDROP	52.4 ± 2.8	60.9 ± 3.8	51.3 ± 7.1	48.4 ± 1.8
40%	Transformer	23.0 ± 3.5	7.4 ± 6.0	14.5 ± 2.6	6.9 ± 2.6
	Trans-mean	25.7 ± 2.5	9.1 ± 2.3	18.5 ± 1.4	9.9 ± 1.1
	GRU-D	46.4 ± 2.5	64.5 ± 6.8	42.6 ± 7.4	44.3 ± 7.9
	SeFT	26.3 ± 0.9	29.9 ± 4.5	27.3 ± 1.6	22.3 ± 1.9
	mTAND	19.4 ± 4.5	15.1 ± 4.4	20.2 ± 3.8	17.0 ± 3.4
	RAINDROP	52.5 ± 3.7	53.4 ± 5.6	48.6 ± 1.9	44.7 ± 3.4
50%	Transformer	21.4 ± 1.8	2.7 ± 0.2	12.5 ± 0.4	4.4 ± 0.3
	Trans-mean	21.3 ± 1.6	2.8 ± 0.4	12.5 ± 0.7	4.6 ± 0.2
	GRU-D	37.3 ± 2.7	29.6 ± 5.9	32.8 ± 4.6	26.6 ± 5.9
	SeFT	24.7 ± 1.7	15.9 ± 2.7	25.3 ± 2.6	18.2 ± 2.4
	mTAND	16.9 ± 3.1	12.6 ± 5.5	17.0 ± 1.6	13.9 ± 4.0
	RAINDROP	46.6 ± 2.6	44.5 ± 2.6	42.4 ± 3.9	38.0 ± 4.0

实验结果

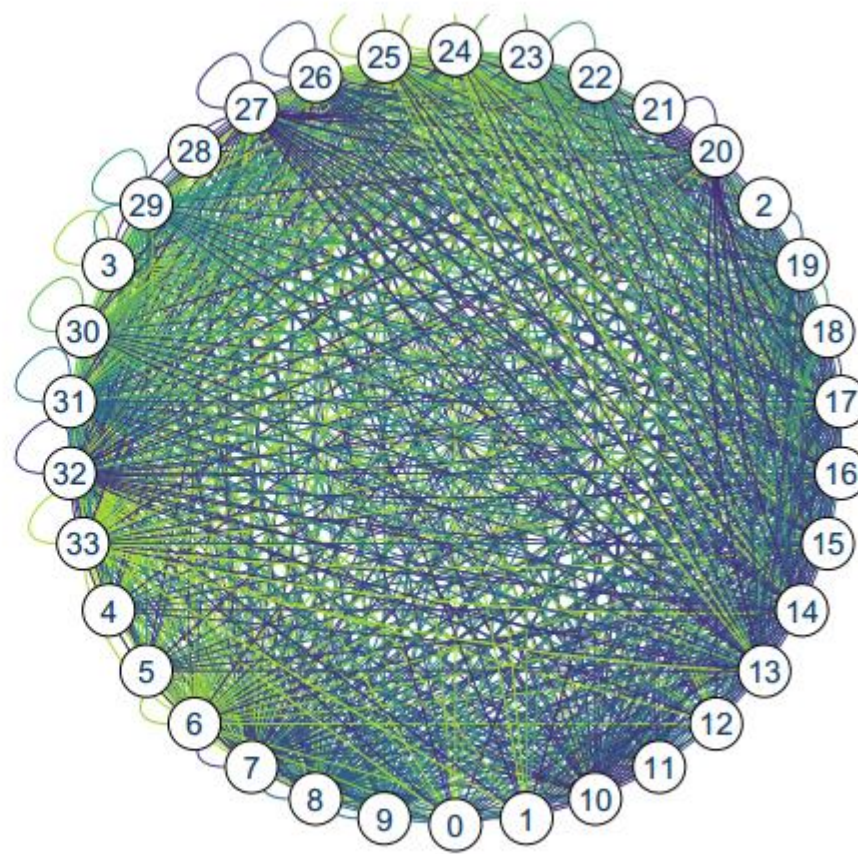
- 在20种设置中，Raindrop在16种设置中取得了比基线更好的性能。
- Trans-mean和GRU-D是最强的竞争对手。

Missing sensor ratio	Methods	PAM (Setting 3: leave-random-sensors-out)			
		Accuracy	Precision	Recall	F1 score
10%	Transformer	60.9 ± 12.8	58.4 ± 18.4	59.1 ± 16.2	56.9 ± 18.9
	Trans-mean	62.4 ± 3.5	59.6 ± 7.2	63.7 ± 8.1	62.7 ± 6.4
	GRU-D	68.4 ± 3.7	74.2 ± 3.0	70.8 ± 4.2	72.0 ± 3.7
	SeFT	40.0 ± 1.9	40.8 ± 3.2	41.0 ± 0.7	39.9 ± 1.5
	mTAND	53.4 ± 2.0	54.8 ± 2.7	57.0 ± 1.9	55.9 ± 2.2
	RAINDROP	76.7 ± 1.8	79.9 ± 1.7	77.9 ± 2.3	78.6 ± 1.8
20%	Transformer	62.3 ± 11.5	65.9 ± 12.7	61.4 ± 13.9	61.8 ± 15.6
	Trans-mean	56.8 ± 4.1	59.4 ± 3.4	53.2 ± 3.9	55.3 ± 3.5
	GRU-D	64.8 ± 0.4	69.8 ± 0.8	65.8 ± 0.5	67.2 ± 0.0
	SeFT	34.2 ± 2.8	34.9 ± 5.2	34.6 ± 2.1	33.3 ± 2.7
	mTAND	45.6 ± 1.6	49.2 ± 2.1	49.0 ± 1.6	49.0 ± 1.0
	RAINDROP	71.3 ± 2.5	75.8 ± 2.2	72.5 ± 2.0	73.4 ± 2.1
30%	Transformer	52.0 ± 11.9	55.2 ± 15.3	50.1 ± 13.3	48.4 ± 18.2
	Trans-mean	65.1 ± 1.9	63.8 ± 1.2	67.9 ± 1.8	64.9 ± 1.7
	GRU-D	58.0 ± 2.0	63.2 ± 1.7	58.2 ± 3.1	59.3 ± 3.5
	SeFT	31.7 ± 1.5	31.0 ± 2.7	32.0 ± 1.2	28.0 ± 1.6
	mTAND	34.7 ± 5.5	43.4 ± 4.0	36.3 ± 4.7	39.5 ± 4.4
	RAINDROP	60.3 ± 3.5	68.1 ± 3.1	60.3 ± 3.6	61.9 ± 3.9
40%	Transformer	43.8 ± 14.0	44.6 ± 23.0	40.5 ± 15.9	40.2 ± 20.1
	Trans-mean	48.7 ± 2.7	55.8 ± 2.6	54.2 ± 3.0	55.1 ± 2.9
	GRU-D	47.7 ± 1.4	63.4 ± 1.6	44.5 ± 0.5	47.5 ± 0.0
	SeFT	26.8 ± 2.6	24.1 ± 3.4	28.0 ± 1.2	23.3 ± 3.0
	mTAND	23.7 ± 1.0	33.9 ± 6.5	26.4 ± 1.6	29.3 ± 1.9
	RAINDROP	57.0 ± 3.1	65.4 ± 2.7	56.7 ± 3.1	58.9 ± 2.5
50%	Transformer	43.2 ± 2.5	52.0 ± 2.5	36.9 ± 3.1	41.9 ± 3.2
	Trans-mean	46.4 ± 1.4	59.1 ± 3.2	43.1 ± 2.2	46.5 ± 3.1
	GRU-D	49.7 ± 1.2	52.4 ± 0.3	42.5 ± 1.7	47.5 ± 1.2
	SeFT	26.4 ± 1.4	23.0 ± 2.9	27.5 ± 0.4	23.5 ± 1.8
	mTAND	20.9 ± 3.1	35.1 ± 6.1	23.0 ± 3.2	27.7 ± 3.9
	RAINDROP	47.2 ± 4.4	59.4 ± 3.9	44.8 ± 5.3	47.6 ± 5.2

传感器间依赖关系图的可视化



(a) Negative samples



(b) Positive samples



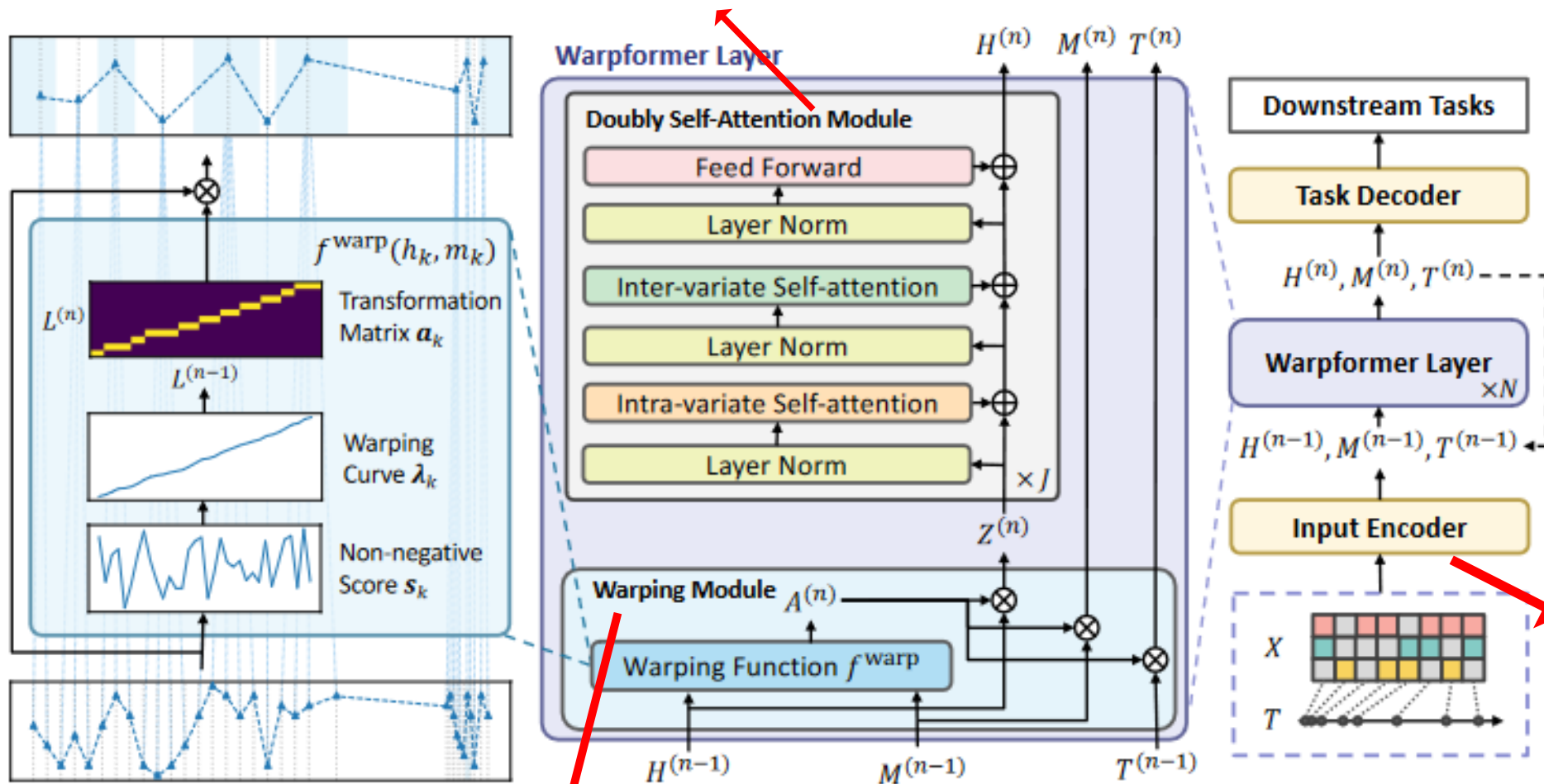
Warpformer: A Multi-scale Modeling Approach for Irregular Clinical Time Series

T	目标	处理不规则观测值、变化时间间隔、任意数量的观测值，实现趋势预测
I	输入	PhysioNet dataset (41维特征)、Human Activity dataset (12维特征) MIMIC -III (61维特征)
P	处理	1.输入编码器训练 2.给定的尺度下统一输入表示 3.双重自注意力聚合 4.训练下游分类器任务
O	输出	患病可能性、活动类型 (2分类、7分类)
P	问题	序列内的不规则性：时间序列信号经常以不规则的间隔被记录 序列间的差异：不同序列之间采样率的显著差异
C	条件	不规则多元时间序列数据
D	难点	如何对不规则多元时间序列的粒度进行有效归一化
L	水平	KDD 2023 CCF A



创新说明

序列内、序列间注意力机制



自适应统一维度表示

获取丰富的和结构化的数据表示

- 序列定义表达

$$\left\{ \left[(t_i^k, x_i^k) \right]_i^{L^k} \right\}_{k=1}^K$$

- 输入矩阵

- 数值矩阵 $X \in \mathbb{R}^{K \times L}$ $X_{k,j} = x_i^k$ or 0

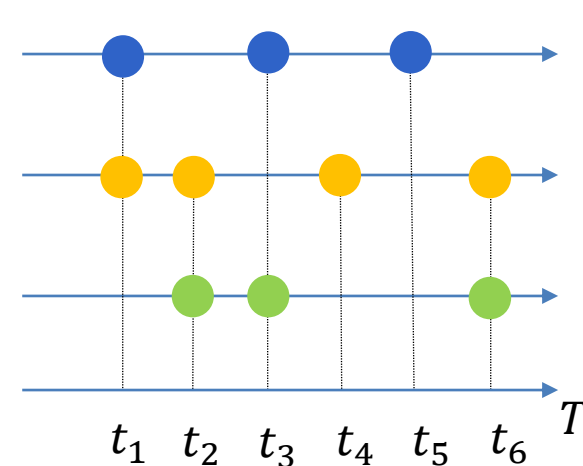
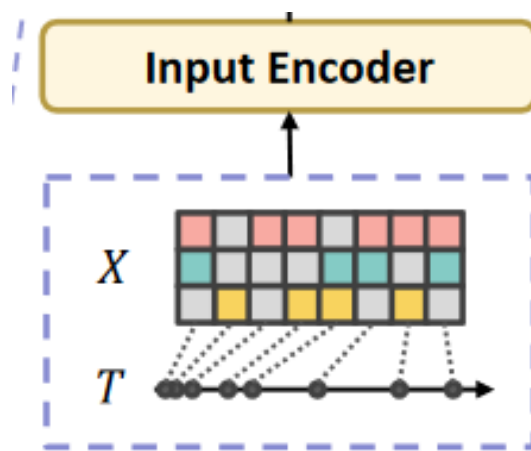
- 类型矩阵 $E \in \mathbb{R}^{K \times L}$ $E_{k,j} = k$ or 0

- 掩码矩阵 $M \in \mathbb{R}^{K \times L}$ $M_{k,j} = 1$ or 0

- 输入编码

$$H = f^{val}(X) + f^{type}(E) + f^{abs}(T) + f^{rel}(T, M)$$

↓ ↓
 绝对时间编码 相对时间编码



$$X = \begin{bmatrix} 0.1 & 0 & 0.2 & 0 & 0.3 & 0 \\ 1.1 & 1 & 0 & 1.2 & 0 & 1.3 \\ 0 & 2.3 & 2.1 & 0 & 0 & 3.1 \end{bmatrix}$$

$$E = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 2 & 2 & 0 & 2 & 0 & 2 \\ 0 & 3 & 3 & 0 & 0 & 3 \end{bmatrix}$$

$$M = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

- Warping Module (规整模块)
 - 将多样的不规则采样序列按照特定的粒度自适应地统一到校准位置，以便进行进一步的表示学习
 - 支持下采样和上采样，以同时对齐粗粒度和细粒度信息
- 模块输入

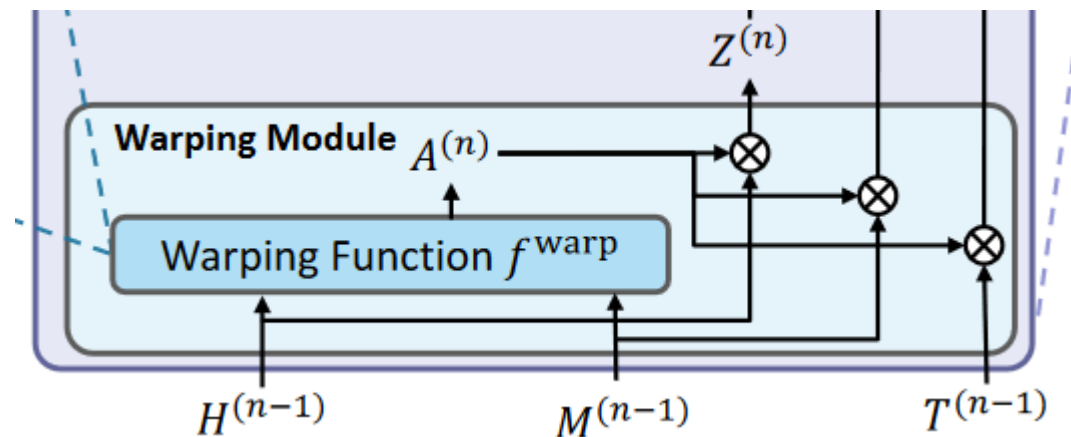
$$(H^{(n-1)}, M^{(n-1)}, T^{(n-1)}) \quad n \in \{1, \dots, N\}$$

$$A^{(n)} = f^{\text{warp}}(H^{(n-1)}, M^{(n-1)}) \quad \longrightarrow \quad \text{上、下采样；实现维度变换}$$

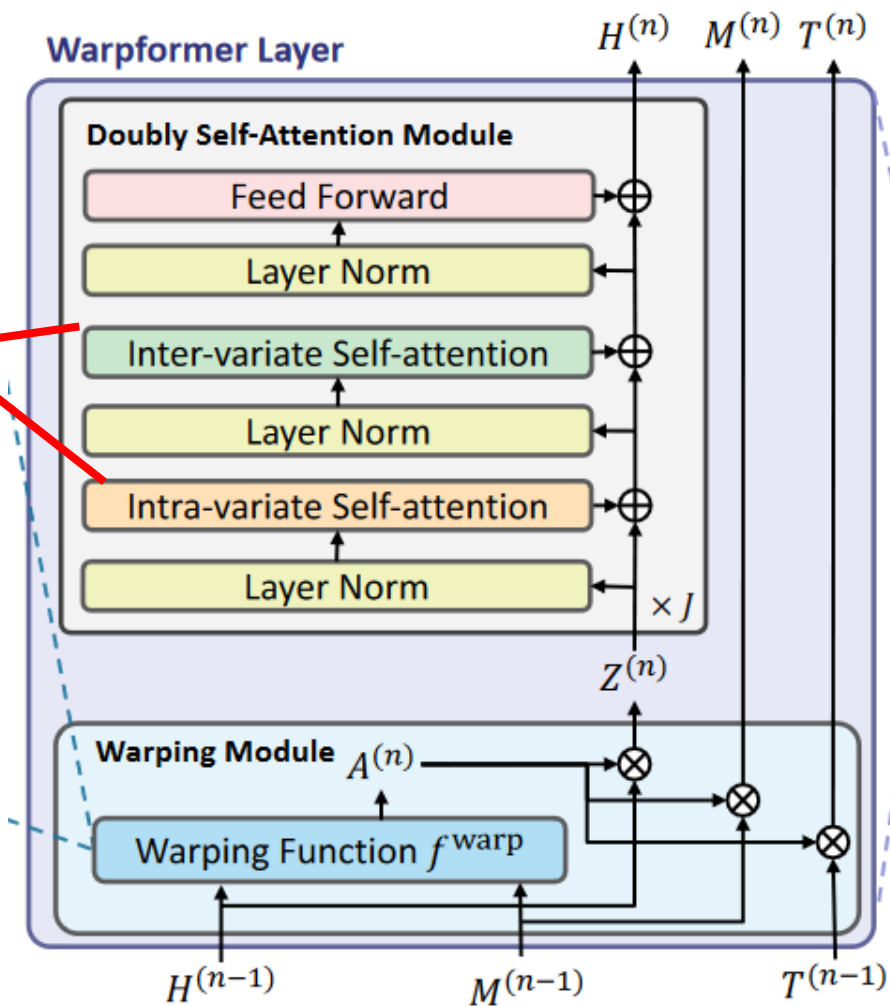
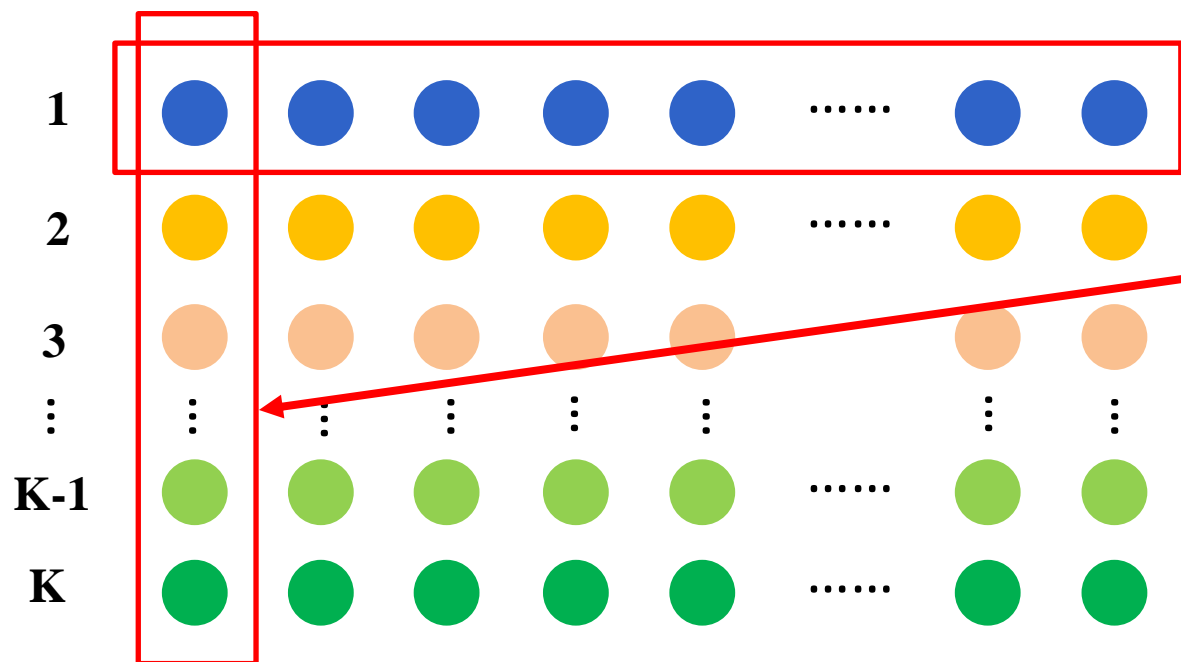
$$Z^{(n)} = A^{(n)} \otimes H^{(n-1)}$$

$$M^{(n)} = A^{(n)} \otimes M^{(n-1)}$$

$$T^{(n)} = A^{(n)} \otimes T^{(n-1)}$$



- Doubly Self-attention Module (双重自注意力模块)
 - 提出了一种新颖的**双重自注意力**机制，作为Transformer编码器的定制扩展
 - 序列内、序列间相关性建模



- 数据集

- **PhysioNet dataset**: 包含4000个实例，侧重于预测住院死亡率。包括4类人口学信息和入住ICU最初48小时内采集的37个生理信号。该数据集中实例长度的中位数为72。
- **Human Activity dataset**: 它由6554个时间序列段组成，共12个通道。该数据集中的所有实例都有固定长度为50个时间点。对特定人类活动进行7种类型的分类。
- **MIMIC -III**: 包含了53423名患者的入院记录。选择61个常用的生物标志物变量和42个在重症监护病房中广泛使用的干预措施。该数据集中实例长度的中位数为34。
 - MOR 临床死亡率-2分类，48小时
 - DEC 失代偿期- 2分类，12小时
 - LOS 住院时间-9分类
 - WBM Next时机-多标签分类，54个生物标志物下一刻活动
 - CIP 临床干预预测-4分类，6小时

• 对比方法

- RNN-Mean, RNN-Forward, RNN- Δt , RNN-Decay, GRU-D, Phased-LSTM
- ODE-RNN, L-ODE-ODE
- IP-Net, mTAND
- SeFT, RainDrop

• 评价指标

- **AUROC**:关注的是模型的分类能力，模型的真正例率(True Positive Rate, 召回率)与假正例率(False Positive Rate)之间的权衡。
- **AUPRC**:关注精度(Precision)和召回率(Recall)之间的权衡。它对于在**正类别样本较少**或非常重要的情况下更具敏感性。

实验结果

- 优于其他的基线算法
- RNN-Mean, 在PhysioNet数据集上表现差; 在基于MIMIC-III的数据集上表现好

MIMIC -III数据集

Model	PhysioNet		Human Activity	
	AUROC	AUPRC	Accuracy	AUPRC
RNN-Mean	55.5 ± 8.7	20.3 ± 7.5	74.9 ± 1.4	65.5 ± 2.4
RNN-Forward	84.2 ± 0.7	52.6 ± 0.7	76.7 ± 0.6	68.7 ± 0.6
RNN- Δ_t	65.7 ± 3.2	25.2 ± 1.5	74.6 ± 0.6	65.3 ± 2.0
RNN-Decay	49.9 ± 0.9	15.7 ± 0.4	75.9 ± 1.3	64.4 ± 1.5
GRU-D	84.7 ± 0.3	52.6 ± 0.3	75.0 ± 1.0	64.3 ± 1.7
Phased-LSTM	78.1 ± 0.5	40.7 ± 3.6	73.5 ± 1.7	60.8 ± 1.0
ODE-RNN	83.7 ± 0.6	53.3 ± 1.1	76.0 ± 1.2	64.9 ± 2.6
L-ODE-ODE	81.2 ± 1.3	46.6 ± 3.6	76.7 ± 1.4	77.1 ± 1.4
SeFT	70.8 ± 0.3	29.9 ± 0.9	75.8 ± 3.3	66.1 ± 3.3
RainDrop	69.2 ± 4.2	28.4 ± 3.6	70.2 ± 1.6	62.3 ± 2.6
IP-Nets	75.6 ± 1.1	37.5 ± 2.3	73.5 ± 1.7	64.7 ± 2.4
mTAND	86.0 ± 0.4	54.6 ± 0.9	81.3 ± 0.3	73.5 ± 0.8
Warpformer	86.6 ± 0.6	56.7 ± 0.7	84.9 ± 0.7	81.1 ± 0.9

Model	MOR		DEC		LOS	
	AUROC	AUPRC	AUROC	AUPRC	AUROC	AUPRC
RNN-Mean	89.5 ± 0.3	62.9 ± 0.7	98.4 ± 0.2	83.6 ± 0.8	76.7 ± 0.1	31.2 ± 0.1
RNN-Forward	88.5 ± 0.3	60.4 ± 1.2	97.1 ± 0.6	75.6 ± 2.2	76.4 ± 0.2	30.8 ± 0.1
RNN- Δ_t	87.4 ± 1.0	55.8 ± 2.4	97.3 ± 0.5	74.5 ± 0.8	76.2 ± 0.2	30.6 ± 0.2
RNN-Decay	89.1 ± 0.4	62.6 ± 1.0	98.4 ± 0.3	84.2 ± 2.4	77.0 ± 0.1	31.6 ± 0.1
GRU-D	87.6 ± 0.5	59.9 ± 0.6	97.6 ± 1.0	77.9 ± 5.6	76.5 ± 0.1	30.9 ± 0.3
Phased-LSTM	86.7 ± 0.3	53.7 ± 0.6	97.5 ± 0.1	78.2 ± 0.4	75.6 ± 0.2	30.2 ± 0.1
SeFT	88.0 ± 0.4	59.3 ± 1.4	98.5 ± 0.2	86.7 ± 0.7	76.1 ± 0.1	31.0 ± 0.1
RainDrop	87.6 ± 0.1	59.1 ± 0.4	98.0 ± 0.2	82.7 ± 0.5	76.0 ± 0.2	30.6 ± 0.4
IP-Nets	88.9 ± 0.3	61.9 ± 0.9	98.3 ± 0.1	85.2 ± 0.6	74.3 ± 3.6	28.8 ± 3.7
mTAND	89.0 ± 0.2	61.8 ± 0.7	97.2 ± 0.3	74.5 ± 3.2	73.8 ± 0.4	28.3 ± 0.4
AdaCare	76.0 ± 0.5	47.7 ± 0.7	94.9 ± 2.0	68.5 ± 9.0	64.9 ± 0.6	22.1 ± 1.2
STraTS	89.3 ± 0.1	61.3 ± 0.3	98.6 ± 0.1	84.0 ± 1.6	76.6 ± 0.2	31.4 ± 0.3
Warpformer	90.3 ± 0.1	64.6 ± 0.4	99.0 ± 0.1	90.0 ± 0.4	77.7 ± 0.2	32.5 ± 0.2

实验结果

- 每个任务的观察窗口长度不同，越长捕捉到的信息更丰富
- 任务本身难度的差异，导致改进效果不同

Model	MOR		DEC		LOS		WBM		CIP	
	AUROC	AUPRC	AUROC	AUPRC	AUROC	AUPRC	AUROC	AUPRC	AUROC	AUPRC
RNN-Mean	89.5 ± 0.3	62.9 ± 0.7	98.4 ± 0.2	83.6 ± 0.8	76.7 ± 0.1	31.2 ± 0.1	79.4 ± 0.3	25.8 ± 0.6	86.9 ± 0.3	45.6 ± 0.2
RNN-Forward	88.5 ± 0.3	60.4 ± 1.2	97.1 ± 0.6	75.6 ± 2.2	76.4 ± 0.2	30.8 ± 0.1	74.6 ± 4.9	22.4 ± 3.0	86.4 ± 0.4	45.2 ± 0.3
RNN- Δ_t	87.4 ± 1.0	55.8 ± 2.4	97.3 ± 0.5	74.5 ± 0.8	76.2 ± 0.2	30.6 ± 0.2	72.5 ± 4.2	20.4 ± 2.1	50.6 ± 0.6	25.2 ± 0.2
RNN-Decay	89.1 ± 0.4	62.6 ± 1.0	98.4 ± 0.3	84.2 ± 2.4	77.0 ± 0.1	31.6 ± 0.1	77.8 ± 2.2	24.5 ± 1.9	76.2 ± 9.6	39.8 ± 4.6
GRU-D	87.6 ± 0.5	59.9 ± 0.6	97.6 ± 1.0	77.9 ± 5.6	76.5 ± 0.1	30.9 ± 0.3	74.4 ± 6.9	22.7 ± 4.9	84.6 ± 1.8	43.3 ± 1.7
Phased-LSTM	86.7 ± 0.3	53.7 ± 0.6	97.5 ± 0.1	78.2 ± 0.4	75.6 ± 0.2	30.2 ± 0.1	77.4 ± 0.3	24.3 ± 0.4	84.8 ± 0.2	44.0 ± 0.1
SeFT	88.0 ± 0.4	59.3 ± 1.4	98.5 ± 0.2	86.7 ± 0.7	76.1 ± 0.1	31.0 ± 0.1	83.8 ± 0.4	33.1 ± 0.3	86.6 ± 0.2	45.0 ± 0.1
RainDrop	87.6 ± 0.1	59.1 ± 0.4	98.0 ± 0.2	82.7 ± 0.5	76.0 ± 0.2	30.6 ± 0.4	80.4 ± 0.4	27.7 ± 0.6	86.9 ± 0.4	44.8 ± 0.2
IP-Nets	88.9 ± 0.3	61.9 ± 0.9	98.3 ± 0.1	85.2 ± 0.6	74.3 ± 3.6	28.8 ± 3.7	81.6 ± 0.1	28.1 ± 0.6	85.2 ± 0.6	44.0 ± 0.6
mTAND	89.0 ± 0.2	61.8 ± 0.7	97.2 ± 0.3	74.5 ± 3.2	73.8 ± 0.4	28.3 ± 0.4	66.1 ± 0.2	16.7 ± 0.1	84.2 ± 0.3	42.1 ± 0.2
AdaCare	76.0 ± 0.5	47.7 ± 0.7	94.9 ± 2.0	68.5 ± 9.0	64.9 ± 0.6	22.1 ± 1.2	51.2 ± 0.2	12.9 ± 0.0	64.4 ± 2.4	36.2 ± 1.9
STraTS	89.3 ± 0.1	61.3 ± 0.3	98.6 ± 0.1	84.0 ± 1.6	76.6 ± 0.2	31.4 ± 0.3	79.1 ± 2.1	26.2 ± 2.4	87.8 ± 0.2	45.6 ± 0.2
Warpformer	90.3 ± 0.1	64.6 ± 0.4	99.0 ± 0.1	90.0 ± 0.4	77.7 ± 0.2	32.5 ± 0.2	85.5 ± 0.1	35.5 ± 0.3	88.0 ± 0.2	46.4 ± 0.2

MIMIC -III数据集



特点总结与未来展望

- 算法优势
 - 引入**图神经网络**算法，建模序列间的动态拓扑关系
 - 改进Transformer模型，同时解决序列内、序列间关系计算问题；且可以实现任意尺度的变化输入（**上、下采样**）
- 算法劣势
 - 动态切片的信息传递会导致过多**孤立节点**存在
 - 将序列归一化至统一的时间长度，未考虑不同采样信号的**采样率存在差异**的问题
- 工作展望
 - 针对不同变量的尺度变化进行适应性调整
 - 图神经网络的应用，是否可以针对**数据特点**进行迁移

- [1] Zhou H, Zhang S, Peng J, et al. Informer: Beyond efficient transformer for long sequence time-series forecasting. Proceedings of the AAAI conference on artificial intelligence[C]. PaloAlto, CA: AAAI, 2021, 35(12): 11106-11115.
- [2] Zhang X, Zeman M, Tsiligkaridis T, et al. Graph-guided network for irregularly sampled multivariate time series [EB/OL]. (2022-08-15) [2024-02-20] <https://arxiv.org/abs/2110.05357>.
- [3] Zhang J, Zheng S, Cao W, et al. Warpformer: A Multi-scale Modeling Approach for Irregular Clinical Time Series [EB/OL]. (2023-08-10) [2024-02-20] <https://arxiv.org/abs/2306.09368>.
- [4] 动态时间规划 [EB/OL].(2013-06-13)[2024-02-20] <https://blog.csdn.net/zouxy09/article/details/9140207?spm=1001.2014.3001.5501>

知人者智，自知者明。胜人者有力，自胜者强。知足者富。强行者有志。不失其所者久。死而不亡者，寿。

谢谢！

