

时间序列分析在 FCC 产品产量预测中的应用

1 问题背景

自 1942 年标准石油开发公司 (SOD)，即现在的埃克森美孚公司第一个催化裂化装置投入实际运行以来，催化裂化技术一直在稳步发展。目前，催化裂化已成为炼油工业重要的二次加工过程，并逐步发展成为了流化催化裂化（Fluid Catalytic Cracking, FCC）的形式。该工艺使原油在适当的温度、压力和催化剂的作用下，发生分解、异构化、氢转移、芳构化和缩合等一系列化学反应。原油被转化为天然气、汽油、柴油等主要产品，以及油浆、焦炭等副产品。

图 1 是 FCC 工艺原则流程图。原料油与未反应完全的循环物料进入提升管后，与高温再生催化剂接触，立即汽化，裂解为轻质油（柴油、汽油等），同时产生油浆和焦炭。为了提高理想产品的收率，在提升管上部采用急冷技术来抑制二次裂化。反应过后的催化剂依次进入第一、第二再生器，进行循环再生。

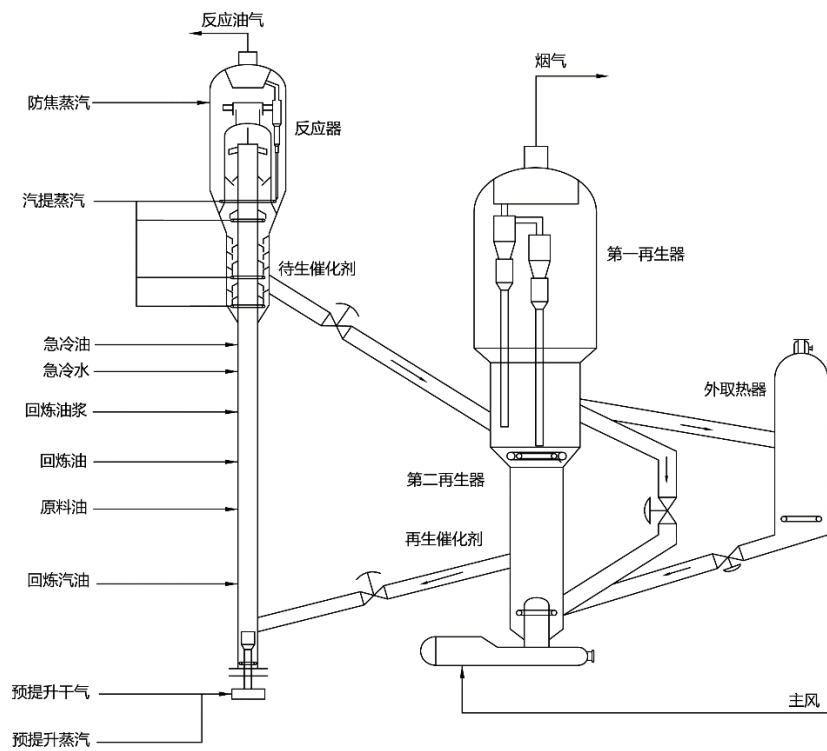


图 1 FCC 工艺原则流程图

为了实现更好地过程优化与异常识别，准确预测产品产量成为了 FCC 领域的重点关注问题。

2 问题描述

由于 FCC 过程通常是动态的，产品产量的预测应该考虑数据的时间特征，因此建立时间序列分析模型进行产量预测。该问题可以描述为：由 FCC 过程 t 时间的各个要素观测值，预测 $t+n$ 时间的产品产量 ($n=1, 2, 3\cdots$)。

在本案例分析中，收集了西南某炼油企业 2016 年 6 月-2017 年 4 月 64 个位点平均每 8 个小时的观测数据（共 912 组），并以此建立时间序列模型预测汽油产量。

3 方法应用

时间序列预测分析就是利用过去一段时间内某事件时间的特征来预测未来一段时间内该事件的特征。这是一类相对比较复杂的预测建模问题，和回归分析模型的预测不同，时间序列模型是依赖于事件发生的先后顺序的，同样大小的值改变顺序后输入模型产生的结果是不同的。

时间序列模型最常用最强大的工具就是长短期记忆神经网络（Long Short-Term Memory, LSTM），经典的 LSTM 模型结构图如图 2 所示。这是一种改进的循环神经网络，它在普通 RNN 单元上新引入了三个门控单元：遗忘门（forget gate）、输入门（external input gate）以及输出门（output gate）。通过训练，模型可以自适应的牢记重要信息，同时忽略次要信息。

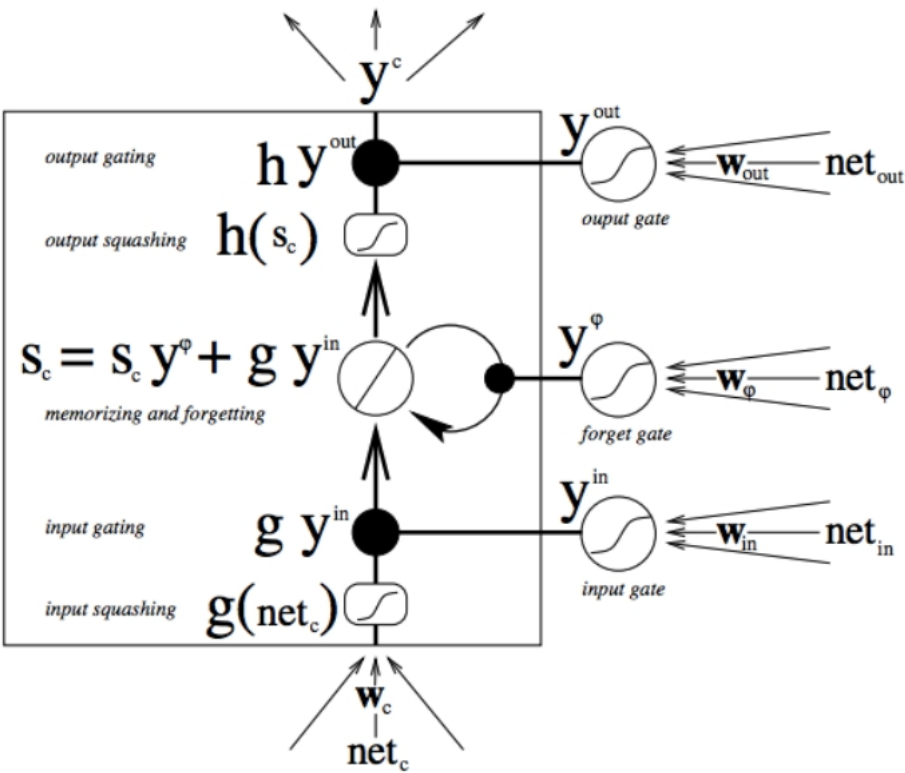


图 2 经典的 LSTM 模型结构示意图

利用 LSTM 对 FCC 过程建立时间序列分析模型，建模步骤如下：

(1) 数据预处理；

将数据进行归一化，避免量纲对数据造成影响。

(2) 将时序数据转换为监督数据；

将时间标签为 t 的预测变量向后滑动一个时间步，作为输出变量，时间标签为 $t-1$ 的所有变量作为输入变量。

(2) 划分数据集；

将原始数据集的 80% 划分为训练集，20% 划分为测试集。

(4) 模型的训练及预测。

4 结果展示

LSTM 模型的在训练过程中，将模型的 MSE 指标作为损失函数，其在训练过程中的变化如图 3 所示，最终收敛到 0.0066。模型的训练值、测试值与工业实际过程数据对比图如图 4 所示。

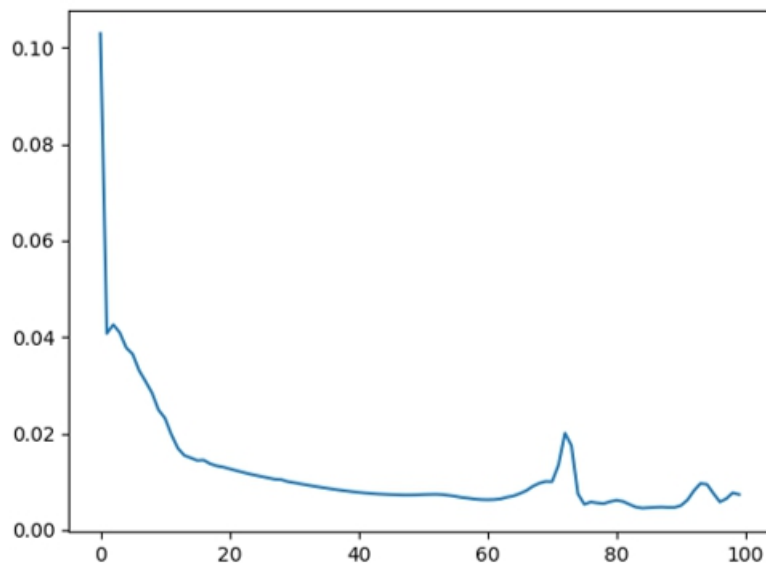


图 3 LSTM 训练过程损失函数图

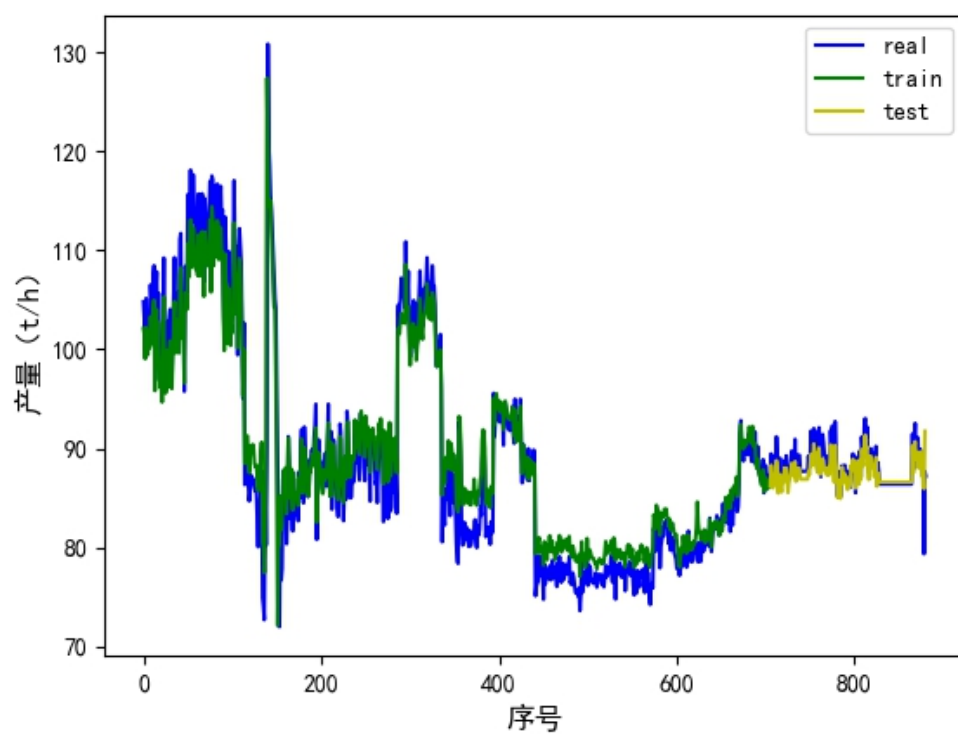


图 4 汽油产量训练值、测试值与工业值的比较图