

北京師範大學

人工智能学院

硕士研究生开题报告

研究 生 姓 名 陈治豪

学 号 202321081007

导 师 姓 名、职 称 余先川 教授

系 所 人工智能学院

专 业 计算机科学与技术

研 究 方 向 模式识别与数据挖掘

入 学 时 间 2023 年 9 月

论 文 题 目 基于神经随机微分方程的

不规则采样时间序列分类方法研究

一、立论依据：

1 论文选题的依据

1.1 选题背景

时间序列数据广泛存在于现代科学与工程领域，包括天文光变曲线观测[1]、医疗健康监测[2]和工业设备监控[3]等众多应用场景。随着数据采集技术的进步和存储能力的提升，时间序列数据呈现出高维度、高复杂度和高动态性等特点，给数据分析和模式识别带来巨大挑战[4]。特别是在实际应用中，时间序列数据往往是观测时间不可控的，而是呈现出不规则采样的特性，同时蕴含着难以直接观测的内在随机性，这进一步增加了数据处理和分析的难度。

不规则采样时间序列(Irregularly-Sampled Time Series, ISTS)是指观测点在时间轴上分布不均匀的序列数据，其采样间隔可能因多种因素而变化，如传感器故障、资源限制、观测条件变化或特定领域的实际需求等[5]。在天文观测中，光变曲线数据受天气条件和观测设备可用性的影响，形成高度不规则的采样模式[6]；在医疗领域，患者生理指标的监测频率往往根据临床需要而动态调整。这种不规则采样特性使得传统的时间序列分析方法难以直接应用，因为它们通常假设数据是等间隔采集的[7]。

与等间隔采样的时间序列相比，不规则采样时间序列分析面临诸多技术挑战。首先，观测点之间的时间间隔变化本身蕴含着重要的语义信息——医疗监测中采样频率的突然增加可能预示患者状态恶化，天文观测中的密集采样往往对应重要天体事件的发生[1]，简单地忽略这一特性会导致关键信息的损失。其次，不规则采样导致的长时间间隔使系统演变充满不确定性，数据中蕴含的内在随机性——如测量噪声、环境扰动、系统内部的随机波动等，进一步增加了建模难度[8]。此外，不规则采样破坏了许多经典时序分析方法的基本假设，如傅里叶变换要求等间隔采样[6]、卷积神经网络假设固定的感受野等，这些方法在处理不规则时序时要么需要复杂额外的预处理步骤，要么会产生一定的性能下降。

同时，不规则采样时间序列分类任务中普遍存在严重的类别不平衡问题[9]。在实际应用中，关键事件或异常状态往往是稀有的——如天文观测中的特殊天体事件、医疗监测中的疾病发作等，这些少数类样本通常只占总体样本的极小比例。这种不平衡不仅体现在样本数量上，更体现在时序模式的复杂度差异上。少数类往往具有更加多样和复杂的时序演化模式，但由于样本稀缺而难以被充分学习。传统的分类方法在面对这种极度不平衡的数据时，往往偏向于多数类，导致对关键少数类

的识别能力严重不足，这在实际应用中可能造成重大损失。

综上所述，如何在保持时序动态特征完整性的同时有效处理极度类别不平衡问题，已成为不规则采样时间序列分类领域亟待解决的核心难题[10]。这一挑战不仅深刻影响着相关理论方法的创新方向，更直接关系到诸多重要应用领域中智能决策系统的可靠性与实用性。

1.2 论文选题的研究意义

理论意义：目前国内外学者在深度学习应用于时间序列分类任务上已经进行了大量的研究，但这类时间序列任务普遍是等间隔采样的，时间序列的形状可以描绘出数据的基本特征。然而针对不规则采样时间序列的分类研究却十分罕见，这是由于不规则时间序列存在大量信息缺失，以往的深度学习方法难以捕获其时间上的依赖关系。为此，本文研究的不规则采样时间序列分类方法可以进一步完善深度学习应用于不规则采样时间序列的不足之处，充分挖掘潜藏在不规则采样中的时间序列特征，并提出一系列针对不规则采样时间序列分类的新理论和新方法。

实际意义：不规则采样时间序列分类是指将不同时间间隔采集的序列数据标记为对应的类别，这是利用时序数据进行模式识别中最关键的一步。利用深度学习方法完成不规则时序自动分类可以促进时域数据分析与人工智能科学的深入交叉，加速各类时序数据的自动化处理速度，提升分类识别的准确性。进一步的，本文的研究有利于推动多个应用领域的智能化发展——在医疗健康监测中，可以准确识别患者状态变化和疾病发展趋势；在天文观测中，可以高效完成变星分类和天体事件检测。本文提出的基于神经随机微分方程和连续时间类别敏感 Transformer 的方法，为处理不规则采样时序数据提供了新的技术路径。该方法不仅能够有效建模时序的连续动态特性，还能解决类别不平衡这一普遍存在的问题，为医疗监测、天文观测、等领域的不规则采样时间序列分类任务提供了可行的解决方案。

2 国内外研究现状分析

2.1 基于离散时间建模的不规则时间序列处理方法

基于离散时间建模的时间序列分析方法经历了从基于统计学的传统方法到现代深度学习模型的演进过程。传统方法主要依赖于数据的统计特性，而深度学习方法则通过自动学习特征表示提升了分析性能。

早期的时间序列分析主要依赖于统计模型。自回归积分移动平均模型(ARIMA)是处理线性相关时间序列的经典方法，通过结合自回归、差分和移动平均组件来捕捉数据的趋势、季节性和随机

波动[11]。指数平滑法是另一种应用广泛的方法，基于“近期数据比远期数据更重要”的原则，为时间序列中的观测值分配指数递减的权重[12]。隐马尔可夫模型(HMM)则用于建模时间序列背后的隐状态转换，特别适用于状态变化显著的场景[13]。

对于具有非线性和非平稳特性的时间序列，动态时间规整(Dynamic Time Warping, DTW)提供了一种计算两个时间序列相似度的有效方法，即使它们在时间轴上存在伸缩或位移[14]。支持向量机(SVM)通过核函数将时间序列映射到高维空间进行分类，适用于样本量较小但维度较高的情况[15]。小波变换则提供了时间-频率联合分析能力，能同时捕捉时间序列的局部和全局特征[16]。集成经验模态分解(EEMD)进一步提供了自适应多尺度分析能力，能将复杂非线性非平稳信号分解为有限个本征模态函数[17]。

然而，这些传统方法普遍存在几个局限性：它们通常需要强假设（如线性相关或平稳性）、对系统的先验知识，且特征提取过程往往依赖专家经验，难以处理高维复杂数据。此外，它们中的大多数方法设计初衷是针对等间隔采样数据，面对不规则采样时间序列时需要额外的数据预处理步骤，这可能导致信息损失。

为克服传统方法的局限性，研究者转向深度学习技术。循环神经网络(RNN)及其变体长短时记忆网络(LSTM)和门控循环单元(GRU)通过循环连接机制建模时间依赖，成为时间序列分析的重要工具[18,19]。然而，RNN系列模型仍面临长序列处理中的梯度消失问题，且计算效率较低。

卷积神经网络(CNN)通过卷积操作提取局部特征，并通过池化操作实现平移不变性，在时间序列分析中展现了优越性能[20]。TCN(Temporal Convolutional Network)结合了扩张卷积和残差连接，实现了更大的感受野和更稳定的训练过程[21]。然而，CNN结构假设数据在时间轴上均匀分布，这与不规则采样时间序列的本质相悖。

时间序列分析中引入图表示学习，为捕捉复杂时间动态和节点间相关性提供了新思路。空时图卷积网络(STGCN)将图卷积与时间卷积相结合，有效建模时空相关性[22]。DCRNN(Diffusion Convolutional Recurrent Neural Network)通过随机游走模拟交通网络中的扩散过程，同时整合循环结构捕捉时间依赖[23]。TGCN(Temporal Graph Convolutional Network)则融合图卷积与门控循环单元，实现节点间空间依赖和时间动态的联合建模[24]。ASTGCN(Attention based Spatial-Temporal Graph Convolutional Networks)通过引入空间和时间注意力机制，增强了模型对关键节点和时间点的感知能力[25]。这类方法将时间序列视为图上的动态信号，通过节点间的连接关系捕捉变量间相互影响。然而，对于不规则采样数据，图构建本身仍是一个挑战，且这些方法通常需要预定义的图结构作为先验知识。

注意力机制通过动态分配权重，使模型能够聚焦于输入序列中的关键部分。Transformer 凭借其自注意力机制和并行计算优势成为时间序列分析的强大工具[26]。与 RNN 相比，直接建模序列中任意两点间的依赖关系，避免了长距离传播问题。针对时间序列特性，各种变体被提出：Informer 通过 ProbSparse 自注意力机制降低了计算复杂度，适用于长序列预测[27]；Autoformer 引入自相关机制分解时间序列，提升了长期趋势捕捉能力[28]；LogTrans 提出对数稀疏注意力，实现更高效的长序列建模[29]。

为解决不规则采样问题，多种基于注意力的方法被相继提出。

MTANS(Multi-Task Attention-based Neural Signals)模型通过多任务学习框架和特化的注意力机制，有效处理医疗信号等多变量时间序列数据[30]。Warpformer 提出时间扭曲注意力机制，通过学习时间点之间的相对位置编码，适应不同采样率的数据[31]。PRIME(PRogressive Irregular time series ModEling)通过层次化架构和位置感知注意力，实现了对不规则采样模式的有效建模[32]。Contiformer 则引入连续表示学习，将离散时序观测点映射到连续函数空间，突破了传统注意力机制对等间隔输入的依赖[33]。

2.2 基于连续时间建模的不规则时间序列处理方法

随着深度学习技术的快速发展，如何有效处理不规则采样时间序列已成为时间序列分析领域的核心挑战。传统的离散时间序列方法在面对不规则采样数据时往往表现不佳，这是因为简单的时间离散化会破坏数据的连续性，导致重要的时间依赖信息丢失[34]。为解决这一问题，研究者们开始探索基于连续时间过程的建模方法，以更准确地捕捉系统的动态演化特性。

早期的连续时间建模研究主要基于经典的统计学方法。线性高斯状态空间模型 (Linear Gaussian State-space Model, LGSSM) [35]通过卡尔曼滤波技术处理缺失数据，为后续方法奠定了理论基础。切换状态线性动力系统 (Switching State Linear Dynamical System, SSLDS) [36]引入离散切换状态来捕获非线性动态。高斯过程状态空间模型 (Gaussian Process State-space Model, GP-SSM) [37]和多任务高斯过程 (Multi-task Gaussian Process, MTGP) [38]采用非参数化方法，能够灵活地建模复杂系统。连续时间贝叶斯网络 (Continuous Time Bayesian Network, CTBN) [39]和连续时间隐马尔可夫模型 (Continuous Time Hidden Markov Model, CTHMM) [40]则从概率图模型角度提供了新的建模框架。随机优化控制框架 (Stochastic Optimization Control Framework, SOCF) [41]将问题转化为随机最优控制，通过求解 Hamilton-Jacobi-Bellman 方程学习系统动态。

深度学习的兴起为连续时间建模带来了革命性的突破，彻底改变了处理不规则时间序列的研究范式。连续时间递归神经网络 (Continuous Time Recurrent Neural Network, CTRNN) [42]作为这一领域的开创性工作，首次用微分方程描述神经元的连续时间活动，突破了传统离散时间递归神经网络 (Recurrent Neural Network, RNN) 的局限性。该方法通过连续动力系统理论，使网络能够自然地处理任意时间间隔的输入，为不规则采样数据的建模开辟了全新思路。在 CTRNN 的启发下，常微分方程循环神经网络 (Ordinary Differential Equation Recurrent Neural Network, ODE-RNN) [43]进一步将循环神经网络推广为由神经常微分方程 (Neural Ordinary Differential Equation, NODE) [44]定义的连续时间隐藏动力学系统。与传统 RNN 在离散时间点更新隐藏状态不同，ODE-RNN 通过求解微分方程来演化隐藏状态，使其能够在任意时间点进行预测，实现了对不规则采样的自然处理。这种方法不仅保留了 RNN 的序列建模能力，还避免了插值带来的信息损失。潜在常微分方程 (Latent Ordinary Differential Equation, Latent ODE) [45]则将这一思想与生成模型相结合，通过变分自编码器框架学习连续时间的潜在表示，能够同时进行状态推断和轨迹预测，在医疗时间序列分析等领域展现出巨大潜力。

在具体架构设计方面，研究者们致力于将经典的深度学习结构扩展到连续时间域。连续时间门控循环单元 (Continuous Time Gated Recurrent Unit, CTGRU) [46]创新性地将门控循环单元 (Gated Recurrent Unit, GRU) 的门控机制与常微分方程结合，使用连续时间衰减函数替代离散更新，既保持了 GRU 的长期依赖建模能力，又实现了对不规则时间间隔的自适应。连续时间长短期记忆网络 (Continuous Time Long Short-Term Memory, CT-LSTM) [47]则引入了相位调制机制，通过学习每个神经元的时间尺度来适应不同频率的时间模式，在事件序列和长期依赖建模任务中表现出色。从完全不同的视角出发，神经霍克斯过程 (Neural Hawkes Process, NHP) [48]基于点过程理论，通过神经网络参数化的自调制机制建模事件序列的时间依赖性，特别适合处理事件驱动的不规则时间序列数据。

近年来，神经微分方程领域取得了显著进展。神经控制微分方程 (Neural Controlled Differential Equation, Neural CDE) [49]将连续路径作为控制信号，通过控制微分方程建模隐藏状态演化，在处理高度不规则数据时表现优异。自适应神经控制微分方程 (Adaptive Neural Controlled Differential Equation, ANCDE) [50]和神经粗糙微分方程 (Neural Rough Differential Equation, NeuralRDE) [51]进一步提升了计算效率和长期依赖建模能力。特别值得关注的是，神经随机微分方程 (Neural Stochastic Differential Equation, Neural SDE) [52]引入了随机性，提出了朗之万型、线性噪声型和几何型三种稳定的 SDE 形式，在分布偏移下展现出卓越的鲁棒性。连续循环单元 (Continuous Recurrent Unit, CRU) [53]假设隐藏状态遵循线性随机微

分方程，通过卡尔曼滤波推导闭式解，实现了高效计算。

尽管上述方法在连续时间建模方面取得了重要进展，但仍存在诸多局限性。首先，早期的统计模型虽然理论完备，但表达能力有限，难以捕获复杂的非线性动态，且大多依赖强假设条件。其次，基于神经常微分方程的方法虽然提升了建模能力，但计算成本高昂，特别是在长序列和高维数据上，数值求解器的迭代过程严重影响了训练和推理效率。第三，现有的神经控制微分方程方法虽然在理论上优雅，但实际应用中对初始条件敏感，且难以处理观测噪声较大的情况。最后，大多数方法缺乏对不确定性的有效量化，这在实际应用中尤为重要，特别是在高风险决策领域。因此，如何在保持模型表达能力的同时提高计算效率、增强鲁棒性并提供可靠的不确定性估计，成为当前亟需解决的关键问题。

2.3 基于不规则时间序列的不均衡分类方法

在实际应用中，不规则时间序列分类任务经常面临类别分布严重不均衡的问题。例如，在医疗诊断中，疾病患者往往远少于健康人群；在故障检测中，异常事件的发生频率远低于正常状态；在金融欺诈检测中，欺诈交易仅占极小比例。这种类别不均衡问题在不规则采样条件下变得更加复杂，因为稀有类别的样本不仅数量少，而且采样模式可能更加不规则，导致传统分类方法严重偏向多数类，难以有效识别少数类模式。

数据层面的解决方案主要通过重采样技术平衡类别分布。合成少数类过采样技术 (Synthetic Minority Over-sampling Technique, SMOTE) [54]通过在少数类样本之间的特征空间中生成合成样本，避免了简单复制导致的过拟合。然而，SMOTE 在时间序列场景中直接应用存在困难，因为它忽略了时间依赖性。自适应合成采样 (Adaptive Synthetic Sampling, ADASYN) [55]根据学习难度自适应生成样本，为边界区域生成更多合成样本。聚类基础过采样 (Cluster-Based Over Sampling, CBOS) [56]先对少数类聚类，再在簇内过采样，保持了数据分布特性。基于支持向量机的 SMOTE (Support Vector Machine based SMOTE, SVMSMOTE) [57]利用 SVM 识别边界样本，提高了合成样本质量。近年来，基于生成对抗网络的过采样方法[58]通过学习数据分布生成高质量少数类样本，也为时间序列数据的合成提供了新思路。

欠采样方法通过减少多数类样本来平衡数据集。EasyEnsemble[59]通过多次随机欠采样训练多个基分类器，充分利用了多数类信息。RUSBoost (Random Under-Sampling Boost) [60]结合随机欠采样和 AdaBoost，在每次迭代中动态调整采样策略。随机过采样示例生成 (Random Over-Sampling Examples, ROSE) [61]结合了过采样和欠采样的优点，通过平滑自助法生成合成样本。动态重采样 (Dynamic

Resampling, DRS) [62]根据训练过程中的模型表现动态调整采样策略, 避免了固定策略的局限性。

算法层面的方法通过修改损失函数或学习策略来处理不均衡问题。成本敏感学习 (Cost-Sensitive Learning) [63]为不同类别的错误分配不同成本, 使模型偏向少数类。焦点损失 (Focal Loss) [64]动态调整样本权重, 让模型更关注难分类样本, 在目标检测等任务中取得巨大成功。类平衡损失 (Class-Balanced Loss) [65]基于有效样本数重新加权, 理论上保证了长尾分布下的性能。标签分布感知边际损失 (Label Distribution-Aware Margin Loss, LDAM) [66]调整不同类别的分类边界, 为少数类提供更大的决策空间。分布平衡损失 (Distribution-Balanced Loss, DBLOSS) [67]同时考虑类别频率和样本难度, 在语义分割等任务中表现优异。类平衡交叉熵 (Class-Balanced Cross Entropy, CBCE) [68]通过简单的反频率加权实现类别平衡, 易于实现且效果显著。

混合方法结合了数据层和算法层的优势。SMOTEBoost[69]将 SMOTE 与 Boosting 结合, 在每轮迭代中对少数类过采样。两阶段加权采样 (Two-stage Weighted Sampling, TWS) [70]先通过过采样平衡类别, 再根据样本难度重加权。元学习方法近年来在不均衡分类中展现出巨大潜力。学习重加权样本 (Learning to Reweight Examples) [71]通过元学习自动学习样本权重。元学习语义增强 (Meta-learning Semantic Augmentation, MetaSAug) [72]学习如何为少数类生成语义相关的增强样本。元采样 (Meta Sampling, MESA) [73]通过元学习框架自动学习最优采样策略, 避免了手动设计的局限性。

特征表示学习为不均衡分类提供了新的解决思路。大边际局部嵌入 (Large Margin Local Embedding, LMLE) [74]通过学习判别性特征表示, 最大化类间边际并保持类内紧凑性。自监督多任务表示学习 (Self-supervised Multi-task Representation Learning, SMRT) [75]设计多个自监督辅助任务, 学习对类别不均衡更鲁棒的特征表示。平衡元 softmax (Balanced Meta-Softmax) [76]在特征空间中平衡不同类别的表示, 提高了长尾识别性能。

然而, 现有的不均衡分类方法在处理不规则时间序列时面临独特挑战。首先, 传统的重采样方法难以保持时间依赖性, 简单的插值或复制可能破坏时间动态。其次, 不规则采样本身就包含重要信息, 例如在医疗场景中, 检查频率的变化可能反映病情严重程度, 而重采样可能丢失这些信息。第三, 少数类样本的不规则性往往更强, 使得学习其时间模式更加困难。

因此，将连续时间建模与不均衡分类相结合成为必然趋势。连续时间表示能够自然地处理不规则采样，避免了离散化带来的信息损失；同时，通过在连续时间域中设计针对性的平衡策略，可以更好地保持时间动态特性。例如，可以在神经微分方程框架下设计类别感知的损失函数，并通过连续时间过程生成高质量的少数类轨迹。这种结合不仅能够提高分类性能，还能提供更好的可解释性和不确定性量化，为高风险应用领域提供可靠的决策支持。

二、研究方案:

1. 研究目标、研究内容和拟解决的关键问题:

1.1 研究目标

本研究拟解决不规则采样时间序列分类中存在的不规则时空建模困难、少数类识别率低等问题。通过构建神经随机微分方程与连续时间 Transformer 协同框架，设计类别感知的分类模型，以实现对不规则时序的精确连续表示建模，提高天文观测和医疗检测等领域不规则不均衡时序数据分析的准确性和可靠性。

具体而言，针对不规则采样和长间隔预测问题，采用线性神经 SDE 构建连续路径表示并精确计算状态转移；针对不规则时序的时序依赖建模问题，引入连续时间 Transformer 架构，通过连续时间的注意力机制有效捕捉长时间、不均匀跨度的复杂模式；针对类别不平衡问题，结合混合重采样和类别感知分组注意力机制，从数据层和模型层协同优化。最终构建一个对复杂时序数据具有强大适应能力的分类框架，为相关领域提供高效可靠的数据分析工具。

1.2 研究内容

(1) 基于 Neural SDE 和连续时间 Transformer 的不规则时序连续时间建模

针对单独使用 Neural SDE 缺乏全局依赖建模能力、单独使用连续时间 Transformer 无法刻画系统内在动力学的根本缺陷，提出两者结合的架构。Neural SDE 虽能通过微分方程描述连续演化过程，但其马尔可夫性假设限制了对历史信息的利用，难以捕获远距离时间点之间的复杂依赖；连续时间 Transformer 虽能通过注意力机制建模任意时间点间的关联，但缺乏对系统动力学规律的显式建模，在长时间外推时容易失真。通过将微分方程的动力学建模能力与 Transformer 的全局依赖捕获能力相结合，既保证了物理规律的连续性约束，又实现了跨时间尺度的灵活建模，解决了单一方法在不规则时序建模中的根本局限。

(2) 基于混合重采样和 CGA 的类别不平衡时序分类

针对只使用重采样导致模型过度依赖合成样本、只使用 CGA 在极度不平衡下难以学到有效表示的困境，提出数据-模型协同优化策略。仅依靠混合重采样虽能平衡训练集分布，但合成样本与真实样本的分布差异可能误导模型学习，且无法解决测试集仍然不平衡的问题；仅依靠 CGA 虽能增强少数类特征学习，但在训练样本极度稀缺时，注意力机制缺乏足够的样本多样性来学习鲁棒的类别特定模式。通过重采样为 CGA 提供更丰富的训练信号，同时 CGA 的类别感知机制补偿合成样本的分布偏差，实现了从数据增强到特征学习的全流程优化，有效解决了极度不平衡

场景下的分类问题。

1.3 拟解决的关键问题

(1) 如何实现不规则采样时序的连续建模

不规则采样时间序列的核心挑战在于观测点时间分布的随机性和采样间隔的巨大差异，传统深度学习模型依赖固定时间步长假设，难以自然处理这种不规则性。现有方法往往采用简单的插值或时间编码技术，不仅造成信息损失，更无法准确捕获时序数据中的动态演化规律。在医疗监测数据根据临床需要采集、天文观测受天气和设备限制形成高度不规则的观测模式等实际场景中，数据呈现出复杂的时空耦合特性，如何构建一个统一的深度学习架构，既能自然处理不规则采样特性，又能准确建模连续时间动态，同时在长时间间隔情况下保持预测精度并量化预测不确定性，是本研究需要解决的首要关键问题。

(2) 如何解决不规则采样时序的类别不均匀的问题

在时间序列分类任务中的类别不平衡是一个普遍且严重的问题，在天文光变数据、医疗诊断等关键应用中，目标类别的样本往往只占总体的极小比例。这种不平衡不仅体现在样本数量上，更体现在时序模式的复杂度差异上——少数类往往具有更加多样和复杂的时序演化模式，但由于样本稀缺而难以被充分学习。现有解决方案通常将时序建模和类别平衡作为两个独立问题分别处理，忽视了它们之间的内在联系，简单的重采样技术难以保持时序数据的连贯性和物理意义，通用的类别平衡损失函数未能考虑时序特有的上下文依赖性，如何在设计专门的架构，使少数类在特征空间中获得充分表达的同时保持时序建模的准确性，实现不同类别间的有效信息交互，从数据层和模型层协同优化构建端到端的解决方案，是本研究面临的另一个关键挑战。

2. 拟采取的研究方法（或技术路线、实验方案）及可行性分析

2.1 研究框架概述

本研究提出一种创新的两阶段串联架构来处理不规则采样时间序列的不均衡分类问题。该架构的核心思想是将连续时间建模与类别感知机制有机结合，通过先建模、后平衡的策略系统性地解决不规则采样和类别不均衡的双重挑战。

具体而言，第一阶段采用基于神经随机微分方程的连续时间建模方法，将不规则采样的离散观测嵌入到连续时间随机过程中。这一阶段包含两个关键组件：(1) 线性神经随机微分方程 (LNSDE) 作为基础框架，通过解析解形式实现高效的状态演化建模；(2) 连续时间注意力机制，通过重参数化技术在保持 Transformer 并行计算

优势的同时，实现对任意时间间隔依赖关系的精确捕捉。这两个组件的协同作用确保了对不规则时间序列动态特性的准确建模。

第二阶段在连续时间表示的基础上，通过混合重采样策略和改进的类别感知分组注意力 (CGA) 机制处理类别不均衡问题。数据层面采用时序感知的智能重采样，在平衡类别分布的同时保持时间连贯性；模型层面设计双路径注意力架构，为不同类别，特别是少数类提供专属的特征学习通道。

这种串联结构的合理性体现在三个方面：首先，连续时间建模为后续的类别平衡处理提供了高质量的特征表示，避免了在原始不规则数据上直接进行重采样可能破坏时间依赖性的问题；其次，在连续时间域中进行类别平衡操作能够更好地保持时序的连贯性和动态特性；最后，两阶段的解耦设计使得每个模块可以独立优化，提高了方法的灵活性和可扩展性。通过 LNSDE 的解析求解能力和连续注意力机制的全局建模能力，第一阶段为第二阶段的类别平衡处理奠定了坚实基础，确保了整个框架在处理复杂时序分类任务时的有效性和鲁棒性。

2.2 不规则采样时间序列的连续时间建模

(1) 线性神经随机微分方程(LNSDE)框架

不规则采样时间序列在实际应用中普遍存在，传统的离散时间模型通过插值或时间编码处理不规则采样，不可避免导致信息损失。本研究采用线性神经随机微分方程 (LNSDE) 作为连续时间建模的基础框架，将离散观测嵌入到连续随机过程中。

LNSDE 将系统状态演化建模为：

$$dz_t = f_\theta(z_t, t)dt + g_\phi(z_t, t)dW_t \quad (1)$$

其中 z_t 为隐状态， f_θ 为漂移函数， g_ϕ 为扩散函数， dW_t 为维纳过程增量。为实现高效推理，采用线性化形式： $dz_t = A_t z_t dt + b_t dt + \sigma_t dW_t$

这种线性化设计的优势在于存在解析解：

$$z_t = \Phi(t, s)z_s + \int_s^t \Phi(t, r)b_r dr + \int_s^t \Phi(t, r)\sigma_r dW_r \quad (2)$$

其中 $\Phi(t, s)$ 为状态转移矩阵。解析解避免了数值积分的误差累积，在长时间跨度预测中尤为重要。

(2) 连续时间注意力机制

为克服传统 Transformer 在处理不规则时间序列时的局限，本研究设计了连续

时间注意力机制。该机制通过常微分方程建模注意力分数在连续时间域的演化，能够精确捕捉任意时间间隔内的依赖关系。

对于查询时刻 t_j 和历史时刻 t_i ，连续注意力权重定义为：

$$\alpha_i(t_j) = \frac{\int_{t_i}^{t_j} q(\tau) \cdot k_i(\tau)^T d\tau}{t_j - t_i} \quad (3)$$

为保持计算效率，采用重参数化技术将积分映射到标准区间：

$$\alpha_i(t_j) \approx \frac{1}{2} \sum_{p=1}^P \gamma_p \tilde{q}_{i,j}(\xi_p) \cdot \tilde{k}_{i,j}(\xi_p)^T \quad (4)$$

这种设计保持了 Transformer 的并行计算优势，同时实现了连续时间建模能力。通过多头注意力扩展，不同注意力头可以专注于不同时间尺度的依赖模式，进一步提升了模型的表达能力。

2.3 基于混合重采样和改进 CGA 的类别不平衡时序分类

针对不规则采样时间序列分类中的类别分布严重失衡问题，本研究提出基于混合重采样策略和改进类别感知分组注意力 (CGA) 机制的综合解决方案。该方案从数据层和模型层协同出发，系统性地解决少数类识别率低的核心挑战。

在数据层面，我们设计了智能混合重采样策略来平衡类别分布 2。对于少数类样本，采用改进的 SVM-SMOTE 算法进行智能合成。该算法首先识别少数类中靠近决策边界的关键样本，然后通过线性插值生成新的合成样本。具体地，对于少数类样本 x_i 及其 k 近邻中的样本 x_j ，合成样本生成公式为：

$$x_{\text{new}} = x_i + \lambda \cdot (x_j - x_i) \quad (5)$$

其中 $\lambda \in [0,1]$ 为随机插值系数。特别地，在时间序列场景下，我们不仅在特征空间进行插值，还需要保持时序的连贯性。因此，对于时间戳 t_i 和 t_j ，相应的时间插值为：

$$t_{\text{new}} = t_i + \lambda \cdot (t_j - t_i) \quad (6)$$

这确保生成的样本既符合原始数据的特征分布，又保持合理的时间间隔模式。

对于多数类样本，我们采用改进的 Repeated ENN (Edited Nearest Neighbor) 方法进行选择性欠采样[2]。该方法通过迭代清理多数类中的噪声和边界样本，判定规

则为:

$$R(x_i) = \sum_{j \in kNN(x_i)} \mathbb{I}(y_j \neq y_i) > \theta \cdot k \quad (7)$$

其中 $kNN(x_i)$ 表示样本 x_i 的 k 个近邻, $\mathbb{I}(\cdot)$ 为指示函数, $\theta \in (0,1)$ 为噪声判定阈值。当 $R(x_i)$ 为真时, 样本 x_i 被识别为噪声或边界样本并被移除。这种方法保留最具代表性的核心样本, 在减少样本数量的同时保持数据质量。

在模型层面, 我们提出了基于类别感知分组的注意力机制。该机制的核心创新在于为每个类别 c 构建独立的特征表示路径, 同时设计了类间语义交互模块。具体而言, 输入特征 $X \in \mathbb{R}^{N \times d}$ 首先通过类别感知编码器生成类别特定的表示:

$$Z_c = Attention_c(XW_c^Q, XW_c^K, XW_c^V) \quad (8)$$

其中 $W_c^Q, W_c^K, W_c^V \in \mathbb{R}^{d \times d_c}$ 为类别 c 的专属投影矩阵, d_c 为类别特定的隐藏维度。这种设计确保了少数类能够获得充分的表示能力, 不会被多数类的特征所掩盖。

为了避免类别间信息的完全隔离, 我们设计了基于语义相似度的类间交互机制。不同类别特征表示之间的语义相似度计算为:

$$S_{ij} = \frac{Z_i^T Z_j}{\|Z_i\| \|Z_j\|} \quad (9)$$

基于相似度矩阵 S , 类间信息交换通过门控机制进行:

$$\tilde{Z}_i = Z_i + \sum_{j \neq i} \sigma(S_{ij} - \tau) \cdot \text{Gate}_{ij}(Z_j) \quad (10)$$

其中 $\sigma(\cdot)$ 为 sigmoid 函数, τ 为相似度阈值, Gate_{ij} 为可学习的门控函数。这种设计既保护了各类别 (特别是少数类) 的独特特征, 又能够利用类别间的互补信息提升整体分类性能。

此外, 我们还引入了自适应的加权交叉熵损失函数:

$$\mathcal{L} = - \sum_{i=1}^N \sum_{c=1}^C w_c \cdot y_{ic} \log(\hat{y}_{ic}) \quad (11)$$

其中权重 $w_c = \frac{N}{N_c \cdot C}$ N 为总样本数, N_c 为类别 c 的样本数, C 为类别总数。这种机

制确保模型在整个训练过程中持续关注少数类的学习，避免在训练后期被多数类主导。通过上述数据层和模型层的协同优化，本方法能够有效解决不规则时序分类中的类别不平衡问题。

2.4 可行性分析

(1) 理论基础成熟可靠

本研究涉及的核心理论方法均有坚实的数学基础和广泛的研究支撑。神经随机微分方程 (Neural SDE) 理论已在金融建模、物理系统模拟等领域得到验证，其解析解特性和不确定性量化能力为不规则时序建模提供了可靠保障。线性 SDE 的解析解形式已被证明能够避免数值积分的误差累积问题，特别适用于长时间跨度预测。连续时间注意力机制虽然相对较新，但其基于 ODE 的理论框架已在多项研究中被验证有效，重参数化技术确保了计算效率。类别不平衡处理方面，SMOTE、Repeated ENN 等重采样方法已经过大量实证检验，在不同领域被证实能够有效改善少数类识别性能。这些成熟理论的有机结合为本研究提供了坚实的理论支撑。

(2) 技术方案切实可行

本研究提出的两阶段串联架构具有良好的工程可行性。首先，LNSDE 的线性化设计使得状态转移计算可以通过矩阵运算高效实现，避免了复杂的数值积分过程。连续时间注意力机制通过重参数化保持了 Transformer 的并行计算优势，无需修改底层计算框架即可在现有深度学习平台上实现。其次，混合重采样策略和 CGA 机制均可作为独立模块集成到整体框架中，模块化设计提高了系统的可维护性和可扩展性。各模块之间通过标准的张量接口进行数据交互，确保了系统集成的便利性。此外，本研究的关键算法均有开源实现可供参考，降低了技术实现的难度和风险。

(3) 数据资源充分可得

时间序列分类领域已积累了大量公开数据集，为本研究提供了充分的实验资源。这些数据集不仅具有不规则采样特性，还大多数存在类别不平衡问题，完全符合本研究的应用场景。数据集的多样性也保证了方法的泛化性验证。同时，这些数据的收集和使用均遵守相关法律和伦理规范，确保研究的合规性。

(4) 研究条件完备充足

本研究团队在时间序列分析、深度学习和不平衡分类等相关领域具有丰富的研究经验和技术创新。团队成员曾参与多项国家级科研项目，在顶级会议和期刊发表相关论文，具备扎实的理论功底和工程实践能力。硬件方面，实验室配备了高性能

GPU 集群，能够满足大规模深度学习模型的训练需求。软件环境方面，已搭建完善的深度学习开发平台，支持 PyTorch、TensorFlow 等主流框架。此外，团队与国内外知名研究机构保持密切合作，能够及时获取最新研究进展和技术支持。这些条件为研究的顺利开展提供了有力保障。

3. 本研究的特色与创新之处

(1) 针对单一时序方法的根本局限，本研究首次提出将 Neural SDE 的动力学建模与 Transformer 的全局依赖捕获深度融合。通过线性神经随机微分方程的解析解形式避免数值误差累积，确保长时间间隔的精确状态转移；同时引入连续时间注意力机制，通过重参数化技术既保留并行计算优势，又实现任意时间点的精确依赖建模。这种融合架构既满足了物理系统的连续性约束，又捕获了复杂的时序依赖关系，从根本上解决了不规则时序建模中动力学表达与全局依赖之间的矛盾。

(2) 本研究创新性地提出混合重采样与类别感知分组注意力(CGA)的协同优化方案。在数据层采用 SVMSMOTE 和 Repeated ENN 混合策略，为模型提供平衡的训练信号；在模型层设计 CGA 机制，通过类别特定的注意力通道增强少数类表达，同时利用语义相似度驱动的类间交互避免信息隔离。这种双层协同既解决了单纯重采样的合成样本偏差问题，又克服了单纯 CGA 在极度不平衡下的学习困难，实现了从数据增强到特征学习的全流程优化。

(3) 本研究提出先建模、后平衡的两阶段串联架构，将连续时间建模与类别平衡处理解耦处理。这种架构设计相比现有方法，避免了直接在原始不规则数据上操作可能破坏时间依赖性的问题，不仅保证了时序特征的完整性，还使得各模块可以独立优化，显著提高了方法的灵活性和可扩展性，实现了两者的融合与协同优化。

4. 预期的论文进展和成果：

本研究的预期研究成果主要为以下两点：

(1) 基于神经随机微分方程的不规则时序连续建模方法
提出融合线性神经 SDE 和连续时间 Transformer 的创新框架，实现对不规则采样时间序列的精确建模。该方法通过解析解形式避免数值积分误差，在长时间间隔预测中表现优异。

(2) 面向类别不平衡的双路径分类架构
设计数据层混合重采样与模型层类别感知注意力协同的解决方案，有效提升少数类识别性能，同时保持整体分类准确率。

以上两项研究成果计划整理成论文，拟投稿至机器学习或时间序列分析领域的国际会议或期刊。

目前本研究已取得初步进展:

- 完成了连续时间模块的算法实现和部分混合重采样实现
- 在 ASAS、LINEAR、MACHO 天文观测数据集上进行了实验验证
- 初步实验结果显示，相比现有方法在 F1-score 等关键指标上有明显提升
- 论文撰写工作正在进行中，预计近期完成初稿

5. 论文预期研究成果（可考核的工作量）、年度研究计划、可能遇到的困难和问题及相应的解决办法:

5.1 论文预期研究成果

SCI 期刊论文或 CCF 期刊/会议论文 1-2 篇。

5.2 年度研究计划

2024.09 ~ 2024.12 深入调研不规则时间序列分析和类别不平衡学习的相关文献，重点关注神经随机微分方程、连续时间建模等前沿方法。搭建实验环境。

2025.01 ~ 2025.05 实现完整的模型框架，包括线性神经 SDE 模块、连续时间注意力机制和类别感知分组注意力。在不规则时序数据集上进行初步实验，调试优化模型性能。

2025.06 ~ 2025.12 完成全面的对比实验和消融研究，分析各模块的贡献度。整理实验结果，撰写学术论文初稿，准备投稿材料。同时开始撰写毕业论文。

2026.01-2026.06 整理各阶段研究成果，完成毕业论文。

5.2 可能遇到的困难和问题及相应的解决办法

困难：在研究过程中可能因为连续时间 Transformer 模型参数量大，在长序列数据上训练时出现显存不足的情况；其次，极度类别不平衡场景下模型收敛困难，少数类样本的学习效果不稳定。

解决办法：通过梯度累积、混合精度训练等技术降低显存占用，必要时采用模型并行策略；设计动态调整的类别权重和自适应采样策略，确保各类别都能得到充分学习。

三、论文提纲:

题目：基于神经随机微分方程的不规则采样时间序列分类方法研究

摘要

1 绪论

- 1.1 研究背景与意义
- 1.2 研究目的与挑战
- 1.3 论文的研究内容
- 1.4 本文的组织结构

2 国内外研究现状

- 2.1 离散时间建模方法
- 2.2 连续时间建模方法
- 2.3 时间序列不平衡分类方法
- 2.4 本章小结

3 基于 Neural SDE 和连续时间 Transformer 的不规则时序连续时间建模

- 3.1 引言
- 3.2 线性神经随机微分方程框架
- 3.3 连续时间 Transformer 设计
- 3.4 实验设置
- 3.5 实验结果与分析
- 3.6 本章小结

4 基于混合重采样和改进 CGA 的类别不平衡时序分类

- 4.1 引言
- 4.2 时序感知的混合重采样策略
- 4.3 类别感知分组注意力机制
- 4.4 实验设置
- 4.5 实验结果与分析
- 4.6 本章小结

5 结论与展望

- 5.1 结论
- 5.2 展望

参考文献

学术成果

致谢

主要参考文献目录:

- [1] Feigelson ED, Babu GJ, Caceres GA. Autoregressive times series methods for time domain astronomy. *Frontiers in Physics*. 2018; 6:80.
- [2] Zeger SL, Irizarry R, Peng RD. On time series analysis of public health and biomedical data. *Annual Review of Public Health*. 2006; 27:57-79.
- [3] Chen B, Liu Y, Zhang C, Wang Z. Time series data for equipment reliability analysis with deep learning. *IEEE Access*. 2020; 8:103337-103347.
- [4] Lai G, Chang WC, Yang Y, Liu H. Modeling long-and short-term temporal patterns with deep neural networks. In *The 41st international ACM SIGIR conference on research & development in information retrieval*. 2018: 95-104.
- [5] Rubanova Y, Chen RTQ, Duvenaud DK. Latent ordinary differential equations for irregularly-sampled time series. *Advances in neural information processing systems*. 2019; 32.
- [6] Thiebaut C, Roques S. Time-scale and time-frequency analyses of irregularly sampled astronomical time series. *EURASIP Journal on Advances in Signal Processing*. 2005.
- [7] Chen Y, Ren K, Wang Y, et al. Data discretization unification: A new approach for irregular time series analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2023; 45(3): 3456-3470.
- [8] Oh YK, Lim DY, Kim S. Stable neural stochastic differential equations in analyzing irregular time series data. *arXiv preprint arXiv:2402.14989*, 2024.
- [9] Chawla NV, Bowyer KW, Hall LO, et al. SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research*. 2002; 16: 321-357.
- [10] Cao W, Wang D, Li J, Zhou H, Li L, Li Y. BRITS: Bidirectional recurrent imputation for time series. *Advances in neural information processing systems*. 2018; 31.
- [11] Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.
- [12] Hyndman, R., Koehler, A. B., Ord, J. K., & Snyder, R. D. (2008). *Forecasting with exponential smoothing: the state space approach*. Springer Science & Business Media.
- [13] Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2), 257-286.
- [14] Berndt, D. J., & Clifford, J. (1994). Using dynamic time warping to find patterns in time series. *KDD workshop*, 10(16), 359-370.
- [15] Cao, L. J., & Tay, F. E. H. (2003). Support vector machine with adaptive parameters in financial time series forecasting. *IEEE Transactions on Neural Networks*, 14(6), 1506-1518.

- [16] Percival, D. B., & Walden, A. T. (2000). Wavelet methods for time series analysis. Cambridge University Press
- [17] Wu, Z., & Huang, N. E. (2009). Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Advances in adaptive data analysis*, 1(01), 1-41.
- [18] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735-1780.
- [19] Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint arXiv:1406.1078.
- [20] Wang, Z., Yan, W., & Oates, T. (2017). Time series classification from scratch with deep neural networks: A strong baseline. 2017 International Joint Conference on Neural Networks (IJCNN), 1578-1585.
- [21] Bai, S., Kolter, J. Z., & Koltun, V. (2018). An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. arXiv preprint arXiv:1803.01271.
- [22] Yu, B., Yin, H., & Zhu, Z. (2018). Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting. Proceedings of the 27th International Joint Conference on Artificial Intelligence, 3634-3640.
- [23] Li, Y., Yu, R., Shahabi, C., & Liu, Y. (2018). Diffusion convolutional recurrent neural network: Data-driven traffic forecasting. International Conference on Learning Representations (ICLR).
- [24] Zhao, L., Song, Y., Zhang, C., Liu, Y., Wang, P., Lin, T., Deng, M., & Li, H. (2020). T-GCN: A temporal graph convolutional network for traffic prediction. *IEEE Transactions on Intelligent Transportation Systems*, 21(9), 3848-3858.
- [25] Guo, S., Lin, Y., Feng, N., Song, C., & Wan, H. (2019). Attention based spatial-temporal graph convolutional networks for traffic flow forecasting. Proceedings of the AAAI conference on artificial intelligence, 33(01), 922-929.
- [26] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.
- [27] Zhou, H., Zhang, S., Peng, J., Zhang, S., Li, J., Xiong, H., & Zhang, W. (2021). Informer: Beyond efficient transformer for long sequence time-series forecasting. Proceedings of the AAAI Conference on Artificial Intelligence, 35(12), 11106-11115.
- [28] Wu, H., Xu, J., Wang, J., & Long, M. (2021). Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting. *Advances in Neural Information Processing Systems*, 34, 22419-22430.
- [29] Li, S., Jin, X., Xuan, Y., Zhou, X., Chen, W., Wang, Y. X., & Yan, X. (2019). Enhancing the locality and breaking the memory bottleneck of transformer on time series forecasting. *Advances in Neural Information Processing Systems*, 32.

- [30] Li, R., Plested, J., Kanahasabai, G., Robles-Kelly, A., & Rashidi, P. (2022). MTANS: Multi-task attention networks for medical time series classification. *Journal of Biomedical Informatics*, 130, 104086.
- [31] Lim, B., Leek, K., Kim, J., & van der Schaar, M. (2021). Temporal fusion transformers for interpretable multi-horizon time series forecasting. *International Journal of Forecasting*, 37(4), 1748-1764.
- [32] Wang, C., Liang, K., Xu, P., & Li, Y. (2022). PRIME: A few primitives can boost robustness to common corruptions. *arXiv preprint arXiv:2112.13547*.
- [33] Wen, Q., Zhou, T., Zhang, C., Chen, W., Ma, Z., Yan, J., & Sun, L. (2022). Transformers in time series: A survey. *arXiv preprint arXiv:2202.07125*.
[21] Chen Y, Ren K, Wang Y, et al. Data discretization unification: A new approach for irregular time series analysis[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, 45(3): 3456-3470.
- [34] Chen Y, Ren K, Wang Y, et al. Data discretization unification: A new approach for irregular time series analysis[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, 45(3): 3456-3470.
- [35] Shumway R H, Stoffer D S. An approach to time series smoothing and forecasting using the EM algorithm[J]. *Journal of time series analysis*, 1982, 3(4): 253-264.
- [36] Fox E, Sudderth E B, Jordan M I, et al. Bayesian nonparametric inference of switching dynamic linear models[J]. *IEEE Transactions on Signal Processing*, 2011, 59(4): 1569-1585.
- [37] Frigola R, Lindsten F, Schön T B, et al. Bayesian inference and learning in Gaussian process state-space models with particle MCMC[C]. *Advances in Neural Information Processing Systems*, 2013, 26.
- [38] Bonilla E V, Chai K M, Williams C. Multi-task Gaussian process prediction[C]. *Advances in neural information processing systems*, 2007, 20.
- [39] Nodelman U, Shelton C R, Koller D. Continuous time Bayesian networks[C]. *Proceedings of the Eighteenth conference on Uncertainty in artificial intelligence*, 2002: 378-387.
- [40] Yu S Z. Hidden semi-Markov models[J]. *Artificial intelligence*, 2010, 174(2): 215-243.
- [41] Särkkä S. Recursive Bayesian inference on stochastic differential equations[D]. Helsinki University of Technology, 2006.
- [42] Beer R D. On the dynamics of small continuous-time recurrent neural networks[J]. *Adaptive Behavior*, 1995, 3(4): 469-509.
- [43] Rubanova Y, Chen R T Q, Duvenaud D K. Latent ordinary differential equations for irregularly-sampled time series[C]. *Proceedings of the 33rd International Conference on Neural Information Processing Systems*, 2019: 5320-5330.
- [44] Chen R T Q, Rubanova Y, Bettencourt J, et al. Neural ordinary differential equations[C]. *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, 2018: 6572-6583.

- [45] Rubanova Y, Chen R T Q, Duvenaud D K. Latent ordinary differential equations for irregularly-sampled time series[J]. Advances in neural information processing systems, 2019, 32.
- [46] Mozer M C, Kazakov D, Lindsey R V. Discrete event, continuous time RNNs[J]. arXiv preprint arXiv:1710.04110, 2017.
- [47] Neil D, Pfeiffer M, Liu S C. Phased LSTM: Accelerating recurrent network training for long or event-based sequences[C]. Advances in neural information processing systems, 2016: 3882-3890.
- [48] Mei H, Eisner J M. The neural hawkes process: A neurally self-modulating multivariate point process[J]. Advances in neural information processing systems, 2017, 30.
- [49] Kidger P, Morrill J, Foster J, et al. Neural controlled differential equations for irregular time series[J]. Advances in Neural Information Processing Systems, 2020, 33: 6696-6707.
- [50] Jhin S H, Shin H, Hong S, et al. Attentive neural controlled differential equations for time-series classification and forecasting[C]. 2021 IEEE International Conference on Data Mining (ICDM), 2021: 250-259.
- [51] Morrill J, Salvi C, Kidger P, et al. Neural rough differential equations for long time series[C]. International Conference on Machine Learning, 2021: 7829-7838.
- [52] Oh Y K, Lim D Y, Kim S. Stable neural stochastic differential equations in analyzing irregular time series data[J]. arXiv preprint arXiv:2402.14989, 2024.
- [53] Schirmer M, Eltayeb M, Lessmann S, et al. Modeling irregular time series with continuous recurrent units[C]. International conference on machine learning. PMLR, 2022: 19388-19405.
- [54] Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research, 16, 321-357.
- [55] He, H., Bai, Y., Garcia, E. A., & Li, S. (2008). ADASYN: Adaptive Synthetic Sampling Approach for Imbalanced Learning. IEEE International Joint Conference on Neural Networks, 1322-1328.
- [56] Jo, T., & Japkowicz, N. (2004). Class Imbalances versus Small Disjuncts. ACM SIGKDD Explorations Newsletter, 6(1), 40-49.
- [57] Nguyen, H. M., Cooper, E. W., & Kamei, K. (2011). Borderline Over-sampling for Imbalanced Data Classification. International Journal of Knowledge Engineering and Soft Data Paradigms, 3(1), 4-21.
- [58] Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., & Bharath, A. A. (2018). Generative Adversarial Networks: An Overview. IEEE Signal Processing Magazine, 35(1), 53-65.
- [59] Liu, X. Y., Wu, J., & Zhou, Z. H. (2008). Exploratory Undersampling for Class-Imbalance Learning. IEEE Transactions on Systems, Man, and Cybernetics, Part B, 39(2), 539-550.

- [60] Seiffert, C., Khoshgoftaar, T. M., Van Hulse, J., & Napolitano, A. (2010). RUSBoost: A Hybrid Approach to Alleviating Class Imbalance. *IEEE Transactions on Systems, Man, and Cybernetics, Part A*, 40(1), 185-197.
- [61] Menardi, G., & Torelli, N. (2014). Training and Assessing Classification Rules with Imbalanced Data. *Data Mining and Knowledge Discovery*, 28(1), 92-122.
- [62] Lin, S., Yu, P., & Xu, Z. (2018). Dynamic Resampling for Class Imbalance Learning. *IEEE International Joint Conference on Neural Networks*, 1-8.
- [63] Elkan, C. (2001). The Foundations of Cost-Sensitive Learning. *International Joint Conference on Artificial Intelligence*, 17, 973-978.
- [64] Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal Loss for Dense Object Detection. *IEEE International Conference on Computer Vision*, 2980-2988.
- [65] Cui, Y., Jia, M., Lin, T. Y., Song, Y., & Belongie, S. (2019). Class-Balanced Loss Based on Effective Number of Samples. *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 9268-9277.
- [66] Cao, K., Wei, C., Gaidon, A., Arechiga, N., & Ma, T. (2019). Learning Imbalanced Datasets with Label-Distribution-Aware Margin Loss. *Advances in Neural Information Processing Systems*, 32.
- [67] Wu, T., Huang, Q., Liu, Z., Wang, Y., & Lin, D. (2020). Distribution-Balanced Loss for Multi-Label Classification in Long-Tailed Datasets. *European Conference on Computer Vision*, 162-178.
- [68] Cui, Y., Jia, M., Lin, T. Y., Song, Y., & Belongie, S. (2019). Class-Balanced Cross Entropy for Long-Tailed Classification. *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 9268-9277.
- [69] Chawla, N. V., Lazarevic, A., Hall, L. O., & Bowyer, K. W. (2003). SMOTEBoost: Improving Prediction of the Minority Class in Boosting. *European Conference on Principles of Data Mining and Knowledge Discovery*, 107-119.
- [70] Zhang, X., Fang, Z., Wen, Y., Li, Z., & Qiao, Y. (2017). Range Loss for Deep Face Recognition with Long-Tailed Training Data. *IEEE International Conference on Computer Vision*, 5409-5418.
- [71] Ren, M., Zeng, W., Yang, B., & Urtasun, R. (2018). Learning to Reweight Examples for Robust Deep Learning. *International Conference on Machine Learning*, 4334-4343.
- [72] Liu, B., Zhu, X., & Zhang, L. (2019). MetaSAug: Meta Semantic Augmentation for Long-Tailed Visual Recognition. *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 5212-5221.
- [73] Liu, Z., Miao, Z., Zhan, X., Wang, J., Gong, B., & Yu, S. X. (2019). Large-Scale Long-Tailed Recognition in an Open World. *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2537-2546.

- [74] Huang, C., Li, Y., Chen, C. L., & Tang, X. (2016). Deep Imbalanced Learning for Face Recognition and Attribute Prediction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(11), 2781-2794.
- [75] Yang, Y., & Xu, Z. (2020). Rethinking the Value of Labels for Improving Class-Imbalanced Learning. *Advances in Neural Information Processing Systems*, 33, 19290-19301.
- [76] Ren, J., Yu, C., Sheng, S., Ma, X., Zhao, H., Yi, S., & Li, H. (2020). Balanced Meta-Sigmoid for Long-Tailed Visual Recognition. *Advances in Neural Information Processing Systems*, 33, 4175-4186.

四、导师意见

导师签字:

年 月 日

备注: 研究生在撰写开题报告时, 须有专门的部分或段落, 在导师的指导下对自己在论文研究过程中可能会面临的实验室安全、生物安全、网络安全、技术安全、数据安全等各类安全问题进行评估, 并采取相应的防范措施和做好应急预案。

研究生在进行开题答辩时, 导师和考核小组须对研究生在开题报告中所列出的安全问题是否全面、采取的防范措施是否到位、编制的应急预案是否可行等方面进行充分的评估和论证, 并对不足之处提出改进和补充建议, 帮助研究生把好论文研究安全关。