# 工业物联网中涡轮风扇发动机的预测性维护研究案例教学指导手册

**教学适用的课程：**《大数据架构与设计》

**教学适用的专业：**电子信息

**教学目的与用途：**理解实际工业场景下的大数据应用；掌握大数据预处理、数据分析相关知识点；掌握大数据处理与分析的基本流程。

**教学内容：**

（1）理论依据：本案例基于大数据技术对工业物联网中涡轮发动机进行预测性维护。案例整合了常用的数据挖掘与机器学习方法，进行数据预处理与回归、分类及预测，实现了大数据依托的工业物联网决策，对传统工业应用进行智能化处理。

（2）涉及知识点：数据预处理；回归预测；分类预测；预测评估。

（3）分析路径：首先抛出工业大数据背景，并进行数据集展示，介绍相关字段含义。通过数据可视化观察数据变化趋势，进行数据预处理，提供三个参考步骤，包括删除低方差特征、数据归一化与标准化以及数据降维。然后对处理后的数据进行发动机剩余使用寿命预测，并提供三种方法包括：线性回归、广义线性回归以及决策树回归，每一种方法的预测结果采用均方根误差损失函数进行评估。最后进行涡轮风扇发动机工作状态的判定，在对正负样本不平衡的数据集采用SMOTE算法进行平衡。教学中提供三种方法进行状态二分类判定，包括支持向量机、深度神经网络以及逻辑斯蒂回归。模型评估采用准确率，精确率，召回率和AUC值作为评价指标。

**启发思考题：**

涡轮风扇发动机工作状态正常状态与警报状态的数据比例约为12:1，除了对不均衡的正负样本使用SMOTE算法进行抽样，还能想到其他的解决方法吗？（参考答案：调整正负样本权重的方法，比如Focal Loss的使用）

**建议课堂计划：**

1. 时间安排：3课时
2. 学生学习准备：了解常见的数据预处理、回归预测、分类预测方法的理论知识。
3. 分组及讨论内容：分组讨论启发思考题。
4. 案例开场白：工业物联网数据是大数据的一个重要应用场景。工业物联网大数据融合了具有感知、监控能力的各类传感器以及移动通信、智能分析技术，从而提高制造效率、改善产品质量、降低产品成本。这节课我们将接触一个大数据在工业生产中的实际应用。
5. 结束总结：这节课我们以工业物联网大数据为例，进行了数据预处理；通过回归预测与评估，实现了发动机剩余使用寿命预测；又通过分类预测与评估，实现了涡轮风扇发动机工作状态的判定。我们通过平时课程所学知识去解决了实际工业生产中的问题，课后感兴趣的同学可以搜集相关资料，进一步了解大数据在其他行业的应用，相信你们会收获更多。
6. 案例引导建议：在教师简要分析应用场景后，鼓励学生自主考虑如何使用所学知识解决该场景问题。

**参考文献：**

[1] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016.

参考内容：数据预处理理论，线性回归、广义线性回归、决策树、SVM、logistic回归算法

[2] Goodfellow I, Bengio Y, Courville A. Deep learning[M]. MIT press, 2016.

参考内容：深度神经网络DNN模型

**其他教学支持材料：**

1. 本案例提供配套的PPT、视频、数据集与代码等，发布于Github，链接为：https://github.com/Wanghui-Huang/CQU\_bigdata。

2. 本案例涉及到数据预处理以及多种机器学习算法，建议使用python语言进行编写，推荐的工具包有pandas（数据读取与预处理库），scikit-learn（机器学习算法库），matplotlib（可视化绘图库）。