

SVM 算法在预测反应再生系统汽油产量的应用

1 问题背景

催化裂化是重质油在酸性催化剂作用下，在 500℃左右、0.1~0.3 MPa 的压力下发生以裂化反应为主的一系列化学反应生成轻质油的过程。目前，我国催化裂化工艺生产的柴油和汽油约占成品柴油和汽油总量的 30%和 70%左右，已经成为重油加工的最重要方法之一。催化裂化的工艺过程和产品收率的建模分析一直是石油加工领域研究的热点和难点。并且由于催化裂化是一个高度非线性和相互强关联的系统，使用传统的机理模型难以全面地去描述，大数据技术便成为解决这一问题的有力工具。由于催化裂化过程存在非常多的变量，其中原料油性质、反应再生催化剂性质、反应操作工况条件等因素都会影响到反应过程和产物收率，并且变量之间存在强干扰、大滞后和强耦合等特点，为了减小模型复杂度、过滤数据噪音、提高模型的准确度和泛化能力，应用降维技术与机器学习相结合的催化裂化预测模型是一个不错的选择。

2 问题描述

图 1 是催化裂化工艺中的关键系统——反应-再生系统，包含提升管反应器和催化剂再生器。原料喷入提升管反应器下部，在此处与高温催化剂混合、气化并发生裂化反应。产物与催化剂经旋风分离器分离后，进入后续分馏塔分离出汽油、柴油等轻质油产品。结焦的催化剂则进入再生器用空气烧去焦炭后循环使用。本案例的主要目的是利用降维技术与机器学习相结合的预测模型对反应-再生系统的汽油产量进行预测。

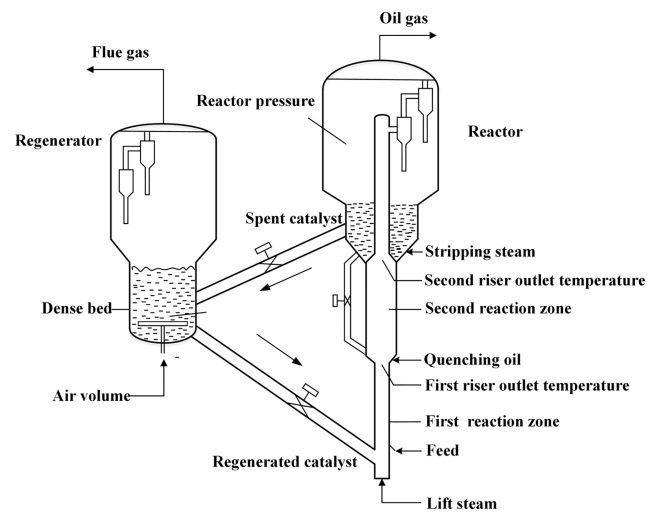


图 1 催化裂化反应-再生系统

3 方法与结果

本案例数据来源为某石化企业的 LIMS 及 DCS 系统中的实际工业数据，共有 912 组样本，每个样本包含 64 个特征变量。数据的特征变量过多，不利于模型的预测准确度和泛化能力，需进行数据降维处理，从而去除掉数据的一些冗余信息和噪声。在降维方法中，PCA（主成分分析法）是一种最常采用的降维算法之一，具有简单高效且能最大程度上保留原始数据信息量的特点。在预测算法中，SVM（支持向量机）通过寻求结构化风险最小来提高模型泛化能力，实现经验风险和置信范围的最小化，即使在样本数量较少的情况下，也能获得良好的预测效果。故本案例采用 PCA 降维方法与 SVM 相结合的催化裂化预测模型对汽油产量进行预测。

结果显示，模型的均方误差 MSE 为 0.159%，R2 系数为 95.3%，预测结果与实际汽油产量相比，误差率小于 1%，且由图 2 模型预测的汽油产量与工业实际值的比较图可以看出，数据点均匀且紧密分布在线条（ $y=x$ ）的两侧。表明构建的 PCA-SVM 预测模型能精确预测催化裂化装置中汽油产量，这对装置操作工况的优化改进具有良好的指导意义，有助于在实际生产中进一步提升催化裂化装置的经济效益。

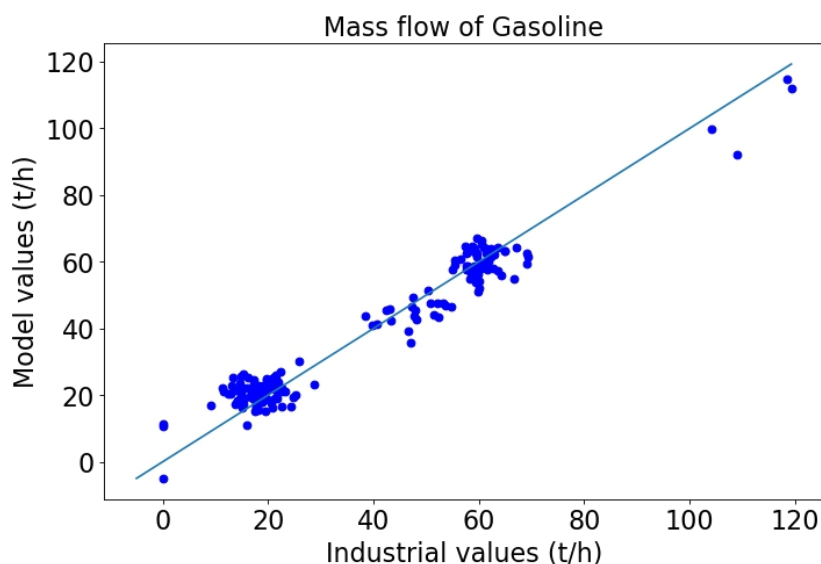


图 2 模型预测汽油产量与工业实际值的比较