

## 附件 2.《Python 语言程序设计》课程期末考核任务说明

### 1 任务介绍

流程工业控制系统时序数据预测任务是一项多变量时间序列预测任务，其中多变量数据源自多个传感器，每个传感器的输出是一个时间序列数据，这些传感器数据来自于控制系统中的多个控制回路。

在工业控制系统中，一个控制回路通常包括一个过程变量（Process Variable, PV）和一个控制变量（Operating Variable, OP）。其中，过程变量是指工业过程或系统中需要监测和控制的实际物理量或状态。可以是温度、压力、流量、速度、液位等，是通过传感器测量并反馈给控制系统的值，用于表示当前的过程状态。控制变量是能够影响工业过程或系统行为的可控变量，控制器通过与设定点进行比较，并计算出误差，然后生成相应的控制信号来调整执行器（针对流量作为过程变量的回路，控制变量可以是阀门开度），以使过程变量达到设定点。干扰变量（Disturbance Variable, DV）是工业过程中，会影响过程变量但不能直接控制的变量。它们是工业过程中的外部因素或变化，可以导致过程变量偏离设定点。

通过对流程工业控制系统时序数据进行预测，可以用于包括故障预警和控制系统优化等多种下游任务。

在流程工业多元时间序列预测任务中，通常只将过程变量作为预测目标。本任务要求在参考文献中的模型或其他模型的基础上进行改进，利用过去 30 个时间步的历史时间序列数据来预测未来 5 个时间步的过程变量数据。

可以作为输入的数据包括：

1. 历史 30 个时间步的过程变量、控制变量和干扰变量数据；
2. 未来 5 个时间步的控制变量数据\*（可选）。

要求的输出数据为：

1. 未来 5 个时间步的过程变量的预测值。

\*注意：如果使用了未来时间步的控制变量数据，模型必须以迭代单步预测的形式完成多步预测任务，并且每次单步预测只能使用该步之前的控制变量序列，因为虽然未来控制变量的值可以通过计算过程变量与设定值间的误差获得，但一次只能获得一步。

## 2 数据集介绍

本工业数据集来自于某厂甲醇精馏装置 DCS 系统数据。数据时间跨度约为 2 天，采样间隔为 10s，共采集 14517 个时间点数据。每个数据点都包括了 23 个传感器的读数，其中包括 9 个控制回路的过程变量、9 个对应的控制变量，以及 5 个干扰变量。

jl\_data\_train.csv 中，每一行代表一个时间点的采样数据，time\_stamp 列为时间戳，PV\_{i} 和 OP\_{i} 列表示第 i 个回路中对应的过程变量和控制变量。DV\_{1-5} 列是工段的干扰变量，在本数据集中，干扰变量不随回路划分。

jl\_adj\_mx.npy 中包含了 23 个传感器变量间的邻接矩阵（选用）。

## 3 完成路径

**路径 1.** 阅读参考文献[1]，修改其提供的开源代码，使其能够处理本任务所提供的数据集。（难度：★）

**路径 2.** 在路径 1 的基础上，尝试修改代码进行模型性能的改进。（难度：★★）

**路径 3.** 在路径 1 的基础上，将代码迁移到华为 MindSpore 环境下实现。（难度：★★）

**路径 4.** 在路径 1 的基础上，阅读参考文献[2]，并自己实现参考文献[2]所提出的模型。（难度：★★）

**路径 5.** 阅读参考文献[3]，修改其提供的开源代码，使其能够处理本任务所提供的数据集。（难度：★★）

**路径 6.** 在路径 5 的基础上，尝试修改代码进行模型性能的改进。（难度：★★★）

**路径 7.** 在路径 5 的基础上，将代码迁移到华为 MindSpore 环境下实现。（难度：★★★）

## 4 参考文献

[1] Shao, Z., Zhang, Z., Wang, F., Wei, W., & Xu, Y. (2022, October). Spatial-temporal identity: A simple yet effective baseline for multivariate time series forecasting. In *Proceedings of the*

31st ACM International Conference on Information & Knowledge Management (pp. 4454-4458).

开源代码: <https://github.com/zezhishao/STID>

[2] Wang, Z., Nie, Y., Sun, P., Nguyen, N. H., Mulvey, J., & Poor, H. V. (2023). ST-MLP: A Cascaded Spatio-Temporal Linear Framework with Channel-Independence Strategy for Traffic Forecasting. *arXiv preprint arXiv:2308.07496*.

[3] Shun-Yao, S., Fan-Keng, S., & Hung-yi, L. (2019). Temporal pattern attention for multivariate time series forecasting [J]. *Machine Learning*, 108(8-9). 开 源 代 码 : <https://github.com/Sanyam-Mehta/TPA-LSTM-PyTorch>