

# Real-time prediction of nuclear power plant parameter trends following operator actions

陈鸿煜

Update date: 2023/5/26

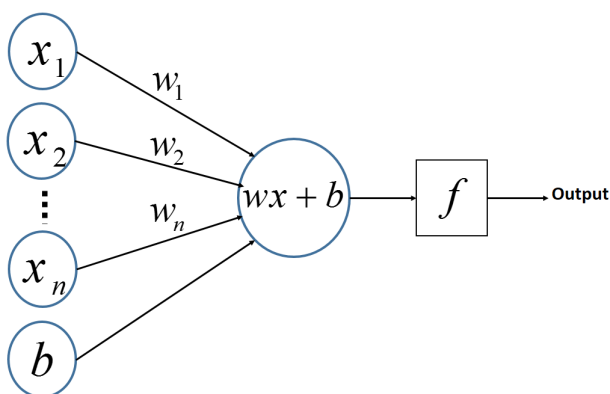
为了在核电站中实现快速准确的未来参数趋势预测，文章提出了一个**由多步预测策略和人工神经网络组成的数据驱动预测模型**(a data-driven prediction model composed of a multi-step prediction strategy and artificial neural networks)。

为了找到最佳模型性能，文章应用了多层感知机（Multi-Layer Perceptron, MLP）、循环神经网络和长期记忆（LSTM）网络，并用核电站模拟器生成的应急运行数据训练了各种候选模型。应用结果表明，具有**多输入多输出策略和LSTM网络的预测模型**能够成功地解决在多种紧急情况下考虑操作员行动的未來参数趋势估计的多变量问题。

## 一、前置知识

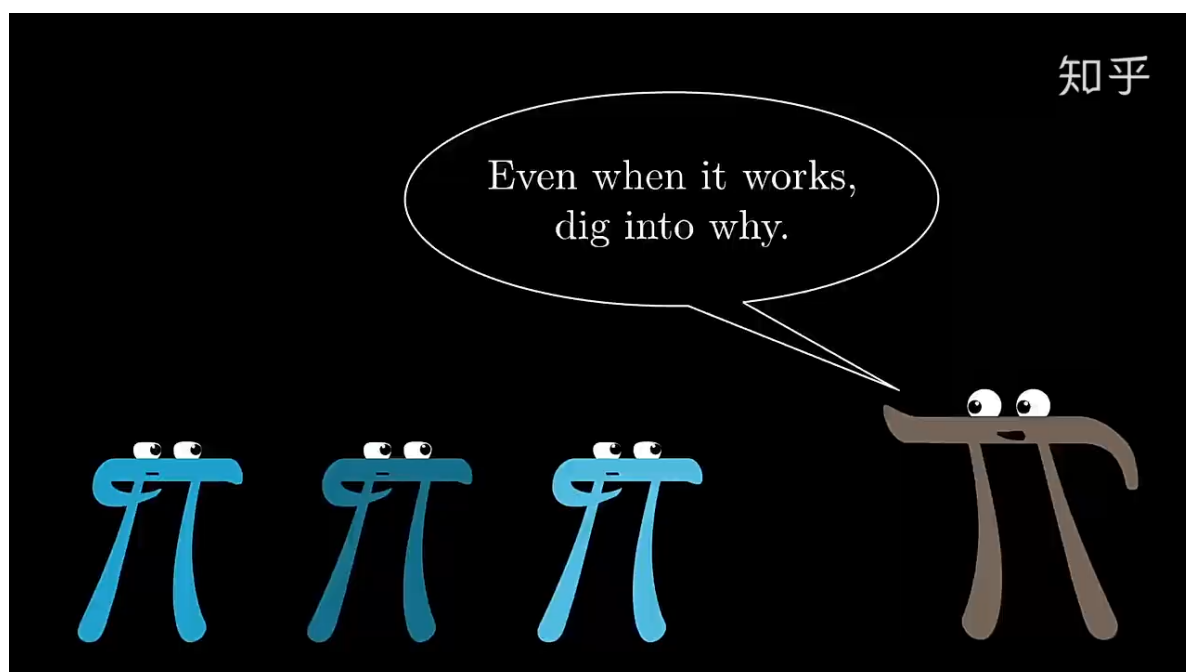
### 1. 多层感知机MLP

首先介绍感知机(Perceptron Learning Algorithm, PLA)，是一个线性的二分类器，但不能对非线性的数据并不能进行有效的分类。示意图如下所示：



<http://blog.csdn.net/xhoies>

单层只能进行二分类，为了进行多分类，就需要增加层数。

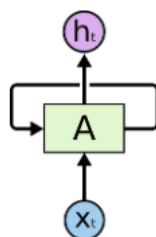


具体如何选择层数，以及如何进行梯度下降，需要进一步学习。

## 2. LSTM

### (1). 概念

LSTM是一种特殊的**递归神经网络**。这种网络与一般的前馈神经网络不同，LSTM可以利用时间序列对输入进行分析；简而言之，当使用前馈神经网络时，神经网络会认为 $t$ 时刻输入的内容与 $t + 1$ 时刻输入的内容**完全无关**，对于许多情况，例如图片分类识别，这是毫无问题的，可是对于一些情景，例如**自然语言处理** (NLP, Natural Language Processing) 或者我们需要分析类似于**连拍照片**这样的数据时，合理运用 $t$ 或之前的输入来处理 $t + n$ 时刻显然可以更加合理的运用输入的信息。为了运用到时间维度上信息，人们设计了**递归神经网络** (RNN, Recurssion Neural Network)，一个简单的递归神经网络如下图所示：



在图中， $x_t$  是在  $t$  时刻的输入信息， $h_t$  是在  $t$  时刻的输出信息，我们可以看到神经元  $A$  会递归的调用自身并且将  $t - 1$  时刻的信息传递给  $t$  时刻。递归神经网络在许多情况下运行良好，特别是在对**短时间序列**数据的分析时十分方便。但是，处理长时间序列数据时，性能会有所下降：

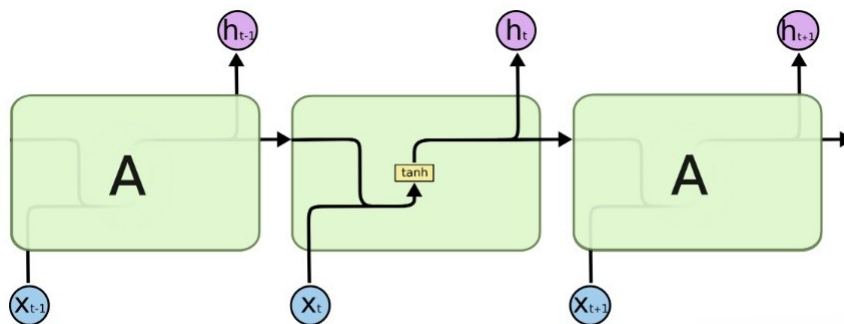
**长期依赖问题**：经过许多阶段传播后的梯度倾向于消失（大部分情况）或爆炸（很少，但对优化过程影响很大）。对于梯度爆炸是很好解决的，可以使用梯度修剪（**Gradient Clipping**），即当梯度向量大于某个阈值，缩放梯度向量。但对于梯度消失是很难解决的。

LSTM从被设计之初就被用于解决一般递归神经网络中普遍存在的**长期依赖问题**，使用LSTM可以有效的传递和表达长时间序列中的信息并且不会导致长时间前的有用信息被忽略（遗忘）。与此同时，LSTM还可以解决RNN中的梯度消失/爆炸问题。

LSTM的设计者提出了“长短期记忆”的概念——只有一部分的信息需要长期的记忆，而有的信息可以不记下来。同时，我们还需要一套机制可以动态的处理神经网络的“记忆”，因为有的信息可能一开始价值很高，后面价值逐渐衰减，这时候我们也需要让神经网络学会“**遗忘**”特定的信息。

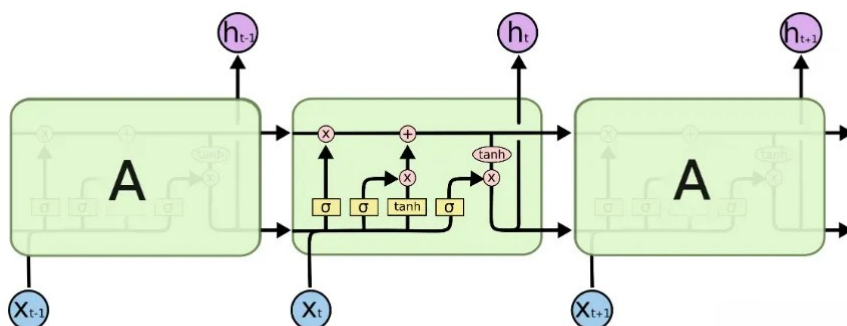
## (2). 具体解释

一个使用  $\tanh$  函数的RNN可以这么表示：



标准 RNN 中的重复模块包含单个层。

可以看到A在  $t-1$  时刻的输出值  $h-1$  被复制到了  $t$  时刻，与  $t$  时刻的输入  $x_t$  整合后经过一个带权重和偏置的  $\tanh$  函数后形成输出，并继续将数据复制到  $t+1$  时刻。与上图朴素的RNN相比，单个LSTM单元拥有更加复杂的内部结构和输入输出：



LSTM 中的重复模块包含四个交互层。