**MCW-RNN方法整理**

# 问题设定

本方法考虑场景为多速率软测量，具体的数据结构如下示意图。过程变量和质量变量中存在多种采样速率，通常质量变量的采样速率慢于过程变量的采样速率，示意图中过程变量中有三种采样速率，质量变量中有一种采样速率，共四种采样速率，采样速率分别为T、2T、4T、8T。

多速率场景下软测量的目标是基于多种快速率的过程变量对慢速率的质量变量进行预测。通常方法采用上采样或下采样的方式将数据整体转化为最快速率或最慢速率下的常规数据结构，进而使用常规建模方式即可。**上采样方式通常采用插值或数据填补方式，会造成误差累计，下采样方式通常使用数据抬升技巧，会导致输入维度较高，同时快速率过程变量占比会明显高于慢速率过程变量。**



**图1 多速率场景数据结构**

# 整体思路

借鉴了时钟循环神经网络（ClockWork RNN, CW-RNN）的思想，CW-RNN提出动机是为了解决RNN无法适用长期时序关系，**将每个隐含层神经元赋予更新周期，每个时间步只对满足更新周期的隐含层神经元进行更新，从而能够将长期和短期的时序关系分别保存在更新周期长和更新周期短的隐含层神经元中**。为了避免慢速率信息受快速率信息影响，在更新时禁止了快速率神经元对慢速率神经元的影响，这样同时能够起到减少参数量的效果。



**图2 CW-RNN思路**

在多速率软测量问题中，输入过程变量本身天然具有不同的采样速率，类似于CW-RNN中的更新周期，因此所提方法**将隐含层神经元根据过程变量的采样速率情况对应进行分组**，每个时间步只对采样到的过程变量对应的隐含层神经元进行更新，具体的更新策略如下图所示。***Wx*中只使用当前时刻采集到的过程变量以及需要更新的隐含层神经元对应的部分，*Wh*中只使用当前时刻需要更新的隐含层神经元对应的部分**。此处并不假设慢速率神经元不受快速率神经元影响，而是认为它们之间地位等同，不同速率神经元之间的耦合信息可以相互交互影响。



**图3 时钟更新策略**

# 实验结果

所提方法命名为MCW-RNN（Modified CW-RNN），同时对比了MLP（只使用含有完整变量的样本建立MLP）、I-MLP（采用插值填补后建立MLP）、L-MLP（采用数据抬升技巧后建立MLP）、RNN（将缺失数据视为0建立RNN）四种方法。

**表1 磨煤机案例三次平均实验结果**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 方法 | QV1 | | QV2 | |
| R2(%) | RMSE(℃) | R2(%) | RMSE(℃) |
| MLP | 54.92 | 1.089 | 86.91 | 0.389 |
| I-MLP | 57.47 | 1.056 | 73.35 | 0.554 |
| L-MLP | 38.78 | 1.269 | 84.91 | 0.417 |
| RNN | 55.32 | 1.084 | 86.41 | 0.396 |
| MCW-RNN | **61.35** | **1.008** | **90.59** | **0.329** |

在两个质量变量上的预测性能相比传统上采样和下采样方法均有明显提升。MLP和RNN性能相当，MLP由于未使用不完整样本数据，因此无法充分使用多速率数据，RNN由于每一时间步对所有隐含层神经元都更新，因此受到快速率变量的主导，慢速率变量信息被抑制。L-MLP同样由于数据抬升后快速率变量占主导，慢速率变量信息会被抑制。I-MLP通过数据填补能够增加样本量实现对QV1性能的提升，但由于数据填补会造成误差累计现象，从而使得错误信息可能会影响模型性能。

# 后续想法

1. 加入attention机制或者换成LSTM（小的改进）；
2. 每种速率的过程变量采用一个单独的LSTM提取时序特征，再将不同速率下提取的时序特征拼接用于预测（之前想过，但感觉思路比较简单）。