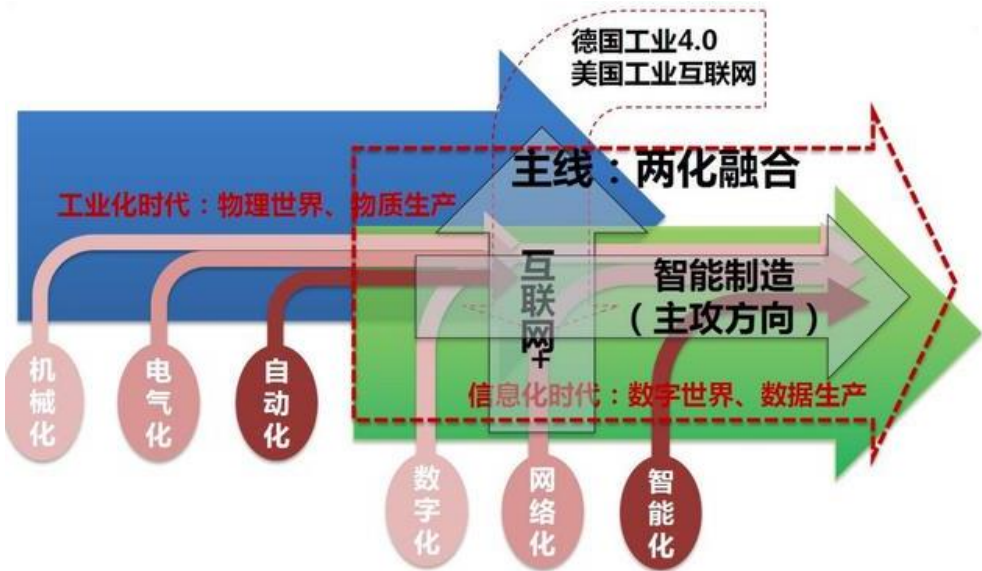


ALTAIR数据分析在工业大数据方面的应用

2022年7月 王甲畏 工程总监

工业4.0时代的特点



“信息化和工业化融合管理体系理论和基本框架”
---中国量化融合服务联盟，工信部两化融合

- **两化融合**：以信息化带动工业化，以工业化促进信息化，走新型工业化道路；
- **智能制造**：实现两化融合的核心途径；
- **人工智能**：基于算法，算力和大数据的综合技术；
- **工业互联网**：支撑企业实现智能制造的关键技术；
- **汽车新四化**：2020年3月，国家提出了加快新基建的建设进度，汽车出现了新四化：智能化、电动化、共享化、网联化；

海量的，异构的工业大数据

数据分析在工业智能领域的应用方向

【**数据的可视化分析**】：让用户理解数据所反映的设备运行状态，能耗，生产力状态等，反映的是发生了什么；

【**原因分析**】：机器能够自我诊断，找出问题发生的原因，同时可以根据历史维护记录和维护标准，告诉用户如何解决故障；

【**预测性维护**】：通过机器学习技术，在机器出现问题之前就可以感知或分析出可能要出现的问题，提前进行设备维护，比如机床刀具更换，反映的是将要发生什么；

【**采取措施，优化运营**】：通过对历史运营数据、工厂排产情况，和设备配置参数等的分析，提前检测出能耗的异常，并提供降低能耗的措施，反映的是该怎么去做。

描述性分析

诊断性分析

预测性分析

指导性分析

工业智能领域的数据分析分类：

- 基于历史数据的离线大数据分析
- 基于实时数据的在线实时分析



Altair 之于制造业

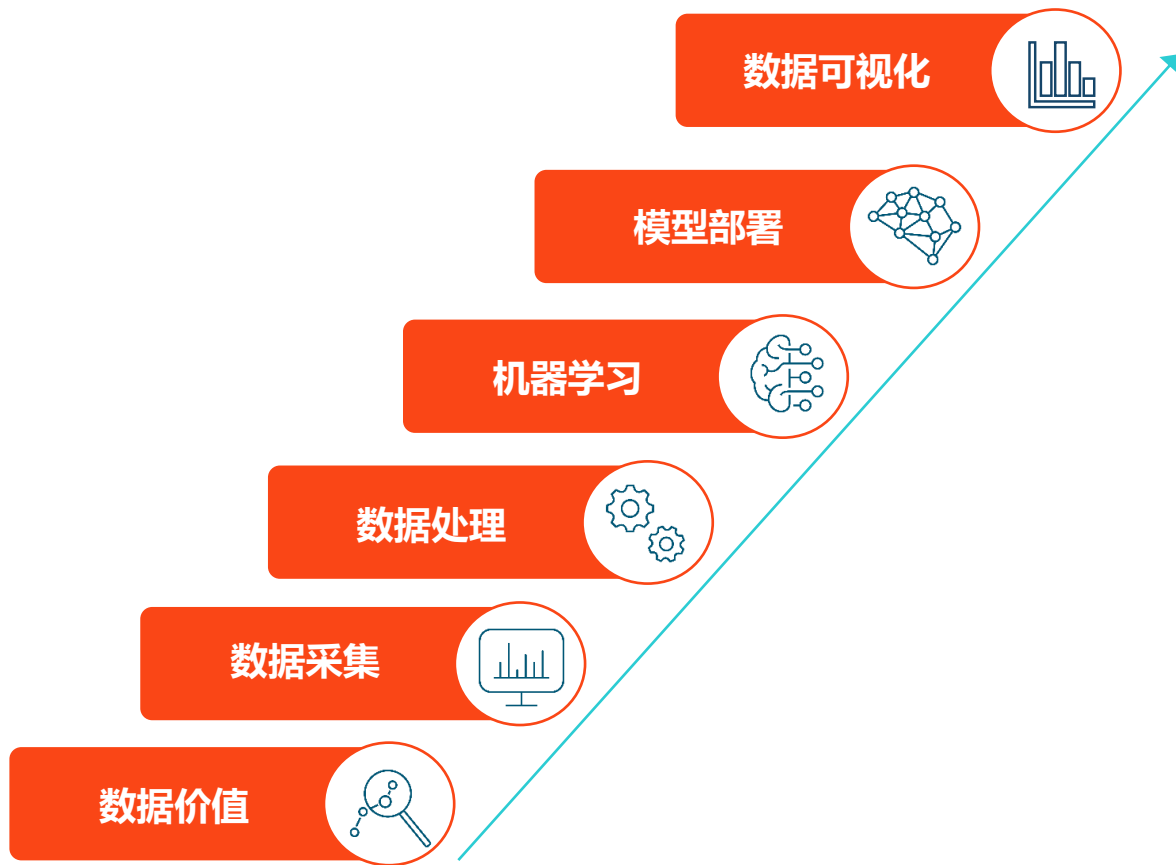
Altair针对计算机辅助工程、高性能计算（HPC）、系统建模、物联网和数据分析的解决方案能够为高性能、创新和可持续的产品和流程进行设计和优化。

Altair数据分析解决方案支持下一代业务和制造运营

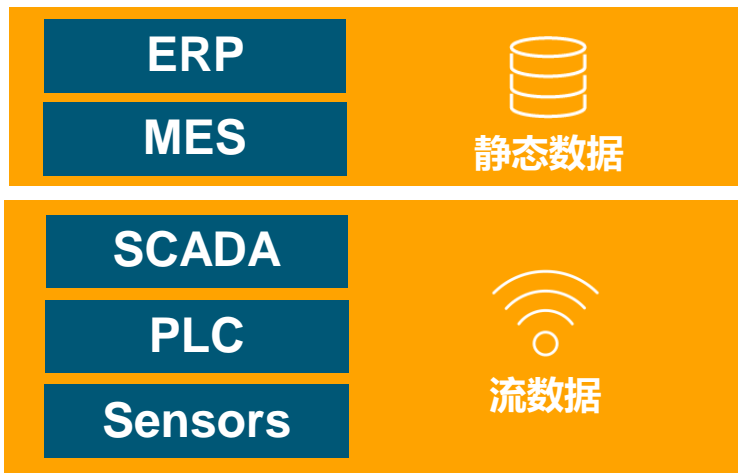
- 工艺改进
- 维保
- 预测性维护
- 产品工程
- 财务
- 销售
- 营销
- 客户支持

工业领域数据分析工作流程

制造业数据分析工作流程

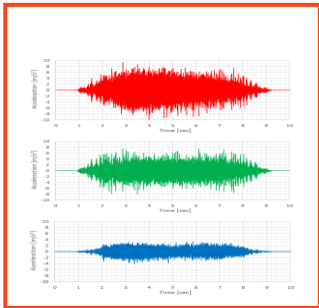


制造业数据分析工作流程

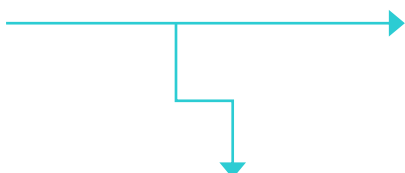
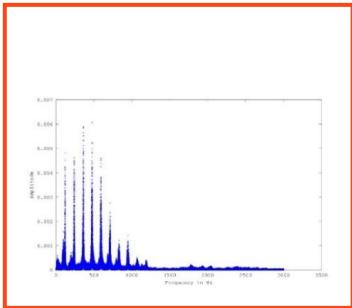


制造业数据分析工作流程

传感器原始数据



处理/训练的数据

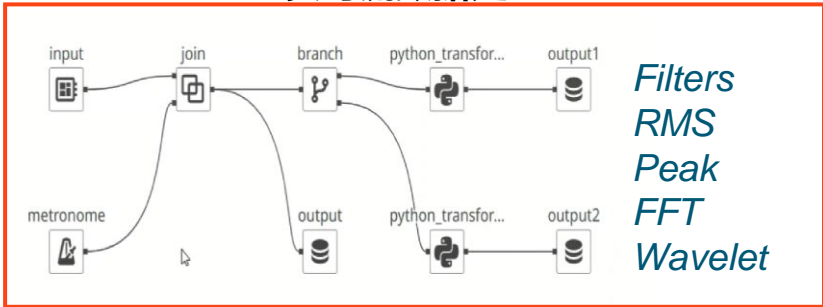


数据处理 

数据采集 

数据价值 

实时流数据处理



制造业数据分析工作流程

机器学习



数据处理



数据采集



数据价值



适用于各类
技术水平人员

易于模型构建—快速
部署

高性能可视化环境

制造业数据分析工作流程

模型部署



机器学习



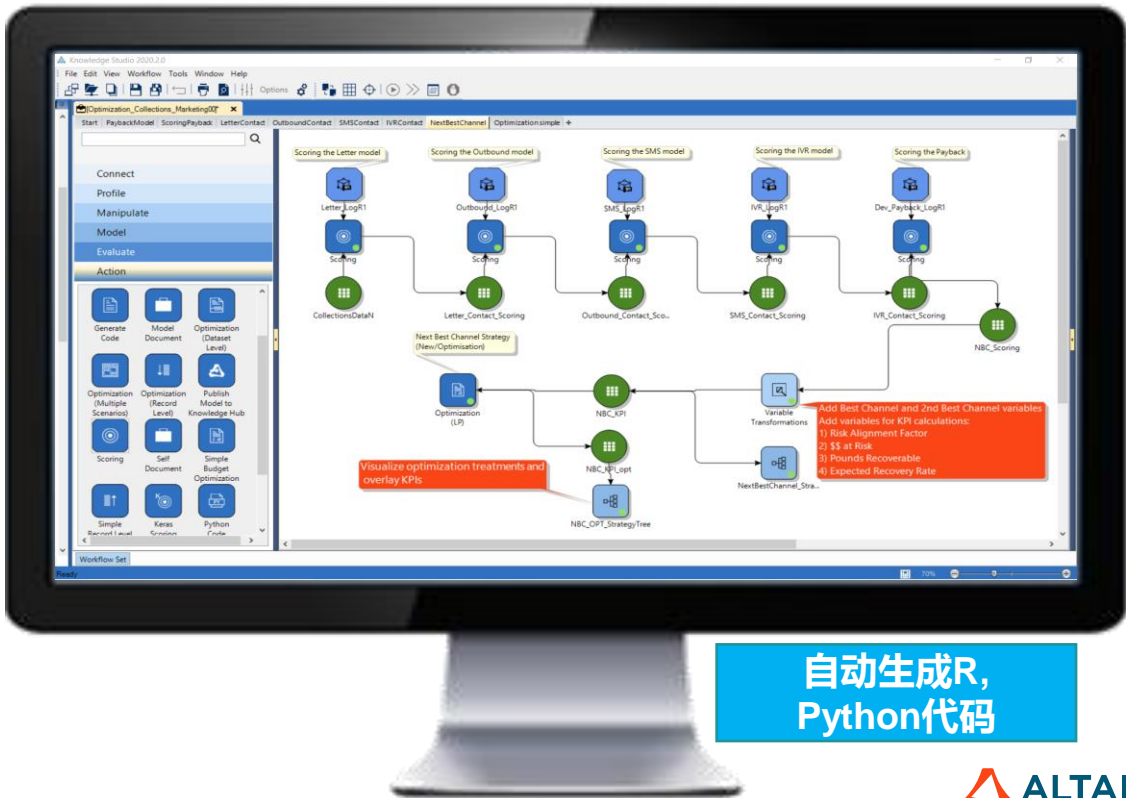
数据处理



数据采集



数据价值



自动生成R,
Python代码

制造业数据分析工作流程

数据可视化



模型部署



机器学习



数据处理



数据采集



数据价值



流数据源

时间序列数据库

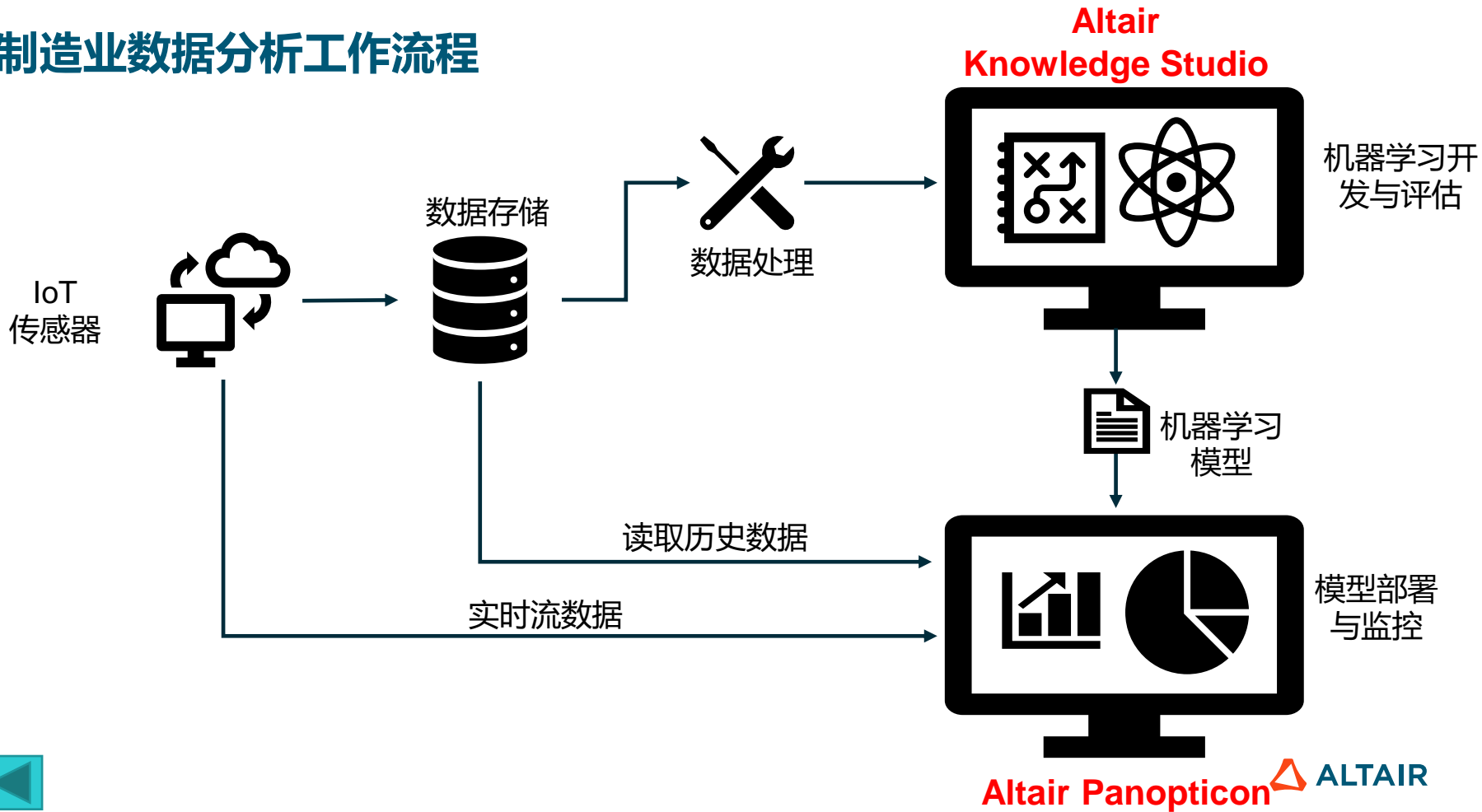
历史 / 综合
SQL / NoSQL 数据

实时数据可视化

- 监控实时流数据
- 以纳秒级精度分析时间序列历史数据
- 构建和部署分析仪表盘
- 用户设置的主动警报



制造业数据分析工作流程



▶ 流数据实时计算



智能工厂：机器学习用于生产流程监控



数据驱动的制造过程分析解决方案



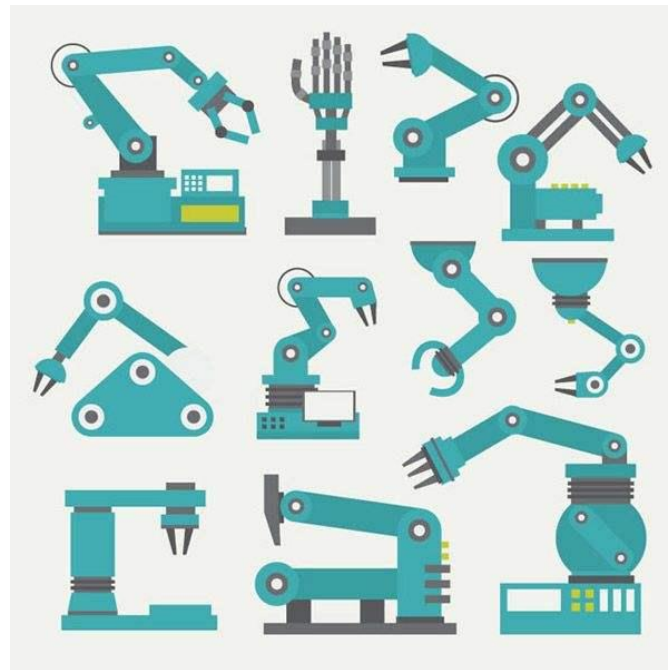
预测性维护



保修分析

机械臂预测维护

- 预测性维护是工业4.0内容的一部分，也是工业互联网的重要应用，2019年的汉诺威工业博览会直接把预测性维护列为工业智能的新风向之一
- 在自动化工业生产环境中，机械手臂的周期性检修和破损维修往往会对企业造成极大的误工损失。在传统防御性维修中，企业往往是人为的制定周期性检修计划，再逐个排查机械手臂零件，以防止一些极小问题最终造成难以接受的损失。
- 一家全球领先的自动化生产厂商选择了Altair Knowledge Studio，客户希望借助我们的软件完成数字化转型，并建立起合理的维修预测模型。以便客户能够最小化由机器破损带来的误工损失，减少不必要的运营花费，延长工业生产周期。
- 计划外的停机时间每年估计使工业制造商损失500亿美元。计划外停机的42%是设备故障的结果。计划外停机会导致过多的维护、修理和设备更换。---华尔街日报



机械臂预测维护

在这个短片中，我们主要是来演示如何应用Knowledge Studio来进行预测维护

异常状态检测

- 异常检测（或离群值检测）是从事件或观察结果中识别出罕见现象，这些事件由于与大多数数据有着显著差异而被怀疑为异常。通常，异常数据会**与某些问题或罕见事件关联**，例如**结构缺陷，设备故障**等。从业务角度来看，识别这些事件通常非常关键，这种关联让人们有必要找出哪些数据点可以视为异常。
- 任何机器，无论是旋转机器（泵，压缩机，燃气轮机或蒸汽轮机等）还是非旋转机器（热交换器，阀门等）都将最终变成欠健康状态。有可能不会真的发生故障或停机，但是设备无法处于最佳状态。这表明可能需要进行一些维护才能恢复最佳运行状态。简而言之，**确定设备的“健康状态”是状态监测的领域**。
- 实施状态监测最常见方法是监控机器上的每个传感器测量值，并对其施加最小和最大值限制。如果当前值在范围之内，则机器运行状况良好。如果当前值超出范围，则机器运行不正常，并发送警报。
- 但实际上，硬编码警报上下限会**发送大量错误警报**，也就是实际上是机器处于健康状态却发出警报。也会有未警报的情况，也就是明明有问题但未被警报。第一个问题不仅浪费时间和精力，而且浪费设备的可用性。第二个问题更为关键，因为它会导致实际损坏以及相关的维修成本和生产损失。

异常状态检测

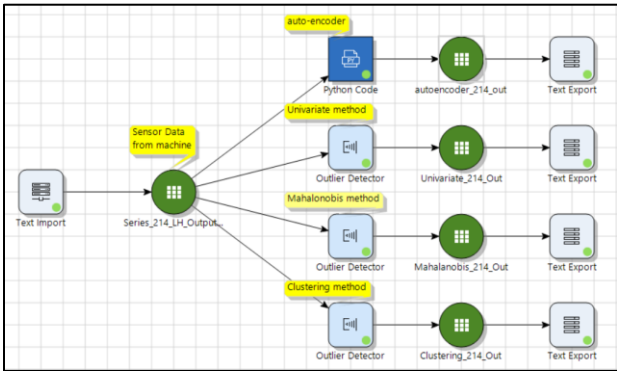
- 马氏距离，聚类，自动编码器 (DL) 方法

目标/问题陈述：使用生产企业整个生命周期中的传感器来识别机器组件中存在的异常行为或模式。异常数据可以指示关键事件，例如技术故障或潜在机会，例如机器行为的变化，从而为企业节省了资金并避免了停机。

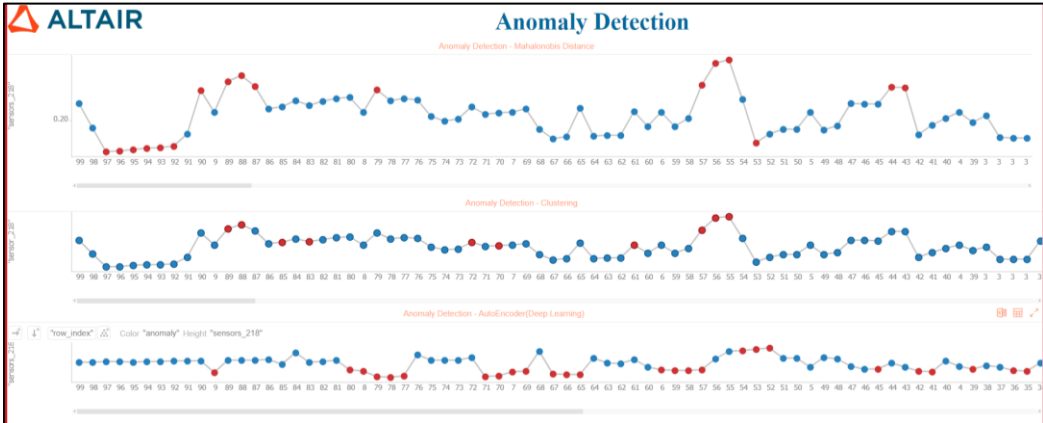
Knowledge Studio的方法：使用统计分析（如马氏距离，聚类等），我们可以隔离出从传感器收集的异常数据。使用Panopticon识别实时流数据中的异常情况，以便做出更好的决策并避免不必要的停机时间。

异常状态检测

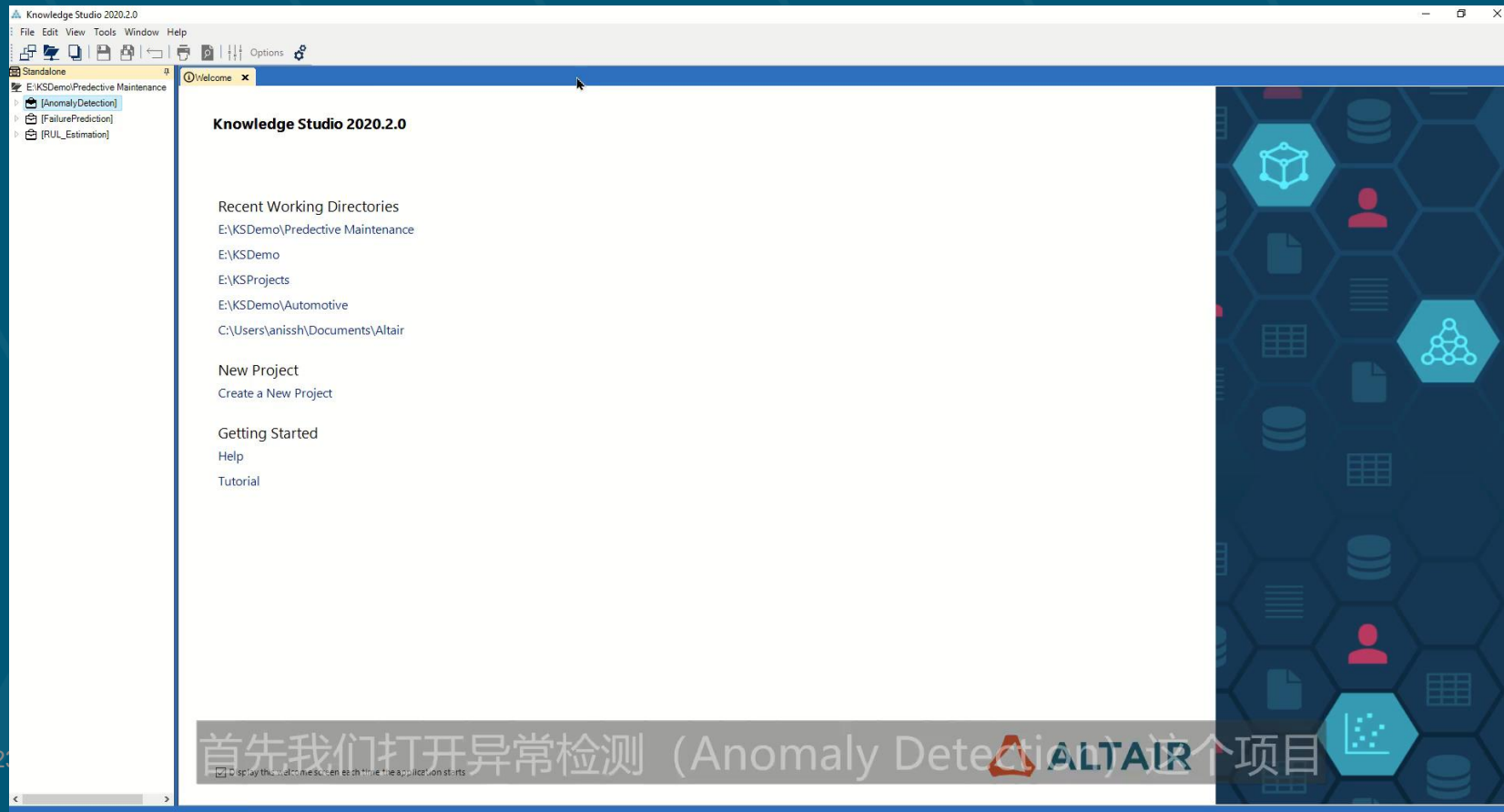
- 马氏距离, 聚类, 自动编码器 (DL) 方法



Time Index	series_214_LH Output Current	Outlier Flag
0	0.00631714	No
1	0.00631714	No
2	0.00631714	No
3	0.00631714	No
...
17	0.005984659	Yes
18	0.00631714	No
...
241	0.006551108	Yes
242	0.006064454	Yes
243	0.00631714	No



异常检测



首先我们打开异常检测 (Anomaly Detection) 这个项目



剩余使用寿命

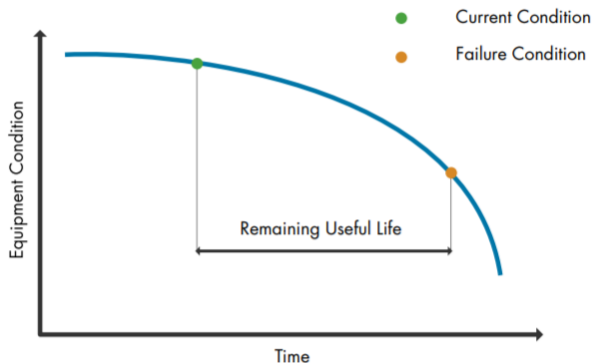
- 相似性模型，降阶模型，生存模型

目标/问题陈述： 预测在故障排除之前剩余的运行周期数发动机将继续正常运行的最后一个循环后的工作循环数。

剩余使用寿命（RUL）是一台机器在需要维修或更换之前可能要运行的时间。

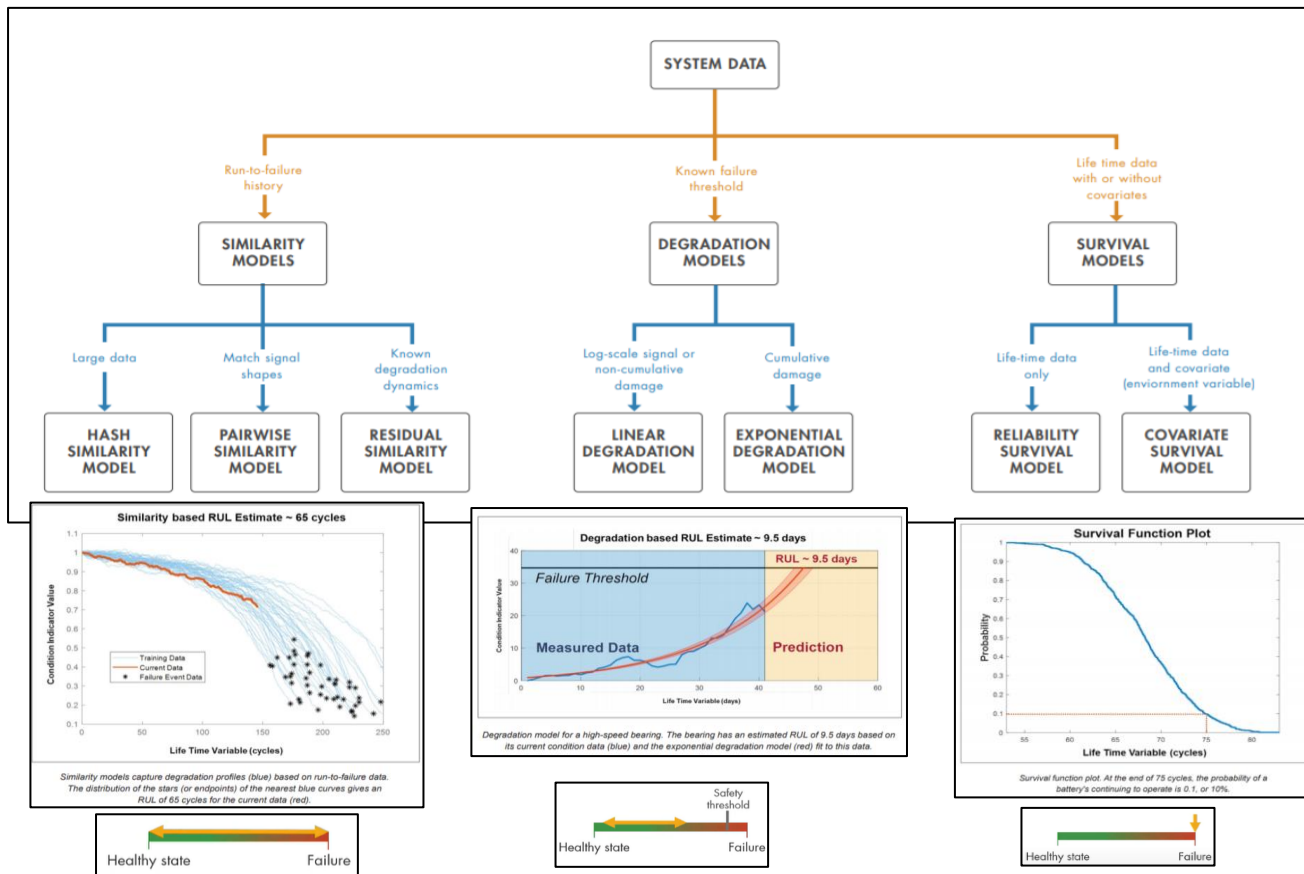
从飞机上的传感器收集的数据可提供有关各个零件状况的信息。

Equipment Deterioration Profile



剩余使用寿命

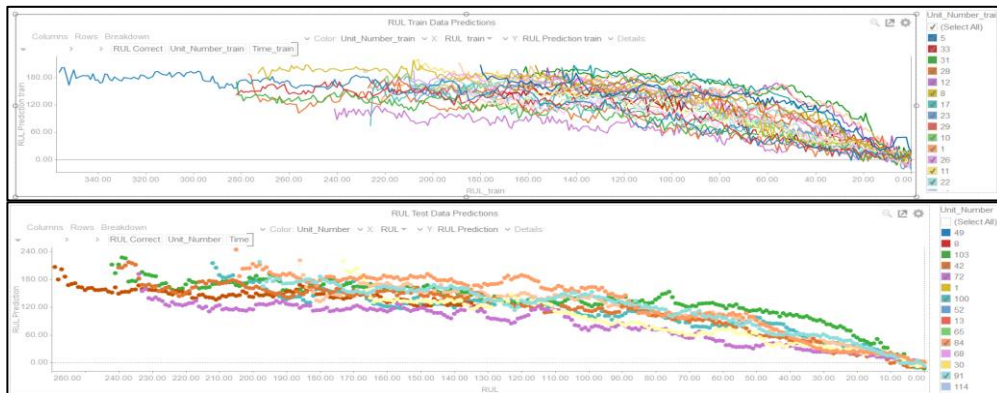
相似性模型, 降阶模型, 生存模型



剩余使用寿命

相似性模型

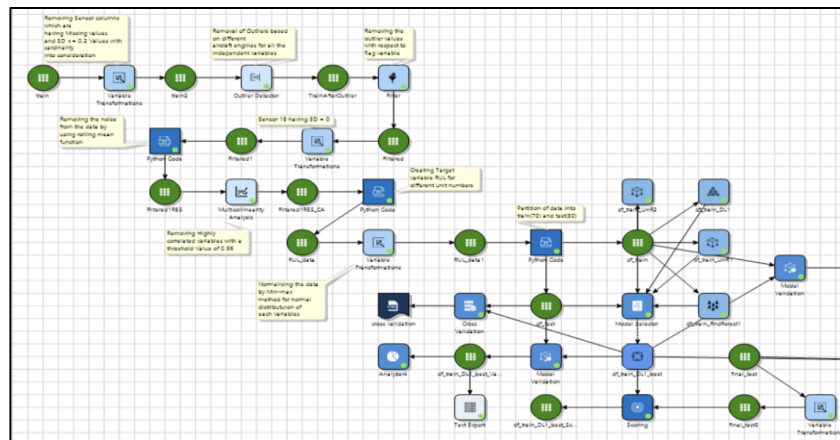
预计的RUL随时间的进展



Prediction summary on Test Data

Engine No.	Current Cycle	Actual RUL (cycles)	Predicted RUL (cycles)
3	119	30	33
6	58	163	166
21	38	166	161
135	120	65	74
216	53	136	136

KS Workflow Similarity models

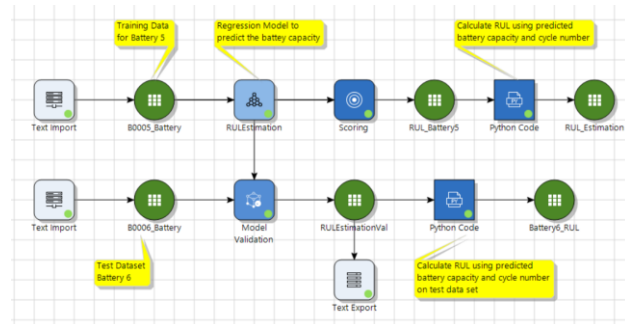
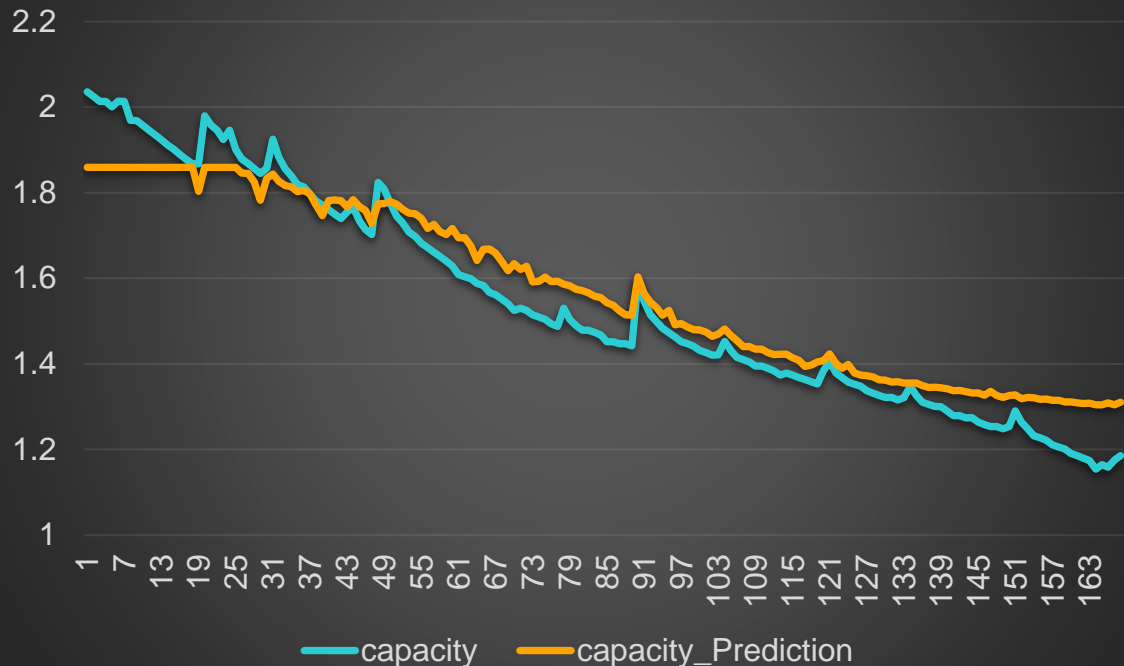


剩余使用寿命

降阶模型

© Altair Engineering, Inc. Proprietary and Confidential. All rights reserved.

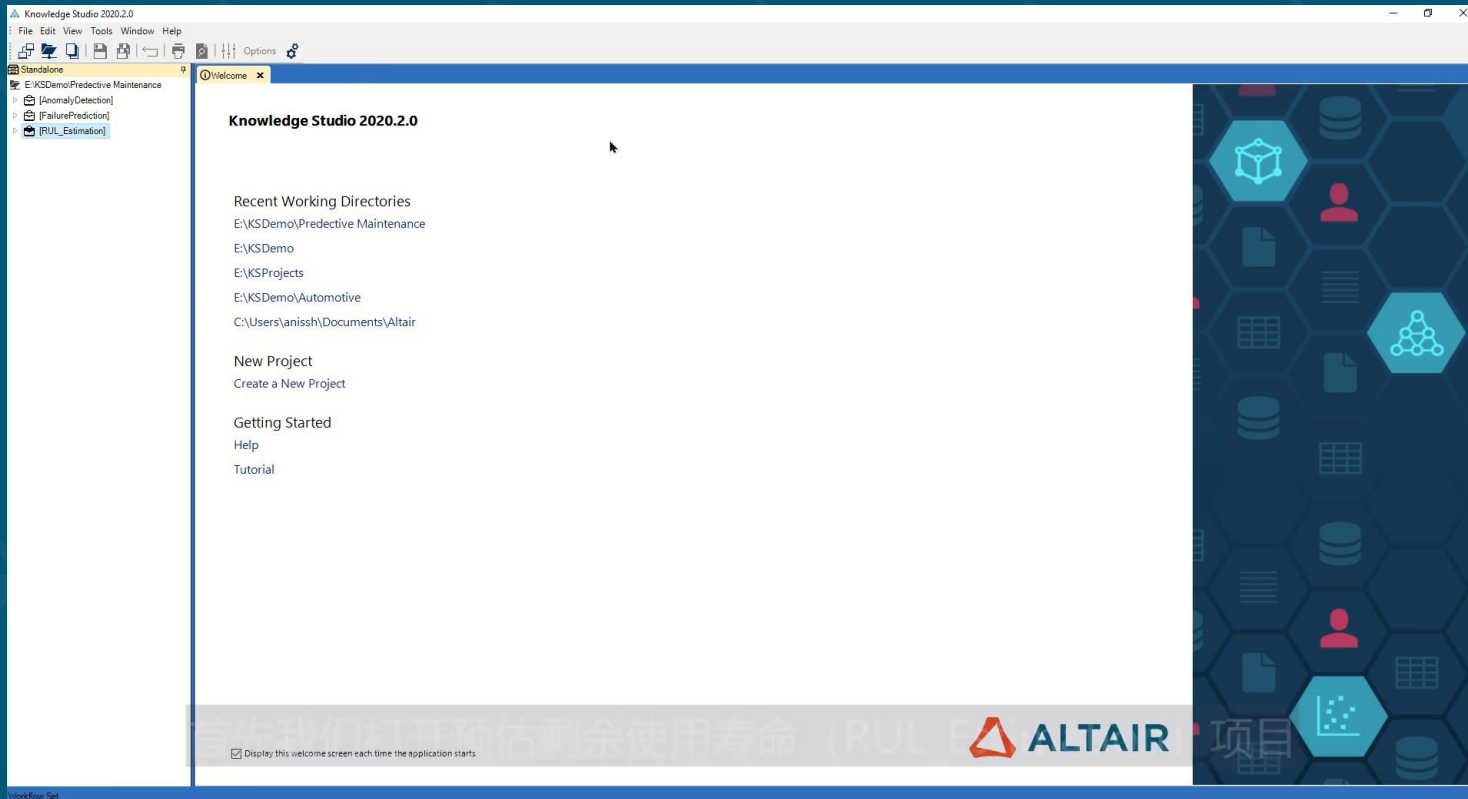
预测能力vs. 实际能力



Threshold Value = 1.35

Parameter	Values
Actual RUL	43
Prediction RUL	33
RUL error	10

剩余寿命预测



案例1：F1赛车性能和状态监控



案例2：福特

福特利用机器学习，加快冲压工艺选择，提供产品通过率

挑战：

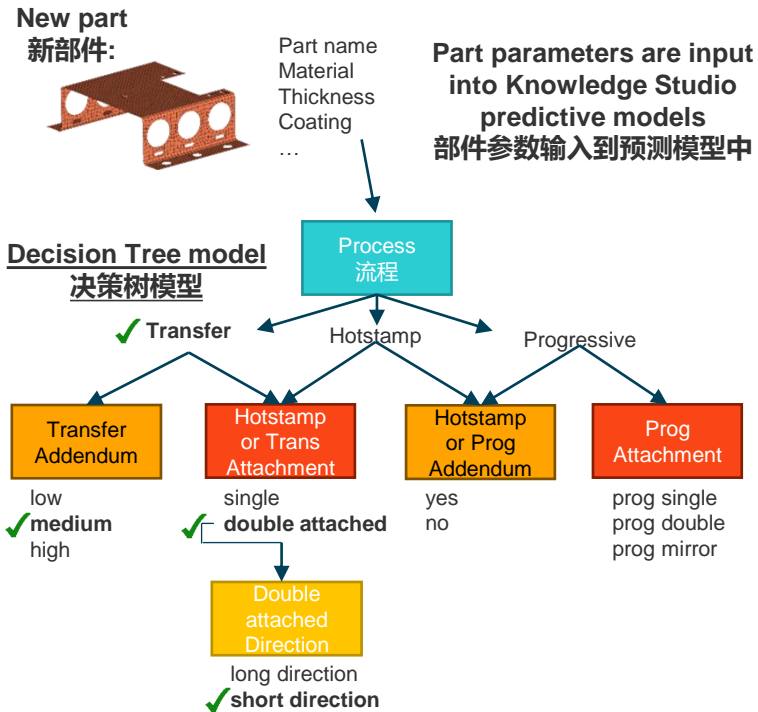
冲压工程师必须手动选择每个零件的冲压工艺

通过机器学习来解决

1. 构建基于机器学习的模型来预测可能的流程
2. 自动化决策工作流程，模型预测正确的冲压流程和附件
3. 将模型部署到冲压工作流程中，提高产品成功率

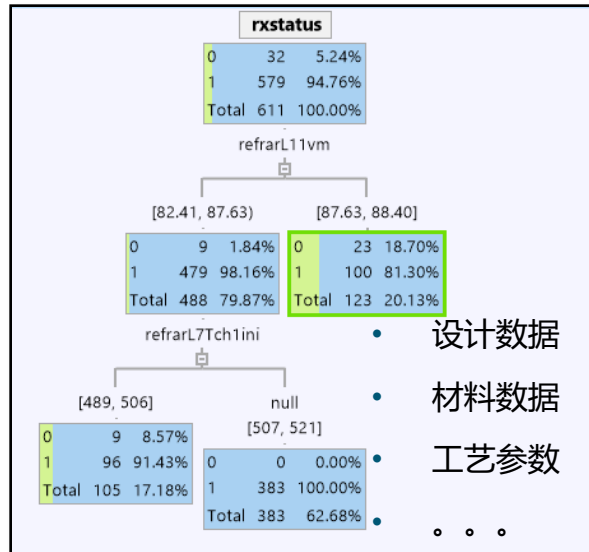
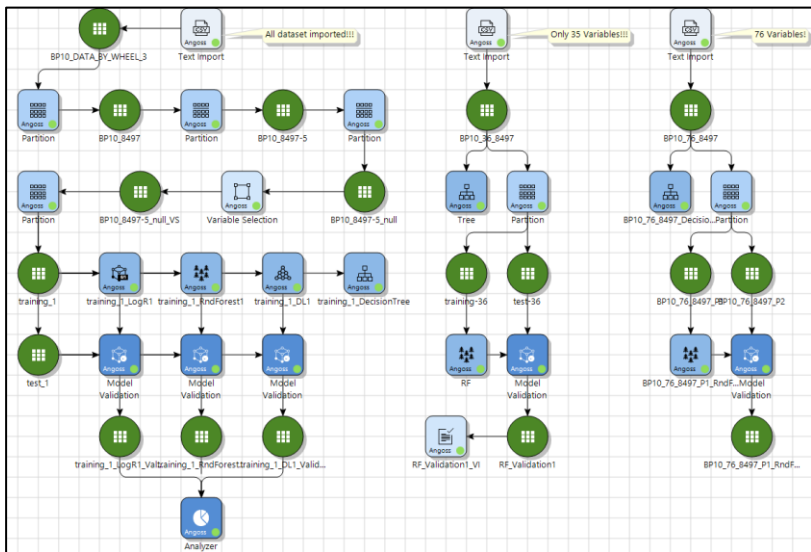
结果：

机器学习模型可自动预测正确的冲压过程，为福特工程师节省时间和费用。该模型还可用于培训新工程师。



The Decision Tree models predicts the right stamping processes, attachment and addendum
决策树模型预测了正确的冲压流程

案例3：利用机器学习，降低轮毂废品率



经过机器学习方法后，试验阶段：

- 生产882个轮毂，废品率 2.38%

没有使用机器学习改进流程之前的最佳结果：

- 生产361个轮毂，废品率 3.05%；
- 经过ML方法改进后，生产同样数量，废品率 1.44%。

减少编程所花的时间，以便有更多的时间用于有价值的数据分析

使用机器学习方法，产品废品率减少50%。

智能监控与预测维护



1. 机器上的传感器实时采集数据
2. 必须实现数据的实时传输



4. 基于历史数据，建立机器学习模型来预测分析（故障，宕机，剩余寿命）

连接机器

远程监控

预测分析

自动维护订单

3. 实时状态监控（设备，产品等）



5. 自动化生产计划，维护单

Altair Knowledge Works产品简介

Monarch

桌面级，自助式
数据准备



市场领先的，多种数据结构的访问与获取



面向每一位企业员工的数据市场

Knowledge Hub

企业级云端数据准备，
数据市场



便捷、灵活的部署方式



超过14,000客户的信任

Knowledge Studio

预测分析，数据科学



众多的企业应用案例



现代，企业级的软件框架与解决方案



极易上手，便捷易用性



值得数据科学家与商业分析师信赖的工具

Panopticon

实时，数据流可视化



数据治理与管控



真正的实时数据可视化

Altair® Monarch®

桌面级，自助式数据准备



可访问任何复杂的数据



点击鼠标即可完成数据准备

80+ Pre-Built Functions

No Scripting or Macros



合并及连接不同类型数据



最小化人工操作及人为误差



半结构化报表捕获

PDF & TXT →
Rows & Columns



自动化周期性、重复性的工作

Reusable 'Recipes'



可以导出到任何BI和数据分析平台



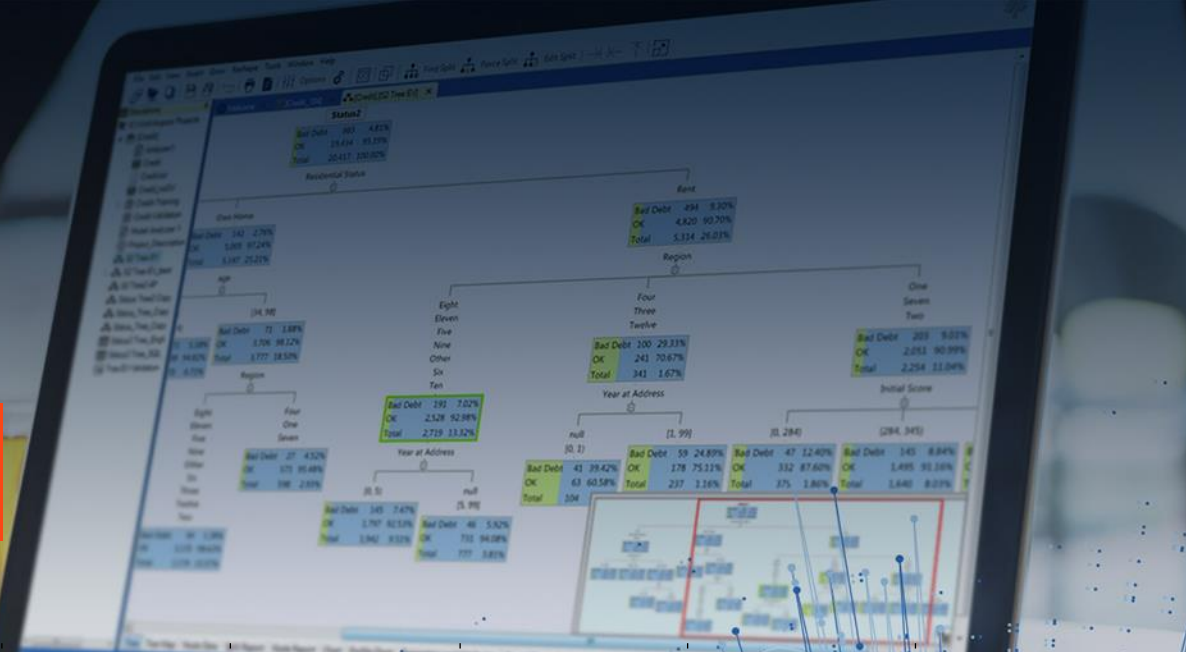
结果可审计，可跟踪

变化历史

传承与沿袭

Altair® Knowledge Studio®

预测分析，机器学习与
数据科学



可视化，
机器学习驱动的
预测方法

模型训练，验证
嵌入自定义代码



非常方便的
建模与部署

可重复的工作流
数据挖掘效率 +40%



行业最佳的
决策树

手工，自动增加
方便理解



自动生成代码

更少的编码时间
方便学习



高级统计学家
和商业用户

更快洞察
应用业务规则



可以驾驭
大数据

可视化分析HDFS, S3
通过Spark建立高级模型



全流程
可视化

探索数据，评估质量
检测模式和趋势



支持链接任何
数据源

SAS, SSPS, R, Hadoop
TXT, XLS
通过ODBC数据库

Altair Knowledge Studio主要功能

连接

- Excel
- Text
- R
- SAS
- SPSS
- HTML
- SQL Server
- Oracle
- DB2
- Microsoft Access
- DBASE
- FOXPRO
- Teradata

操作

- 数据汇总
- 数据合并
- 数据去重
- 数据筛选
- 数据透视表
- 数据集拆分
- 变量转换
- 变量重要性排序
- 变量筛选
- WOE
- 预测能力度量
- 多重共线性分析
- 异常值检测

模型

- 决策树
- 策略树
- 套袋法
- 提升法
- 聚类分析
- 约束回归
- 深度学习
- 因子分析
- 线性回归
- 逻辑回归
- 购物篮分析
- 偏最小二乘回归
- 主成分分析
- 随机森林
- 正则化
- 拒绝推断
- 打分卡模型
- 生存分析
- 时间序列
- 模型融合
- Keras模型
- 模型评估

部署

- Java
- SAS
- PMML
- Python
- R
- SPSS
- SQL
- XML
- 数据点最优化
- 数据集最优化
- 插入R代码
- 插入Python代码

Altair Panopticon™

实时数据流可视化，复杂事件处理



最快速的，
可视化分析

数据流处理无
需编码



当日和历史
时间系列

纳秒级数据响应



多种图表选项

快照&时间序列

相关性，趋势，异常值

等级分析



可定制化界
面



特别适配商
业分析、资
本市场分析，以及
IIOT场景



实时连接

CEP 引擎 & Tick 数据

消息总线，直播 DBs



大数据连接

Spark, Cassandra

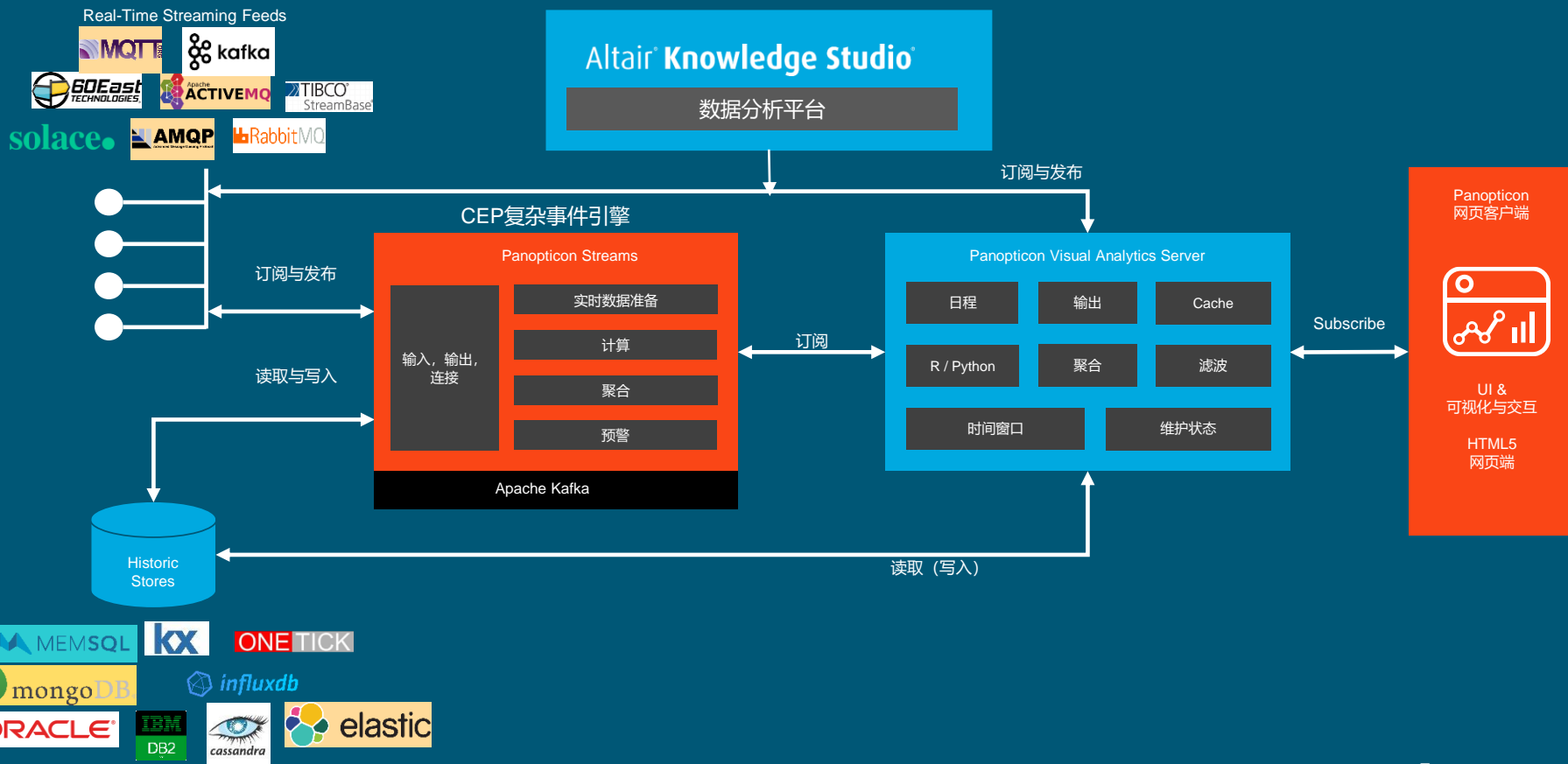
Apache, Impala, Hadoop



企业级

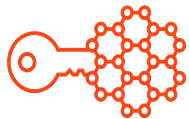
IIoT/IT架构

© Altair Engineering, Inc. Proprietary and Confidential. All rights reserved.



Altair: 变革制造流程分析

无代码数据自动化, 机器学习, 可视化



连接任何数据

- 来自任何来源的数据
- 清理和转换
- 治理与安全
- 自动化和共享
- 云或本地
- 没有编码



实时信号处理

- 实时流
- 时间序列数据
- 嵌入式系统
- 边缘连接
- 分享
- 云或本地
- 没有编码



准确的预测

- 机器学习
- 可解释的AI
- 自动化和共享
- 云或本地
- 没有编码



数据可视化

- 实时仪表盘
- 深度分析
- 分享
- ML模型部署
- 云或本地
- 没有编码

THANK YOU

更多资讯，欢迎关注Altair微信公众号



官网 : www.altair.com.cn

技术博客 : blog.altair.com.cn

邮箱 : KWinfo@altair.com.cn

技术服务热线: 4006-196-186