

如何融合机理模型、专家经验和数据驱动实现工业人工智能？

讨论内容

1. 机理模型、专家经验和数据驱动模型的融合
2. 各种模型的优缺点和适用场景
3. 机理模型与数据驱动模型的融合
4. 数据驱动模型与专家经验的融合
5. 可解释性机器学习的应用
6. 参考资源

1.机理模型、专家经验和数据驱动模型的融合

机理模型

专家经验

可复用无幻觉准确的知识库

数据驱动

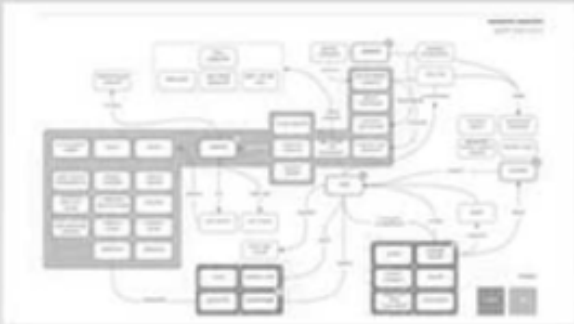
通过大量数据分析和AI算法模型

- 发现长期隐性的异常趋势
- 分析多变量场景下的因子作用
- 拟合机理模型有局限的实际场景


刻画物理世界的三种模型

认知模型

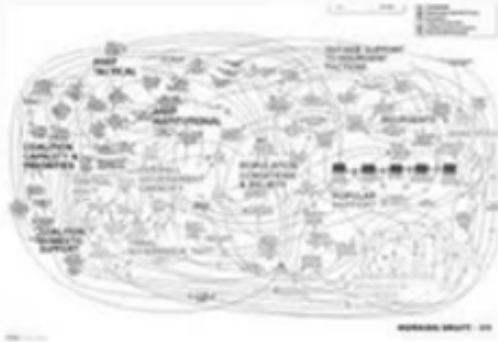
• 概念模型



• 经验知识




• 系统动力学



理论模型

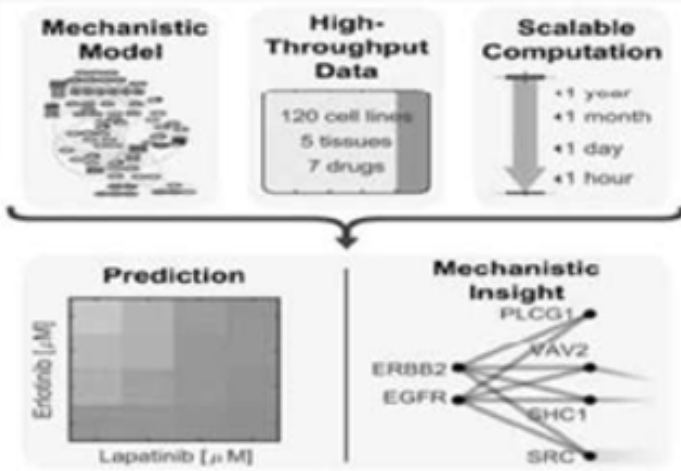
• 唯象模型 (Phenomenological Model)

分析模型
模拟仿真




• 机理模型 (Mechanistic Model)

平衡方程: 物质、能量、动量
结构方程: 物理化学状态方程
联立方程: 基尔霍夫电流定律

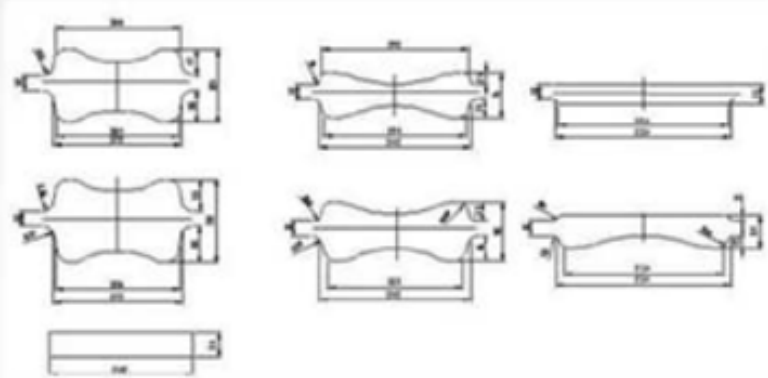


经验模型

• 专家规则



• 经验公式


$$P=r \cdot Z \cdot Z_r \cdot \sigma_3 \cdot \left(1+\frac{f}{3} \cdot \frac{D}{h}\right) \cdot F$$

式中: r 为变形时的受力形态和体积特性系数; Z 为变形时的应力状态不均匀影响系数; Z_r 为变形时的温度不均衡影响系数。

图片来源: 田春华,李闯,等. 工业大数据分析实践

2.机理模型、统计模型、仿真模型和专家规则的优缺点及适用场景

	优点	前提（或限制）	适用场景
机理模型	分析推演能力强	基于大量的简化或强假设 模型参数的可测量性	理论基础、实验条件良好
统计模型	归纳能力强，具备自适用能力	对数据的要求高 预测结果具有一定的不确定性	大量类似的场景 概念逻辑清楚，但缺乏具象的关系
仿真模型	计算推演能力强，可以计算不同场景下的行为（包括极端情形）	物理过程模型和输入不会与现实运行互动（一旦给定，就不再改变）	假设推演分析（What-if 分析）
专家规则	可解释性强	规则的模糊与不完备	逻辑简单明了，需要实时计算

图片来源： 田春华. 工业大数据分析算法实战. 2022年6月. 机械工业出版社

3. 机理模型与数据驱动模型的融合

方面	校准机理模型	后处理机理模型	获取数据驱动模型特征	两种模型融合
特征	提供参数的点估计或者分布估计	采用统计方法修正机理模型； 采用统计方法综合多个机理模型	根据机理获取重要特征，具有可解释性	<ul style="list-style-type: none">机理模型获取全局动态演化过程数据模型刻画局部稳态周期模式
应用场景	卡尔曼滤波	天气预测问题	风机结冰预测	空气质量预测

4. 数据驱动模型与专家经验的融合

- **适用场景：**历史数据没有标记，领域专家提供少量异常样本。
- **融合方法：**
 - 将已有专家规则形式化，采用大量历史数据验证，完善专家规则逻辑。
 - 主动学习（Active Learning），旨在通过智能选择最有价值的数据样本来进行标注，从而在标注成本有限的情况下，最大限度地提高模型的性能。参考链接 [【学习资源】人在环路的机器学习](#)
 - 采样策略包括：不确定性采样、基于代表性的采样、基于密度的采样、信息增益法、多样性采样以及综合性策略
- **可使用的其他方法：**
 - 无监督学习，包括异常样本相似度匹配。参考视频 [【理论和案例】工业大数据分析方法论与故障预诊断和设备健康管理](#)

工业制造领域的问题和可解释机器学习能解决的问题

- 问题类型

- 简单问题：线性关系、低维
- 中等复杂问题：非线性关系、需特征分析
- 高复杂问题：复杂因果关系、多变量交互

- 问题场景

- 质量控制与缺陷检测、设备故障诊断、预测生产性能、预测设备寿命、预测与优化能耗、预测产品合格率、分类故障、优化过程工艺参数、优化生产流程、分类工艺条件、检测生产异常、推荐生产参数、故障预测与预防性维护、因果分析、风险评估

5.1 可解释性机器学习的应用

	决策树	随机森林	线性回归和逻辑回归	支持向量机
特点	<ul style="list-style-type: none">结构简单，易于可视化。每个决策路径对应一个具体规则，便于解释。	<ul style="list-style-type: none">基于多个决策树的集成模型，虽然复杂度高于单一决策树，但仍具有较好的可解释性。提供特征重要性评分，帮助理解哪些因素对预测结果贡献最大。	<ul style="list-style-type: none">简单、直观，适用于线性关系的建模。模型系数可以直接用于解释变量对结果的影响。	<ul style="list-style-type: none">虽然解释性稍低，但对于低维度问题，通过支持向量和边界位置，可以提供一定的可解释性。可用于分类和回归问题。
应用场景	<ul style="list-style-type: none">质量控制与缺陷检测设备故障诊断	<ul style="list-style-type: none">预测生产性能预测设备寿命	<ul style="list-style-type: none">预测能耗预测产品合格率	<ul style="list-style-type: none">分类故障优化工艺参数，找到最优工艺参数的决策边界
问题复杂度	简单	中等复杂	简单	中等到高复杂问题

5.2 可解释性机器学习的应用

	朴素贝叶斯	K-近邻	梯度提升树	因果推断与贝叶斯网络
特点	<ul style="list-style-type: none">基于概率的简单分类模型，特别适用于条件独立性假设成立的场景。提供后验概率，可解释预测结果的置信度。	<ul style="list-style-type: none">基于相似性度量，简单直观。结果可通过查看“邻居”样本的特征来解释。	<ul style="list-style-type: none">性能强大，解释性相对较好，提供特征重要性分析。可以通过树结构或特征重要性图来解释模型的预测逻辑。	<ul style="list-style-type: none">特别适合分析变量间的因果关系。提供可解释的因果路径。
应用场景	<ul style="list-style-type: none">根据历史数据，对生产工艺条件进行分类并解释条件组合的影响。利用传感器数据，检测生产过程中的异常模式。	<ul style="list-style-type: none">根据历史数据，推荐与当前条件最相似的成功生产案例的参数。根据设备历史运行数据，判断当前设备运行状态是否异常。	<ul style="list-style-type: none">用于优化复杂生产流程，找出对产量或质量影响最大的参数。根据时序数据，预测设备的故障时间，并分析关键的故障指标。	<ul style="list-style-type: none">分析生产参数之间的因果关系。评估参数对生产结果的潜在影响。
问题复杂度	简单到中等复杂问题	简单到中等复杂问题	中等复杂到高复杂问题	高复杂

6. 参考资源

- 田春华. 工业大数据分析算法实战. 2022年6月. 机械工业出版社
- 可解释机器学习:黑盒模型可解释性理解指南(第2版) (德)Christoph Molnar (克里斯托夫·莫尔纳)
- 【学习资源】人在环路的机器学习

