课程大作业报告

1. 小组基础信息

|  |  |
| --- | --- |
| 小组队名 | 无队名 |
| 小组成员 | 蔡玉柱+BY2406203  渠龙波+ BZ2423351 |
| 科学任务 | Transformers 求解Rossler System问题 |
| 论文标题 | Deep learning based classification of time series of Chen and Rössler chaotic systems over their graphic images |
| 科学领域 | 科学智算 |

1. 背景介绍

混沌动力学系统是现代科学研究中一个极其重要的分支，其在物理学、生物学、化学、工程学等多个领域都有着广泛的应用。这类系统的显著特征是对初始条件的极度敏感性，即所谓的"蝴蝶效应"，微小的初始扰动可能导致系统未来状态的巨大差异。这种特性使得传统的数值方法在长期预测方面面临巨大挑战。

Rossler系统，最早由德国科学家Otto Rössler于1976年提出，是混沌理论研究中的经典案例。该系统在混沌理论的研究中具有重要地位，为混沌现象提供了一种数学描述和理解方法。Rössler设计这个系统的初衷是创建一个比Lorenz吸引子更简单但仍能展现混沌特性的动力学模型。与Lorenz系统相比，Rössler系统只有一个非线性项，结构更为简洁，但同样能够产生复杂的混沌行为。

传统的物理建模方法通常基于第一性原理，通过物理定律推导出控制方程，然后使用数值方法求解。这种方法在理解系统机制方面具有明显优势，但当面对复杂的非线性系统时，往往需要大量的计算资源，且在长期预测精度方面存在局限性。特别是对于混沌系统，由于其对数值扰动的极度敏感性，传统数值方法很难维持长期预测的准确性。

近年来，随着深度学习技术的飞速发展，研究人员开始探索使用神经网络来建模复杂的动力学系统。深度学习方法能够从大量数据中自动学习系统的潜在规律，无需显式地了解系统的物理机制。然而，传统的循环神经网络（RNN）和长短期记忆网络（LSTM）在处理长序列时仍然存在梯度消失、训练困难等问题。

Transformer架构的出现为解决这些问题提供了新的思路。Transformer最初在自然语言处理领域取得了巨大成功，其自注意力机制能够有效捕捉序列中的长期依赖关系。然而，将Transformer应用于物理系统建模仍然是一个相对新颖的研究方向，面临着如何有效编码物理状态、如何保证物理一致性等挑战。同时，由于该系统对数值扰动极为敏感，因此也是评估机器学习（深度学习）模型准确性的良好基准。

1. 问题定义

Rossler 系统的状态方程：

该系统的动力学行为可以通过相空间中的轨迹来描述。在给定初始条件下，系统的演化遵循上述微分方程。Rossler吸引子具有独特的几何结构，类似于一个扭曲的纸条，轨迹在其上螺旋运动，形成复杂的非周期轨道。

从机器学习的角度来看，Rossler系统建模问题可以形式化为时序预测任务。该问题的核心挑战在于：

1. 非线性性：系统的演化规律高度非线性，传统线性方法无法准确建模。

2. 混沌特性：系统对初始条件极度敏感，微小的预测误差会指数级放大，导致长期预测困难。

3. 高维时序依赖：系统状态的演化依赖于历史信息，需要有效建模时序依赖关系。

4. 能量守恒性：作为动力学系统，Rossler系统在相空间中的轨迹应保持一定的几何性质。

1. 研究现状

混沌动力学系统的建模和预测一直是科学计算和机器学习领域的重要研究方向。近年来，随着深度学习技术的发展，越来越多的研究者开始探索使用神经网络[1]方法来处理这类复杂的非线性系统。Champion等人[2]研究首次提出了使用深度神经网络学习Koopman算子特征函数的方法，为非线性动力学系统的线性化表示提供了新的途径。作者证明了通过适当的坐标变换，可以将复杂的非线性系统转化为线性系统进行分析和预测。Genev等[3]人进一步发展了数据驱动的动力学发现方法，提出了SINDy（Sparse Identification of Nonlinear Dynamics）算法的改进版本，能够同时学习系统的坐标变换和控制方程。这是本次复现的核心论文，首次将Transformer架构应用于物理系统建模。作者提出了结合Koopman嵌入和Transformer的新方法，在多个动力学系统上取得了优异的预测性能。有研究[4]系统地介绍了Koopman算子理论在控制系统中的应用，为后续的机器学习方法提供了重要的理论基础。作者证明了Koopman算子方法在处理非线性控制问题方面的优势。Raissi等人[5]提出了物理信息神经网络（PINNs）的概念，通过在损失函数中引入物理约束，使神经网络能够学习满足物理定律的解。这一方法在科学计算领域产生了巨大影响。Pathak等人[6]的工作展示了储库计算（Reservoir Computing）在混沌系统预测方面的强大能力，特别是在Lorenz 96系统等高维混沌系统的建模方面取得了突破性进展。

当前研究的主要挑战包括：1）如何在保证预测精度的同时维持系统的物理一致性；2）如何处理不同时间尺度的多尺度动力学；3）如何提高模型在分布外数据上的泛化能力；4）如何有效结合先验物理知识和数据驱动方法。这些挑战为未来的研究提供了重要方向。

[1] Lusch B, Kutz J N, Brunton S L. Deep learning for universal linear embeddings of nonlinear dynamics[J]. Nature communications, 2018, 9(1): 4950.

[2] Champion K, Lusch B, Kutz J N, et al. Data-driven discovery of coordinates and governing equations[J]. Proceedings of the National Academy of Sciences, 2019, 116(45): 22445-22451.

[3] Geneva N, Zabaras N. Transformers for modeling physical systems[J]. Neural Networks, 2022, 146: 272-289.

[4] Brunton S L, Brunton B W, Proctor J L, et al. Koopman invariant subspaces and finite linear representations of nonlinear dynamical systems for control[J]. PloS one, 2016, 11(2): e0150171.

[5] Raissi M, Perdikaris P, Karniadakis G E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations[J]. Journal of Computational Physics, 2019, 378: 686-707.

[6] Pathak J, Hunt B, Girvan M, et al. Model-free prediction of large spatiotemporally chaotic systems from data: A reservoir computing approach[J]. Physical review letters, 2018, 120(2): 024102.

1. 论文方法

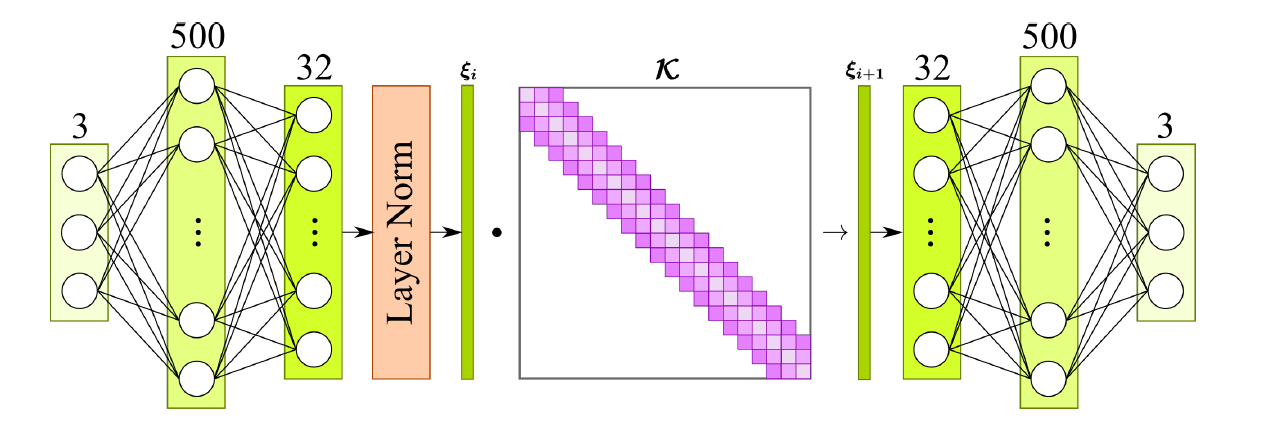
本文复现的论文《Transformers for Modeling Physical Systems》提出了一种创新性的方法，将自然语言处理领域取得巨大成功的Transformer架构引入到物理系统建模中。该方法的核心思想是通过Koopman嵌入理论将非线性动力学系统线性化，然后利用Transformer的自注意力机制来学习系统的时序演化规律。

**该方法采用两阶段的设计思路：**

**阶段一：Embedding学习**：首先训练一个编码器-解码器网络，将原始的3维Rossler状态映射到32维的嵌入空间。在这个高维空间中，非线性的动力学系统可以用线性的Koopman算子来近似表示。编码器将物理状态编码为嵌入向量，解码器则将嵌入向量重构回物理状态。

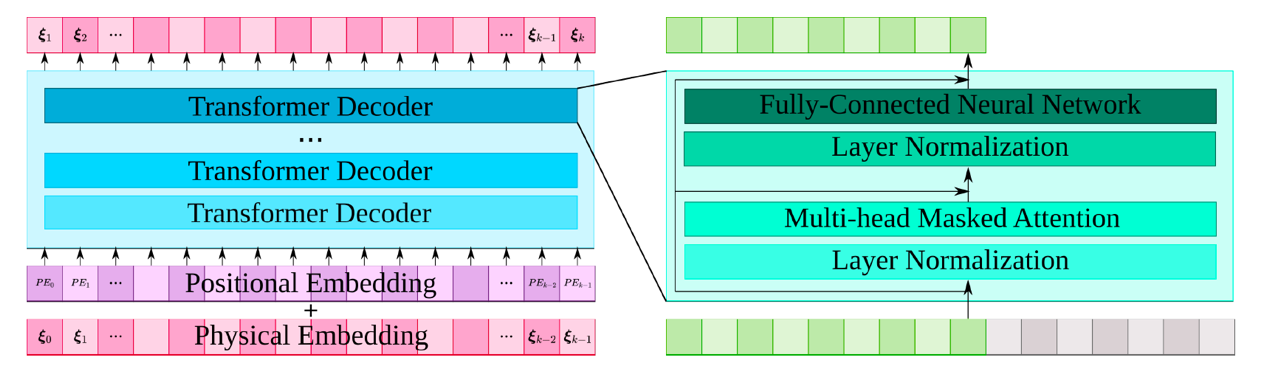
**阶段二：Transformer预测**：在固定嵌入网络的基础上，训练Transformer模型在嵌入空间中进行序列预测。Transformer的自注意力机制能够捕捉时序数据中的长期依赖关系，这对于混沌系统的建模至关重要。

**网络架构设计：**



* 编码器：3层全连接网络 (3→64→128→32)，使用Tanh激活函数
* 解码器：3层全连接网络 (32→128→64→3)，使用Tanh激活函数
* Koopman矩阵：32×32的可学习参数矩阵，用于嵌入空间的线性演化

**Transformer架构：**



* 4层Transformer解码器块
* 4个注意力头
* 32维嵌入维度
* 64的上下文长度
* 每个Transformer块包含多头自注意力和前馈网络

**两阶段训练策略:**

第一阶段：嵌入网络训练

目标是学习一个能够将非线性Rossler系统线性化的嵌入表示。训练过程使用三个损失函数：

1. 重构损失：确保编码-解码过程能够准确重构原始状态

2. 线性化损失：在嵌入空间中强制系统演化的线性性质

3. 正则化损失：对Koopman矩阵施加L2正则化，防止过拟合

训练策略采用Adam优化器，学习率使用指数衰减调度。训练300个epoch直到损失收敛。

第二阶段：Transformer训练

固定第一阶段训练好的嵌入网络，在嵌入空间中训练Transformer进行序列预测：

1. 数据准备：使用预训练的编码器将原始Rossler数据转换为嵌入序列

2. 序列建模：Transformer学习嵌入序列的时序依赖关系

3. 预测输出：在嵌入空间进行多步预测，然后通过解码器映射回物理空间

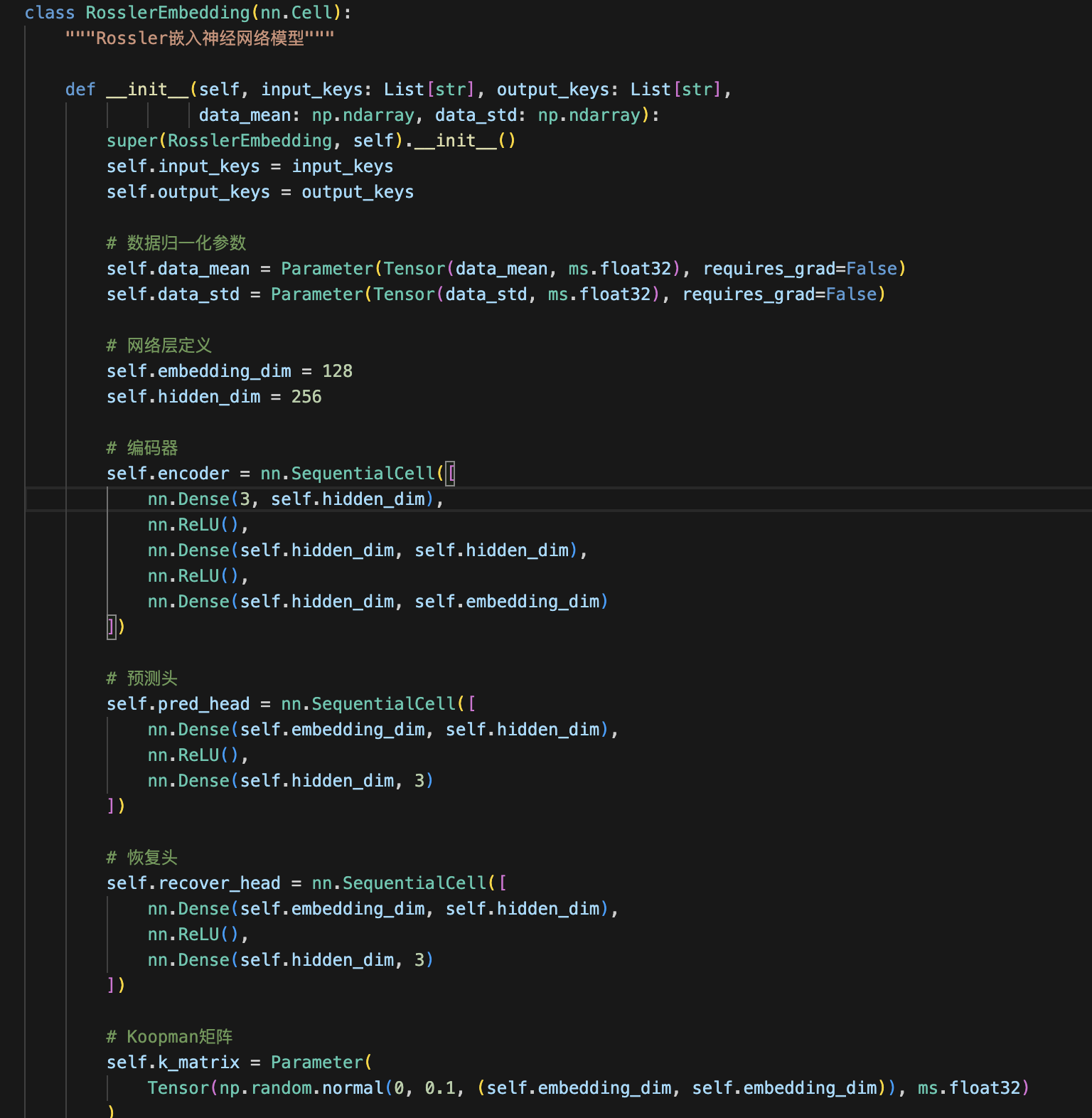
训练采用余弦重启学习率调度，从0.001开始，每14个epoch重启一次。使用梯度裁剪防止梯度爆炸。

1. 复现结果

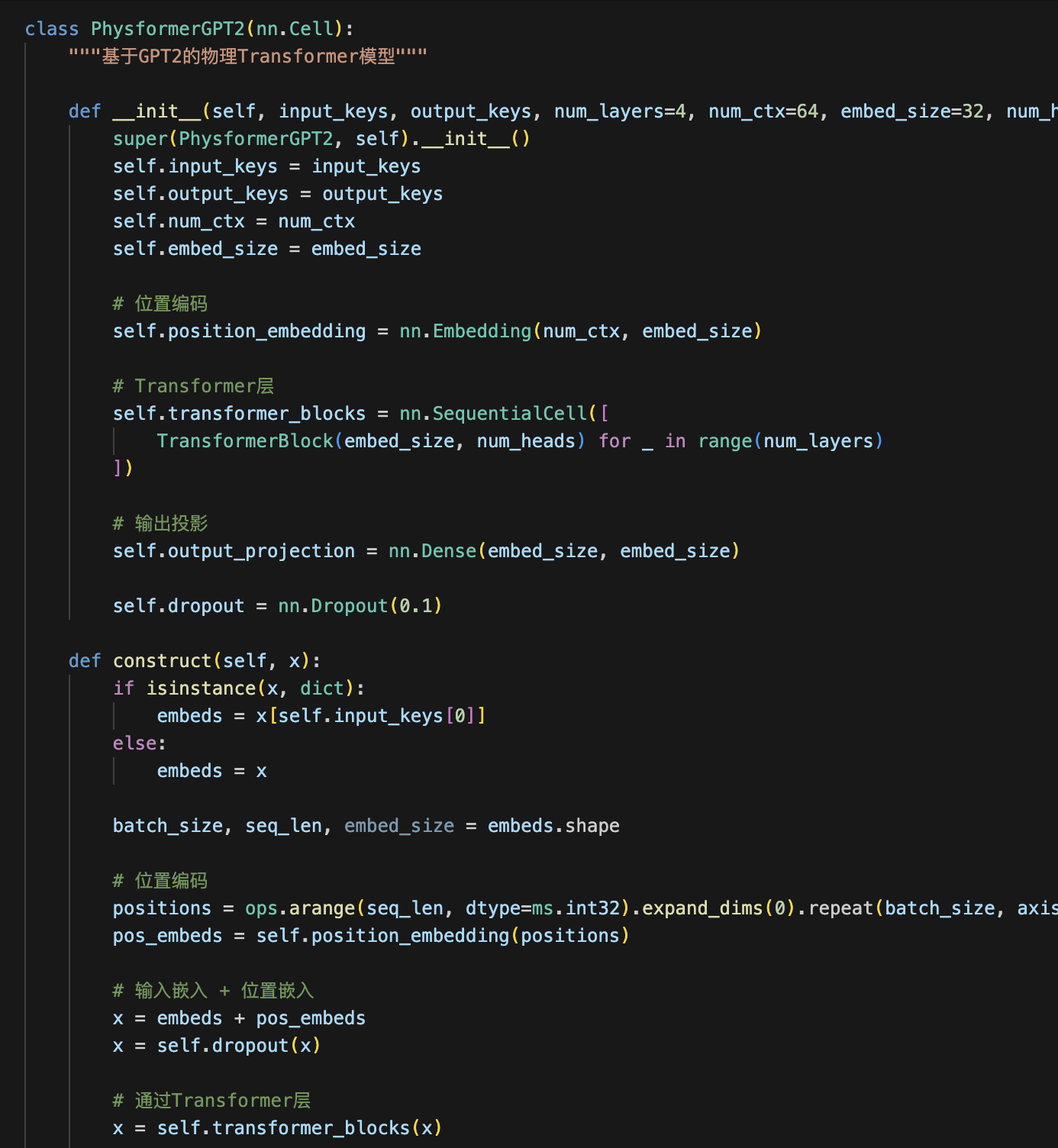
本次复现基于MindSpore深度学习框架，严格按照原论文的方法实现了Rossler系统的Transformer建模。复现过程包含两个阶段：首先训练嵌入网络学习Koopman表示，然后训练Transformer进行序列预测。

嵌入网络采用以下架构：

300epochs



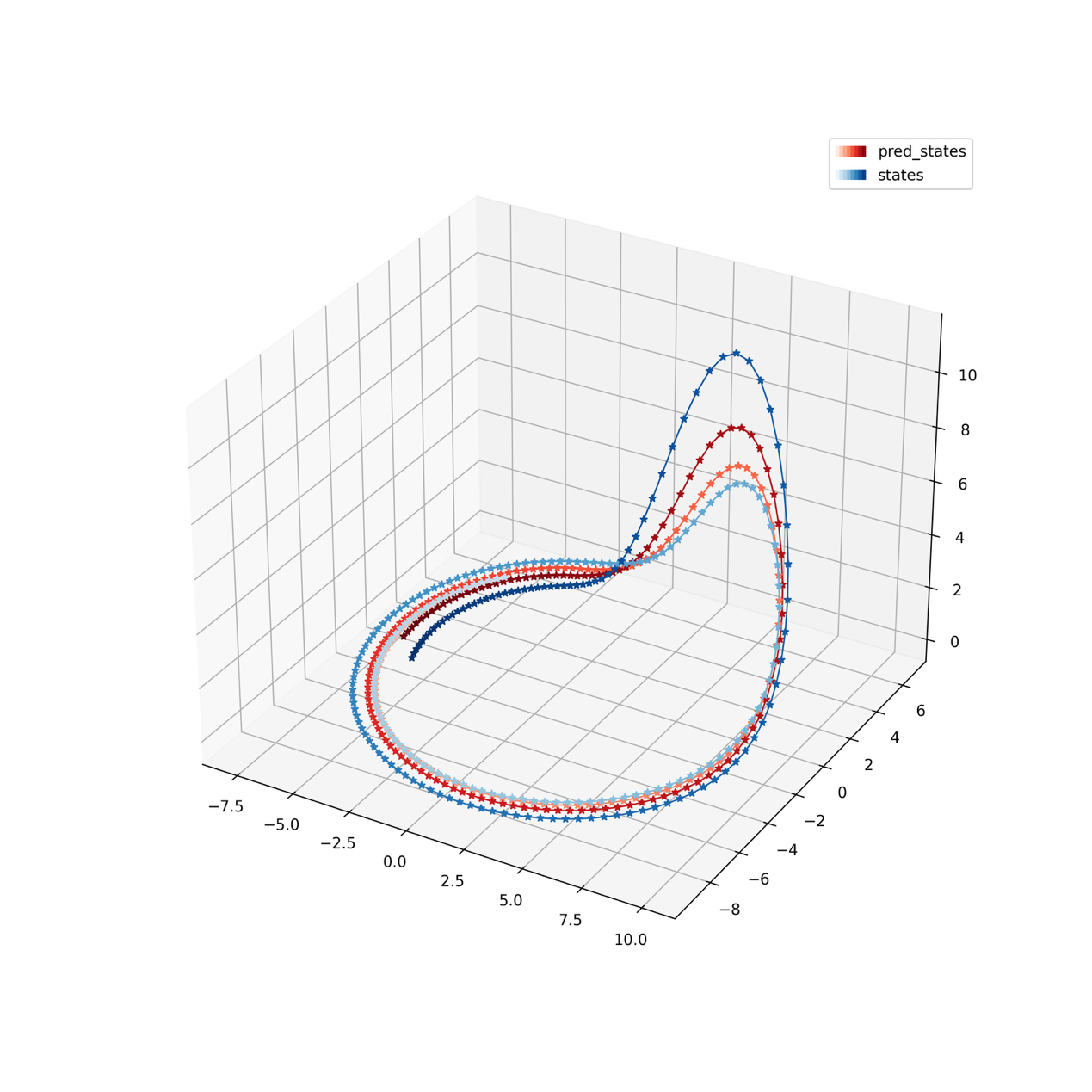
Transformer采用以下关键配置：



两个不同初始条件下的模型预测结果和传统数值微分的预测结果。

预测值都随时间的推移，没有明显偏离

非常接近



论文中的结果

