

# 【设备医生：故障预诊断和设备健康管理PHM】机加工刀具剩余寿命预测

杨晓春 2023-8-16

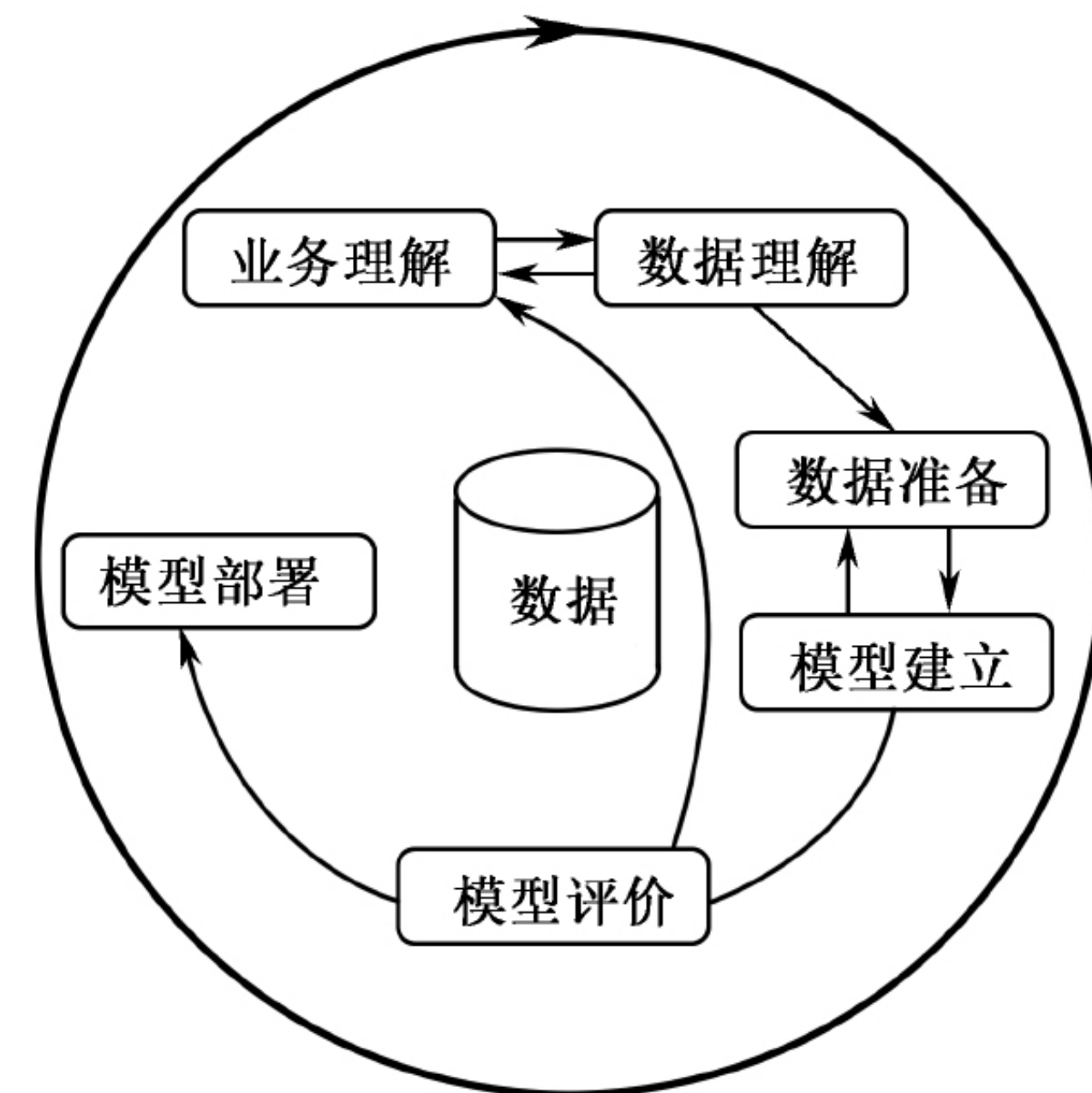
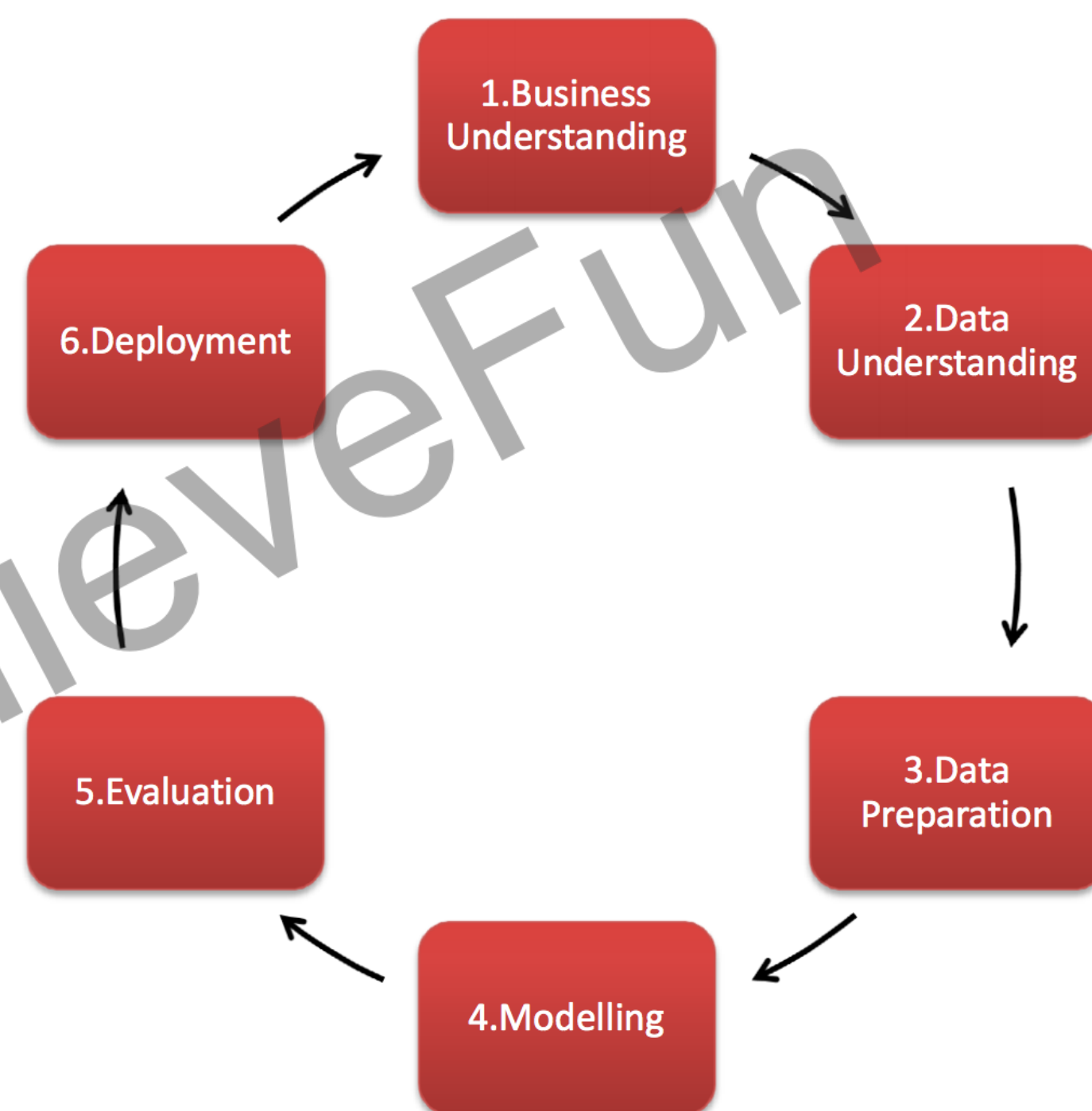


# 分享话题

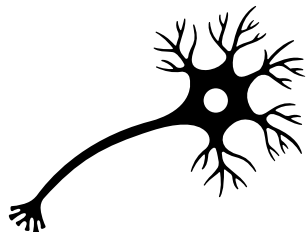
1. 业务理解：机床寿命监测的意义
2. 数据理解：描述、探索和检验数据，数据处理的难点
3. 数据准备：选择、清洗、生成（特征加工）、融合数据
4. 模型建立：选择、设计、建立和评估模型
5. 模型评价：评价结果、确认模型
6. 模型部署：计划、监测和维护

# CRISP-DM模型

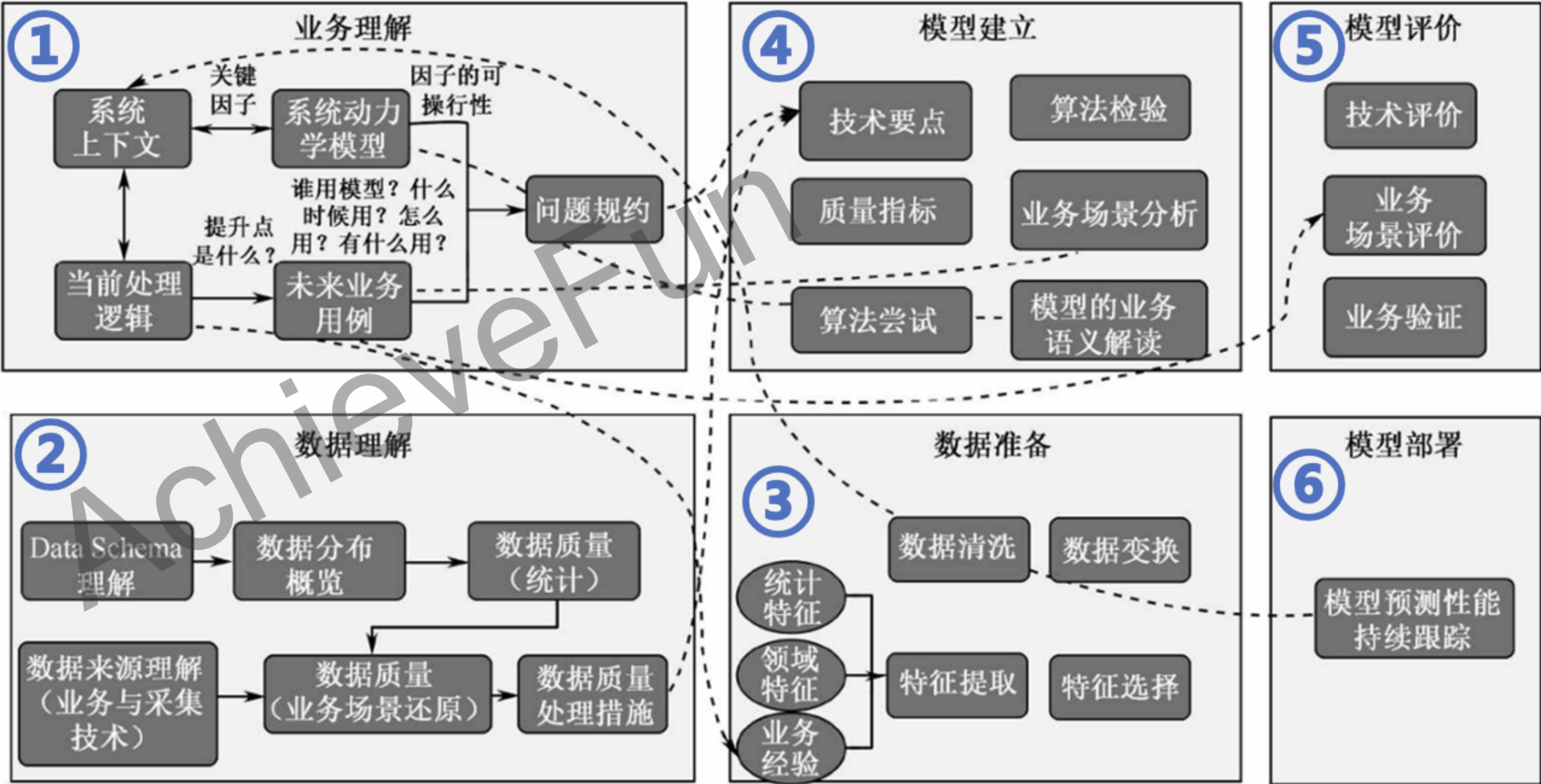
- 欧盟起草的跨行业数据挖掘标准流程（Cross-Industry Standard Process for Data Mining）
- 参考链接 <https://www.sv-europe.com/crisp-dm-methodology/>



# CRISP-DM细化方法



- 忽略哪些步骤容易被忽略?
- 哪些步骤花费较长时间?
- 每个步骤需要哪些理论和技术?



图片来源：田春华,李闯,等. 工业大数据分析实践 (Chinese Edition) . Kindle Edition.



# PHM的流程范式

1. 数据获取
2. 信号处理
3. 诊断
4. 预测
5. 维护决策

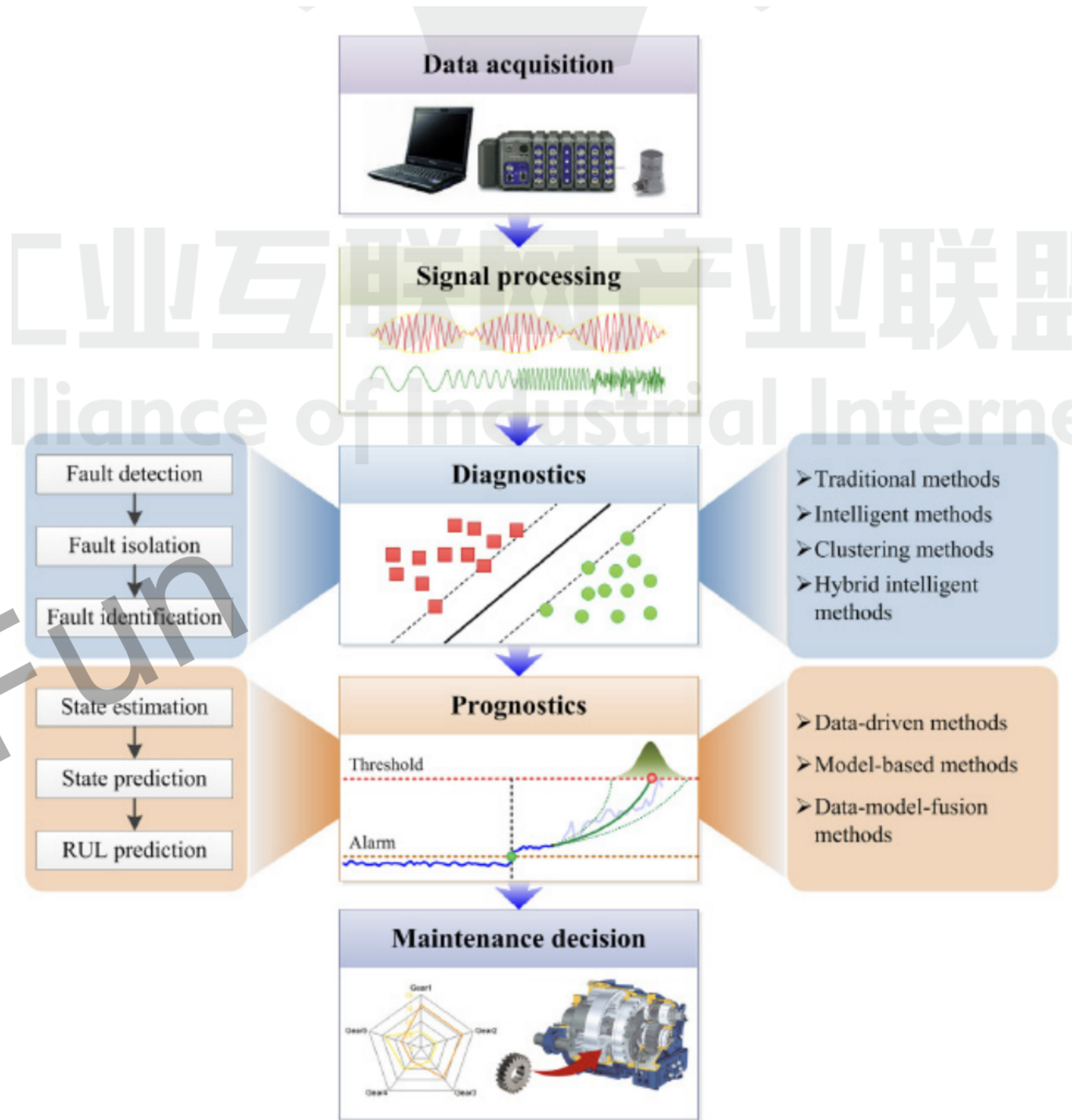


图 1-3-4 PHM 的流程范式<sup>[13]</sup>

图片来源: Tobon-Mejia D A, Medjaher K, Zerhouni N. CNC machine tool's wear diagnostic and prognostic by using dynamic Bayesian networks[J]. Mechanical Systems and Signal Processing, 2012, 28: 167-182.

# 案例来源

- **问题描述：**精准预测机加工刀具的剩余寿命，将有效地优化工作排程且降低刀具采购成本；
- 工业大数据创新竞赛白皮书（2018-2019）  
<http://www.aii-alliance.org/index/c318/n33.html>  
p8-p9
- **编写时间：**2020年8月
- **编写单位：**
  - 中国信息通信研究院
  - 工业大数据创新竞赛组委会
  - 工业互联网产业联盟

## （二）基于多传感器融合的刀具寿命预测方法

团队名称：411 小分队

成员姓名	团队角色	职位
刘思辰	队长	中科院声学所噪声与振动重点实验室博士研究生
刘锋	队员	中科院声学所噪声与振动重点实验室博士后
万伊	队员	中科院声学所噪声与振动重点实验室博士研究生

# 机加工刀具剩余寿命预测：数据资源

- **数据来源：** 机床控制器和外置传感器信号
- 包括加工过程中的**工况信息和传感器数据**。数据来源为实际 CNC 加工过程中，一把全新的刀具开始进行正常加工程序，直到刀具寿命终止时停止数据采集；
- 分成 4 组训练数据 以及 5 组测试数据。4 组训练数据将提供加工过程的 PLC 信号以及震动传感器的原始信号值，同时提供当组实验数据的完整加工「刀具全寿命时间」，单位为分钟。
- 震动传感器的原始信号值因数据量极大，仅提供每 5 分钟任取 1 分钟的片段 数据做为训练样本，并依据时间顺序 1.csv, 2.csv....n.csv 提供，其中最后一个 csv 文档 即为实验结束前最终 5 分钟的撷取数据内容。
- **数据采样频率：** PLC 信号采样频率为 33Hz，震动传感器采样频率 25600Hz。



# 刀具剩余寿命预测方法列表

#	方法名称	团队名称
1	基于经验模态分解的非参数模式匹配刀具寿命预测方法	杨立宁
2	基于多传感器融合的刀具寿命预测方法	411 小分队
3	基于前馈神经网络的刀具剩余寿命预测	沈国锐
4	基于 DTW 退化特征和 CNN+SVR 网络的刀具剩余寿命预测	MZC
5	基于 DWT-LSTM 网络的刀具剩余使用寿命预测	秋裤 ers

图片来源: [http://www.aii-alliance.org/upload/202009/0903\\_084520\\_197.pdf](http://www.aii-alliance.org/upload/202009/0903_084520_197.pdf)

- 特征提取: EMD 经验模态分解
- 特征处理: DTW 动态时间规整
- 数据清洗与信号分解: DWT 离散小波变化
- 训练模型: CNN 卷积神经网络
- 预测模型: LSTM 长短时记忆网络 SVR 支持向量回归



# 方法1:基于多传感器融合的刀具寿命预测方法

Achiever

# 技术方案

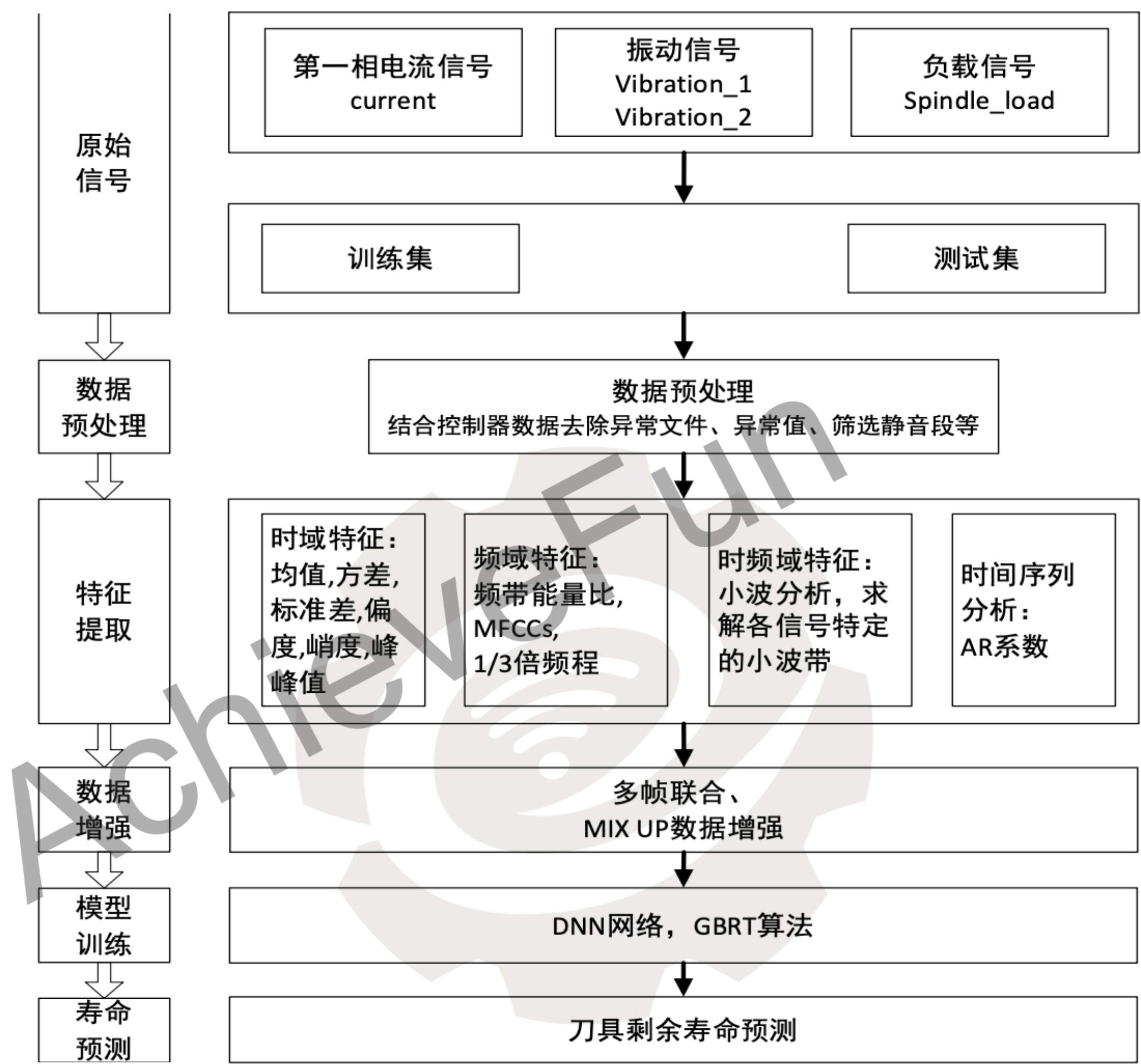


图 1-3-16 技术框架

图片来源: [http://www.aii-alliance.org/upload/202009/0903\\_084520\\_197.pdf](http://www.aii-alliance.org/upload/202009/0903_084520_197.pdf)

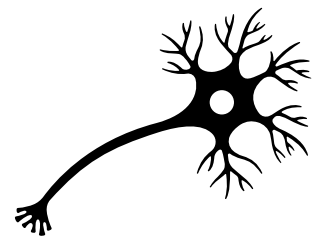


# 1 业务理解： 机床寿命监测的意义

- 在切削加工领域，刀具的状态及其寿命是加工过程中主要关注的对象，刀具寿命(磨耗程度)是影响加工工件质量的关键因子，刀具失效可能造成工件表面粗糙度和尺寸精度的下降，或造成更严重的工件报废或机床受损。
- 采取过度的保护策略又会造成刀具剩余寿命的浪费以及不必要的换刀停机时间浪费。
- 如果能够精准预测刀具的剩余寿命，将有效地优化工作排程且降低刀具采购成本。
- (1) 减少废品率以及加工成本，提升加工质量；
- (2) 降低设备停机时间，提高设备利用率以及生产效率，延长刀具寿命；
- (3) 防止机械事故，保证了加工系统的最大安全；
- (4) 工厂自动化，减少对机床的人为干预。



# 业务理解调查表



类别	问题
业务规划	项目的目标是什么？
	成功的评价指标是什么？
	目前可以投入哪些资源？ 例如： 数据、领域专家
	对应的需求、假设和约束是？
	风险和应急对策是？
	成本和收益的评估方式是？
业务问题理解	领域业务的主要内容是？
	业务流程是怎样的？
	业务存在哪些限制条件？
	有哪些基于机理分析的现有成果？
数据分析问题定义	数据挖掘的目标是？ 例如： 根因分析、预测、运筹优化、规则形式化或者自动化类、假设分析（模拟仿真不同决策下的结果）
	数据挖掘成功的评价指标是？ 例如： 数据量化的价值， 流程优化的价值
项目计划	目前的项目周期是多长时间？
	项目的主导者是谁？



# 2 数据理解

数据说明	文件类型	文件数	包含数据	物理意义
刀具加工过程中，每5分钟任取1分钟数据形成一份文件，置于Sensor中；并将对应时刻的控制器数据放入PLC中。	Sensor 传感器数据 采样频率 25600Hz	48	Vibration_1	X 方向振动
			Vibration_2	Y 方向振动
			Vibration_3	Z 方向振动
			Current	电流
	PLC 控制器数据 采样频率 不稳定于 33Hz	1	Time	记录时间
			spindle_load	主轴负载
			X	X 轴坐标
			Y	Y 轴坐标
			Z	Z 轴坐标
			csv_no	对应的文件序号

- 时序数据库存储

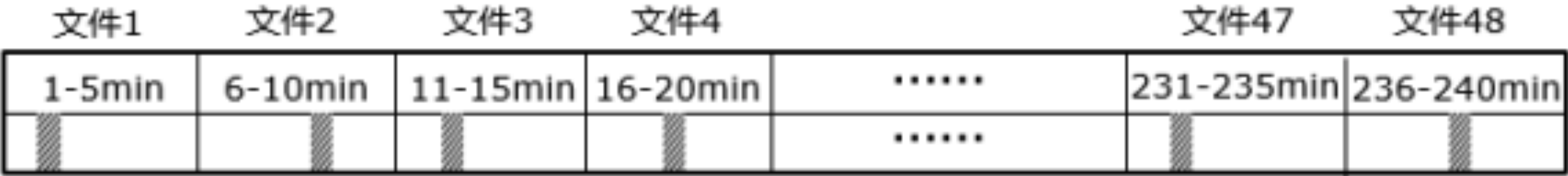


图 1-3-15 “任取 1 分钟”示意图  
图片来源: [http://www.aii-alliance.org/upload/202009/0903\\_084520\\_197.pdf](http://www.aii-alliance.org/upload/202009/0903_084520_197.pdf)

# 数据理解相关问题

类别	问题
数据源	数据来源于哪里？ 例如哪些系统
	数据的接口形式包括哪些？
	数据所属组织和获取难度如何？
数据量	历史数据量的时间跨度是怎样的？
	数据容量是怎样的？
数据基础	目前有哪些相关的数据？ 例如：工艺参数，可以提供的数据类型、关键字段、数据颗粒度、时效性，以及可以提供的数据范围（如时间、空间等）
数据变量	数据的控制量有哪些？
	数据的外生变量（有影响但无法直接控制） 有哪些？
	数据的状态量（中间状态或结果量） 有哪些？
	数据的不可观测量有哪些？ 是否为关键因子？ 是否有替代方案（例如软测量）
	数据的离线检测量有哪些？
	数据的在线检测量有哪些？
数据质量	数据的更新频率是什么？
	数据的缺失率是怎样的？
	哪些数据是人工录入的？ 存在怎样的问题？
	数据的准确率是怎样的？ 例如，传感器质量，考虑传感器的可靠性和安装范式，分析传感器的评价重复性和再现性
	历史数据与当前数据业务场景的差异（缓慢变化维度、合并、演进）



# 需要处理的数据问题

- 传感器数据含噪，质量较差，且控制器数据采样频率不稳定于 33Hz，导致传感器数据与控制器数据难以在时间上保持一致。
- 传感器数据中的 y 方向振动异常情况较多，根据加工机理进行推测，可能是传感器导线接头松动，信号时断时续导致的。
- 鉴于 y 方向的振动异常数据较多，且 y 方向振动应与 x 方向振动有相似性，所以不再考虑 y 轴方向振动，仅对电流 current，与 x 方向振动 v1 与 z 方向振动 v2 数据进行处理。
- 数据集中存在着在某次测量过程中刀具并未进行加工移动的异常文件。

# 确定任务目标

表 2-2 标签的选取

文件序号	A 刀 剩余寿命标签	A 刀 磨损比标签	C 刀 剩余寿命标签	C 刀 磨损比标签
1	235	$235/240=0.979$	180	$180/185=0.973$
2	230	$230/240=0.958$	175	$175/185=0.946$
3	225	$225/240=0.938$	170	$170/185=0.919$
4	220	$220/240=0.917$	165	$165/185=0.892$

图片来源: [http://www.aii-alliance.org/upload/202009/0903\\_084520\\_197.pdf](http://www.aii-alliance.org/upload/202009/0903_084520_197.pdf)

- 磨损比= 剩余寿命/总寿命， 表征刀具的磨损状态。
- 解决一个在含噪的数据标签中进行 0-1 之间的“磨损比”预测的拟合（回归）问题。

### 3 数据准备：数据预处理

1. 联合控制器数据中的刀具坐标，绘制刀具测量各文件时的移动轨迹，删除数据集中刀具并未进行加工的异常空文件；
  2. 将某刀的所有文件进行合并，去除数据中的极大值；
  3. 利用控制器数据中的主轴负载信号筛选出数据集中的静音段，人工选择删去部分静音段；
  4. 令帧长为 0.5s，对该刀的所有数据进行分帧；
  5. 根据现有的若干样本，生成对应的标签文件。
- 由于现有的标签并不十分精准，只有一个大致的时间范围如 1-5min, 6-10min 等，将测量时间选取为该范围的中值，将其设定为 3min, 8min 等以此类推，再统一除以刀具总寿命，以此“磨损比”作为数据的标签，对应样本生成标签文件。

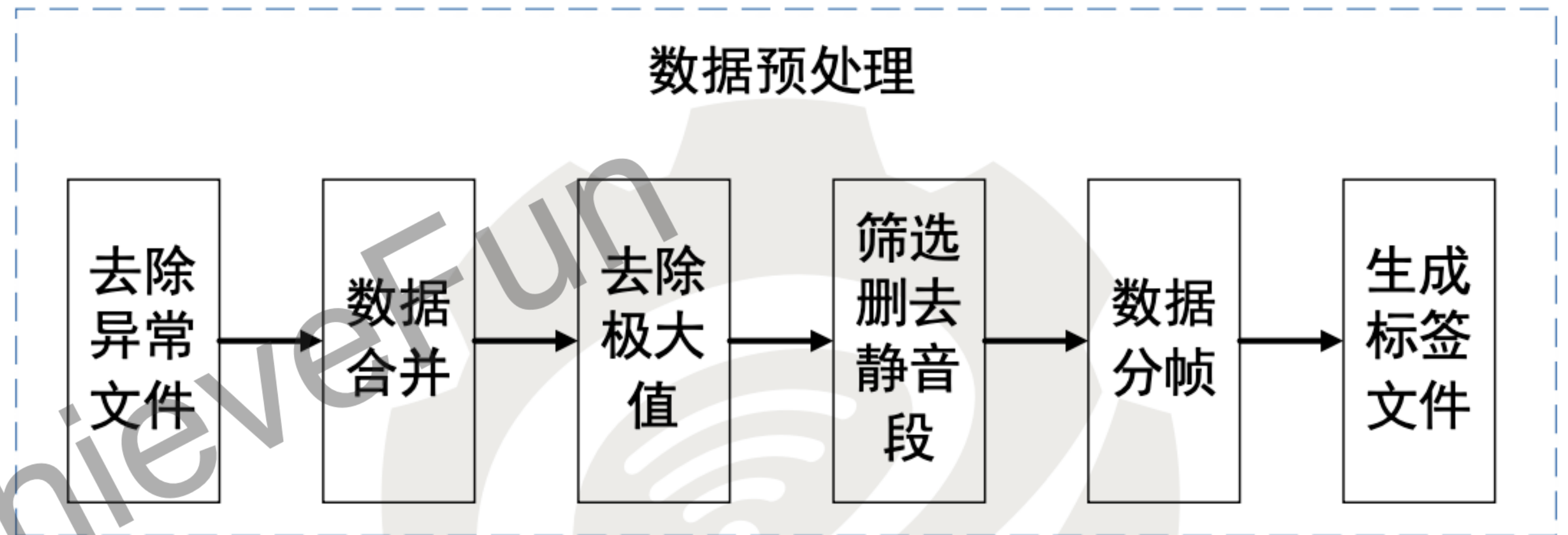


图 1-3-17 数据预处理阶段流程图

图片来源：[http://www.aii-alliance.org/upload/202009/0903\\_084520\\_197.pdf](http://www.aii-alliance.org/upload/202009/0903_084520_197.pdf)



# 3 数据准备：特征加工

- 传统刀具寿命预测方法：试图建立刀具剩余寿命与切削速度、切削深度以及进给量等加工参数之间的关系；
- 如何在只有振动、电流以及控制器数据的情况下建模？
- 数据探索：
  - 刀具的磨损特征在特定的频段内表现较为明显。
  - 研究刀具寿命与振动和电流信号的关系：包括时频域特征，以及特定小波带中提取的时域统计特征。
    - 将刀具磨损类比于结构的损伤，参考结构损伤分析方法，使用时间序列分析，计算振动信号的 AR 系数作为特征。
  - 数据集中的电流数据 c1，x 方向的振动信号 v1，z 方向的振动信号 v2，分别对上述 3 种信号进行小波包分解，对提取出的特定频带的数据进行小波带的时域指标统计。
  - 联合平均负载：从加工机理的角度，选择主轴负载为特征量。联合传感器与控制器数据，对刀具的各个文件求取主轴负载的平均值。

表 2-3 C1 各信号的特征组成

第 1-6 维	第 7 维	第 8-19 维	第 20-45 维	第 46-75 维
均值，标准差，方差，偏度，峭度，峰峰值	频带能量比	MFCCs 系数	1/3 倍频程	AR 系数

表 2-4 各样本的所有特征维度

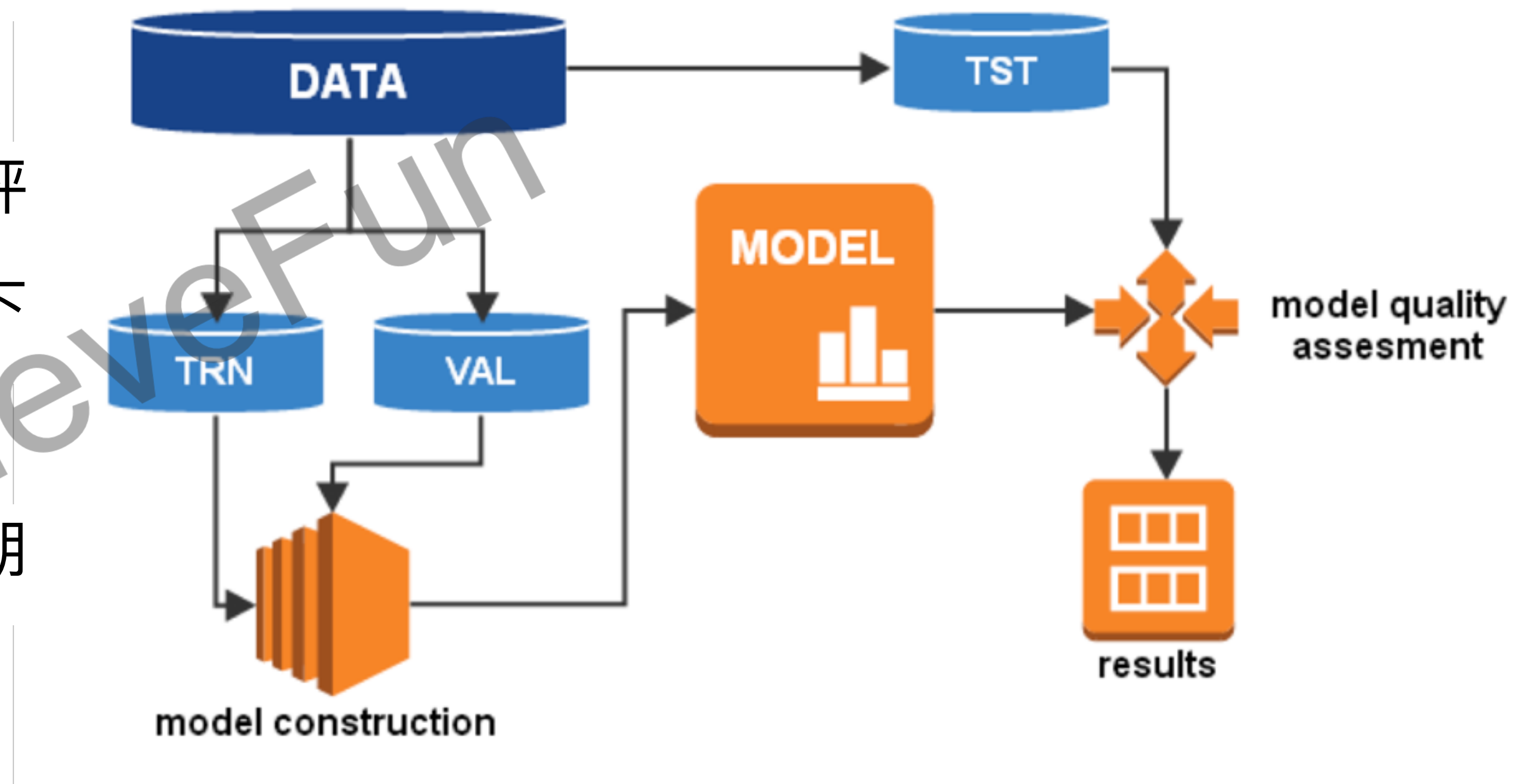
维度	1-75	76-150	151-225	226-231	232-237	238-243	244
信号提取自	c1	v1	v2	c1 小波带	v1 小波带	v2 小波带	平均负载

图片来源：[http://www.aii-alliance.org/upload/202009/0903\\_084520\\_197.pdf](http://www.aii-alliance.org/upload/202009/0903_084520_197.pdf)



# 4 模型建立：机器学习的训练和测试过程

- 训练集：训练模型的数据集。类比：课后练习题
- 验证集：也称为保持或开发集，用于评估先前训练模型在不同超参数值组合下的性能。类比：周考
- 测试集：测试模型的数据集。类别：期末考



## 4 模型建立：选择、设计

- GBRT: Gradient Boosting Regression Tree梯度提升回归树，以决策树为基本学习器的提升方法，预测性能强，作为基准。
- DNN: 深度神经网络，获取特征能力强。
- Bagging: 增强算法的泛化能力
  - 排列组合已知的 3 把刀具的训练数据，训练出多个模型。而后分别预测决赛刀具寿命，对同一文件内的样本求取均值，获得该文件的平均“磨损比”。
  - 将10个文件的平均“磨损比”转化为刀具的寿命，然后去掉预测寿命的最大值与最小值，再求取剩余结果的均值以降低预测误差，来最终预测刀具寿命。



# 优化模型

## 不足之处

- 要想准确预测出刀具的剩余寿命，需要做的不仅是一个当前状态预测，更是一个**趋势性**的分析，联合输入多个样本或许对性能会有所改善。
- 对于回归问题而言，现有数据的标签种类有限，并未均匀的覆盖待预测的0-1 区间。若能将样本进行随机组合，可以丰富填充特征与标签的多样性，或可改善模型。



## 优化方法

- 与音频信号处理相类似，在对各样本提取特征后，不再将数据进行随机打乱，而是按时序的将文件进行排列，多帧联合依次输入以便更好的利用样本的前后帧信息。
- 采用 MIX UP 方法进行数据增强。在每个 batch 输入到网络前，对该 batch 的特征与标签进行同样的线性组合，这不仅可以完善数据集的多样性，更能减少了在预测未见数据时的不确定性，提高系统的泛化能力。

# Mixup

Image Data Augmentation

## Mixup

Introduced by Zhang et al. in [mixup: Beyond Empirical Risk Minimization](#)

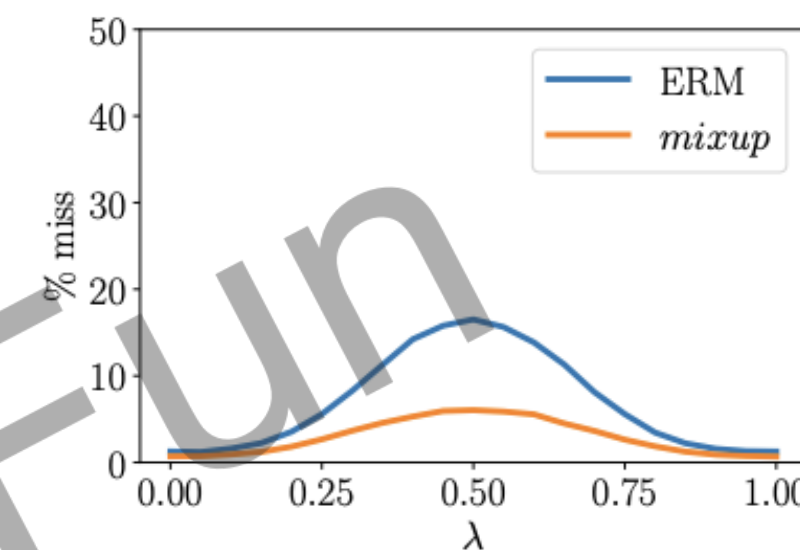
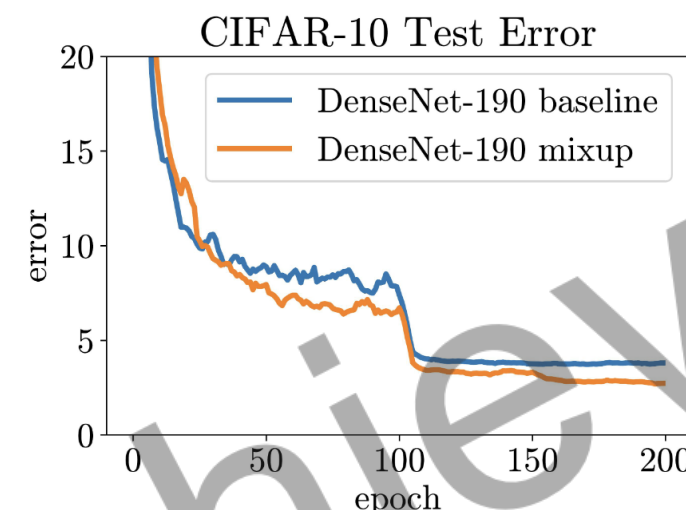
**Mixup** is a data augmentation technique that generates a weighted combination of random image pairs from the training data. Given two images and their ground truth labels:  $(x_i, y_i), (x_j, y_j)$ , a synthetic training example  $(\hat{x}, \hat{y})$  is generated as:

$$\hat{x} = \lambda x_i + (1 - \lambda)x_j$$

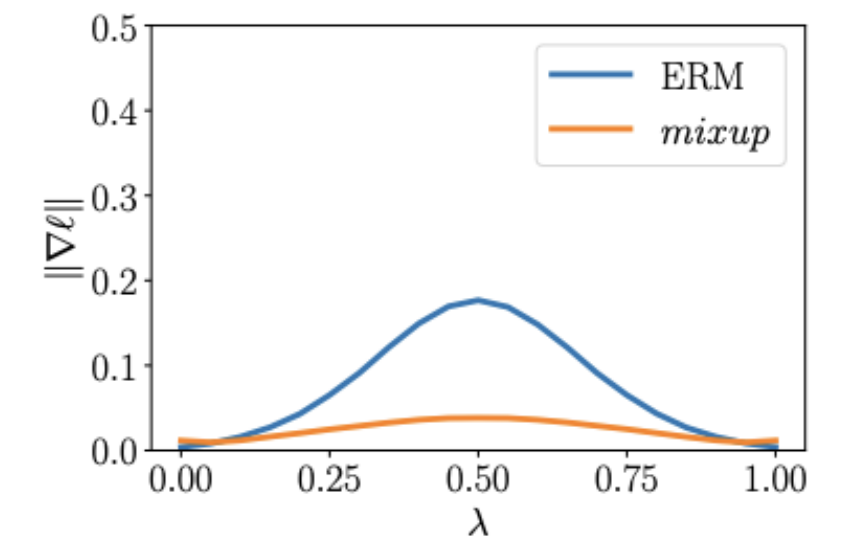
$$\hat{y} = \lambda y_i + (1 - \lambda)y_j$$

where  $\lambda \sim \text{Beta}(\alpha = 0.2)$  is independently sampled for each augmented example.

Source: [mixup: Beyond Empirical Risk Minimization](#)



(a) Prediction errors in-between training data. Evaluated at  $x = \lambda x_i + (1 - \lambda)x_j$ , a prediction is counted as a “miss” if it does not belong to  $\{y_i, y_j\}$ . The model trained with *mixup* has fewer misses.



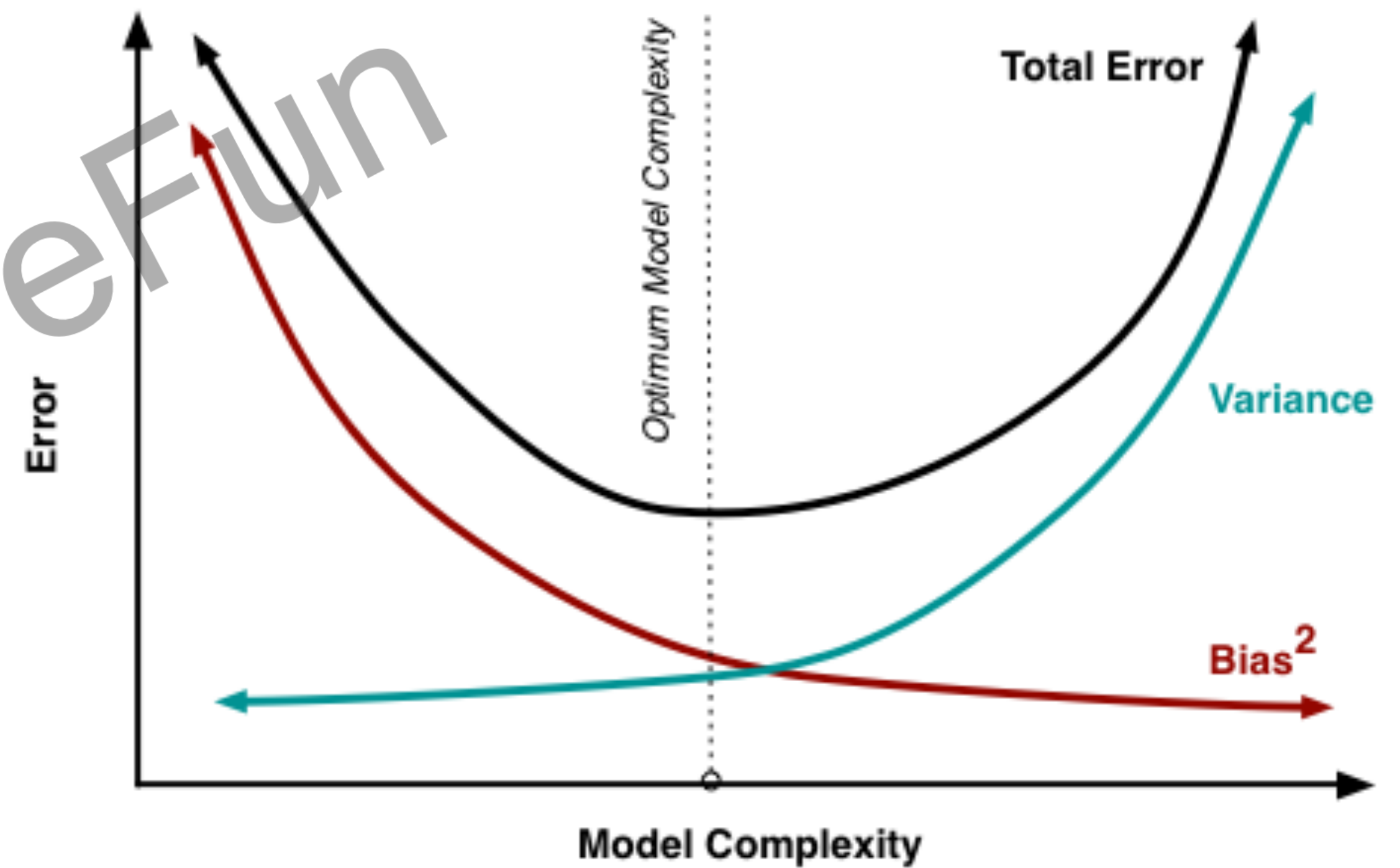
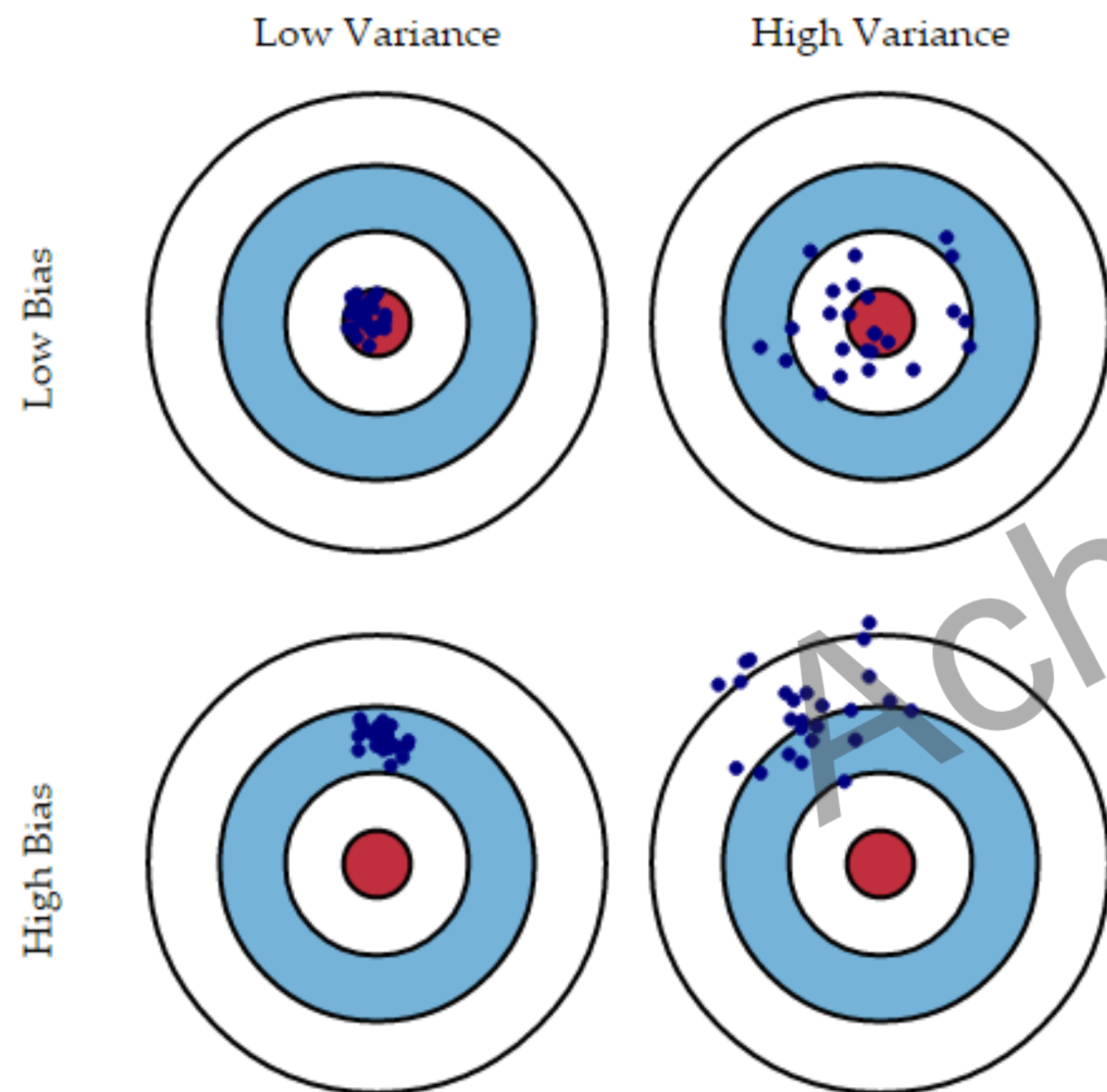
(b) Norm of the gradients of the model w.r.t. input in-between training data, evaluated at  $x = \lambda x_i + (1 - \lambda)x_j$ . The model trained with *mixup* has smaller gradient norms.

- Mixup邻域分布可以是一种数据增强方式，它令模型在处理样本和样本之间的区域时表现为线性。这种线性建模减少了在预测训练样本以外的数据时的不适应性。
- 左图是ERM 和 mixup 在 CIFAR 实验中的测试误差，展示了最佳 ERM 和混合模型的测试误差变化。
- 右图显示了在CIFAR-10数据集上用mixup和ERM两个方法训练的两个神经网络模型的平均表现。两个模型有相同的结构，使用相同的训练过程，使用mixup训练的模型在训练样本之间的模型预测和梯度模值更加稳定。因为Mixup对离散样本空间进行连续化，提高邻域内的平滑性。<https://www.cnblogs.com/lvdongjie/p/14056273.html>作者：张宏毅回答的一些问题：链接：<https://www.zhihu.com/question/67472285/answer/256651581>

图片来源：<https://paperswithcode.com/method/mixup>



# 机器学习模型的问题

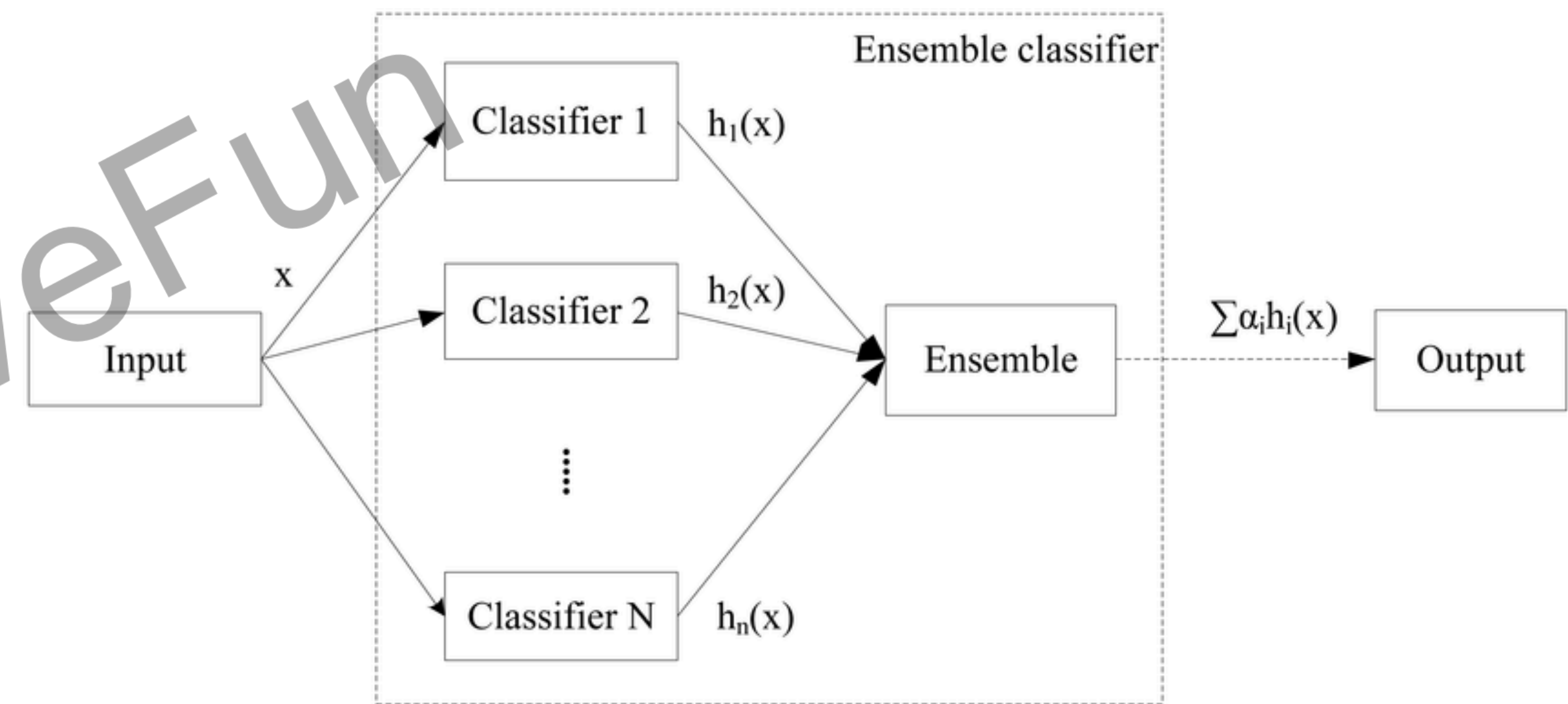


图片来源: <https://towardsdatascience.com/bias-variance-tradeoff-e8995c42b55b>

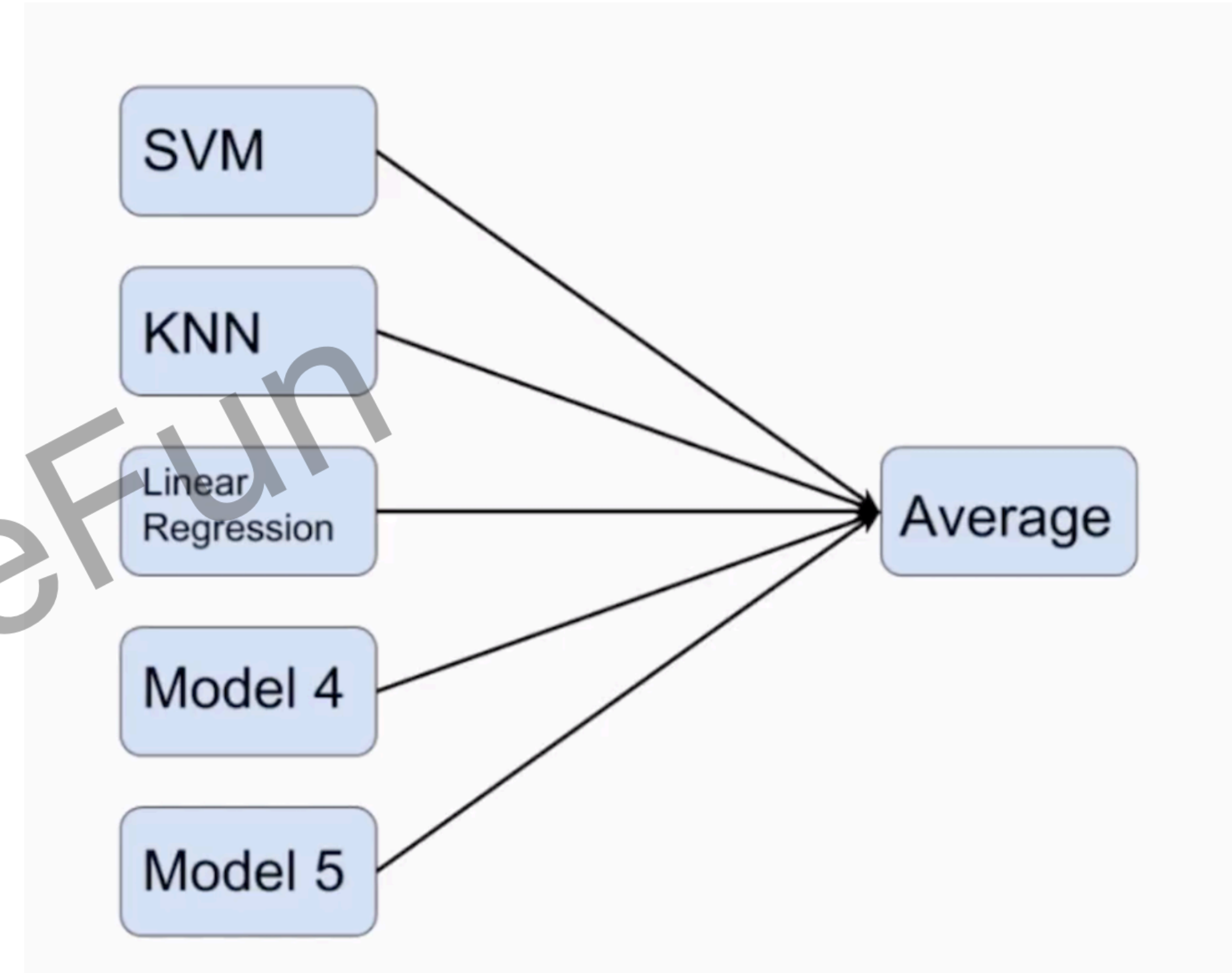
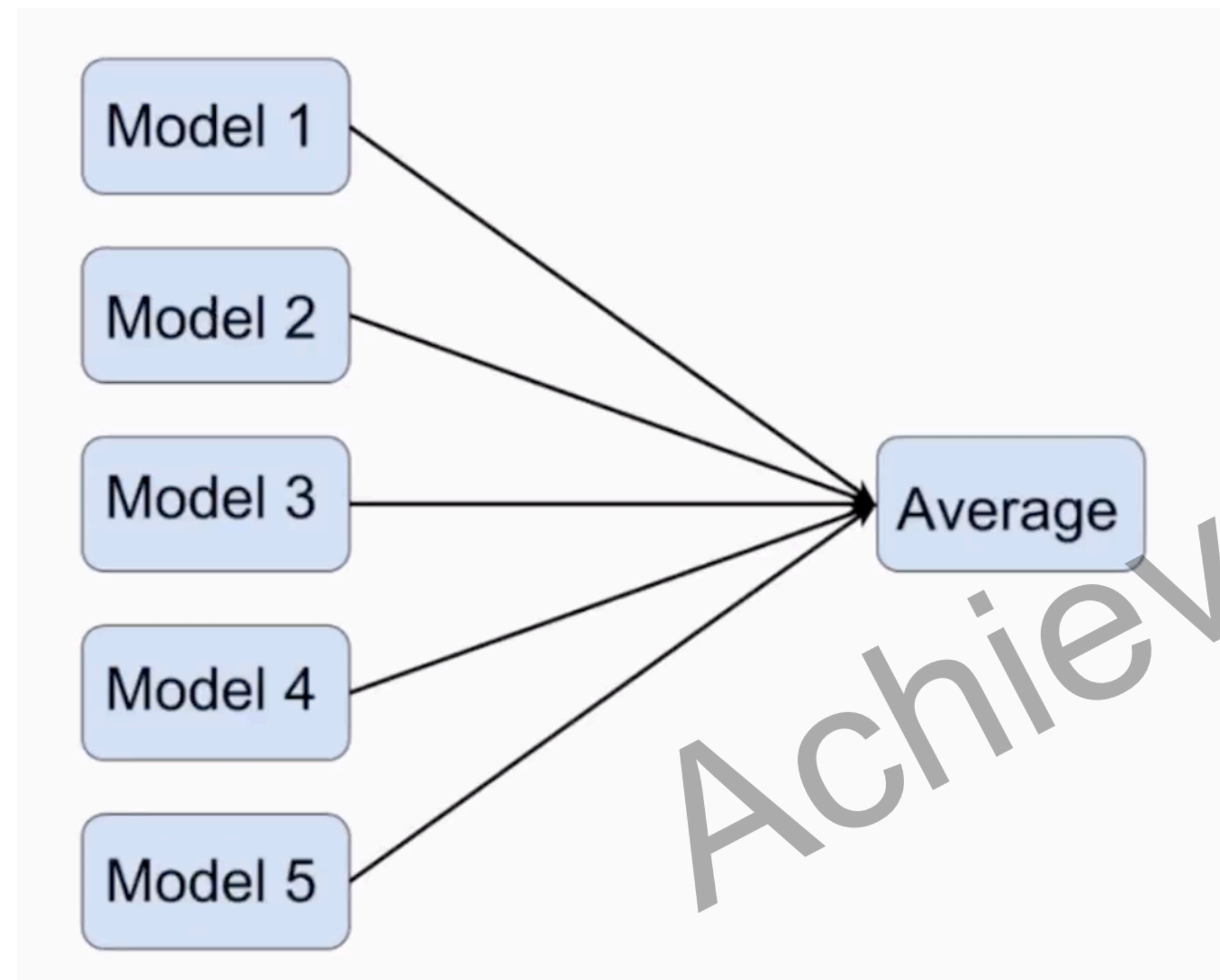


# 集成学习

- 在单个数据集上使用多个协同工作的模型称为**集成 (Ensemble)**，结合多个模型的决策提高单个模型的泛化性和鲁棒性。3个臭皮匠，1个诸葛亮。：)
- 偏差和方差的权衡: 一组预测模型的性质，其中参数估计中具有较低偏差的模型在样本之间具有较高的参数估计方差，反之亦然。
- 减少了方差: 许多模型的总体意见比其中一个模型的单一意见噪音小。
- 使最终的组合模型不易受与局部最小值相对应的给定值的影响。

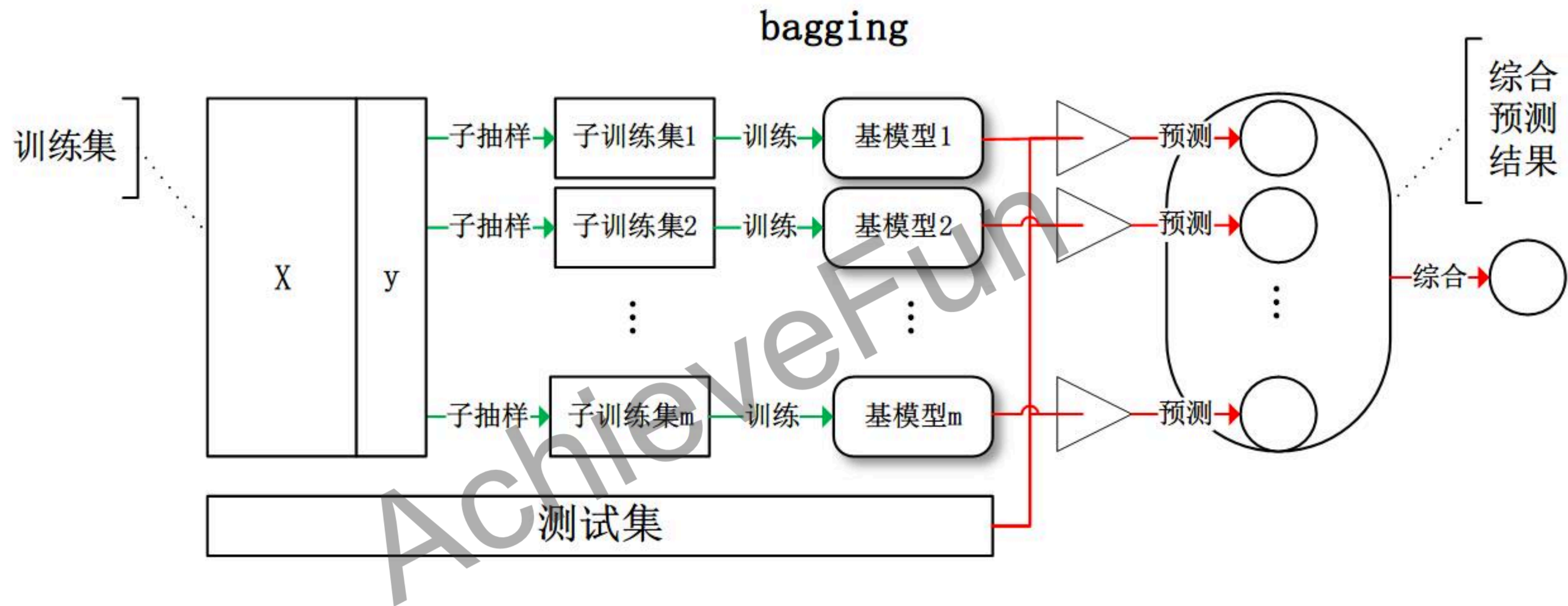


# 集成学习的平均方法



- 平均方法：独立地构建几个估计器，然后平均它们的预测。平均而言，组合估计器通常比任何单基估计器都要好，因为它的方差减小了。

# 集成学习的高级技术： Bagging

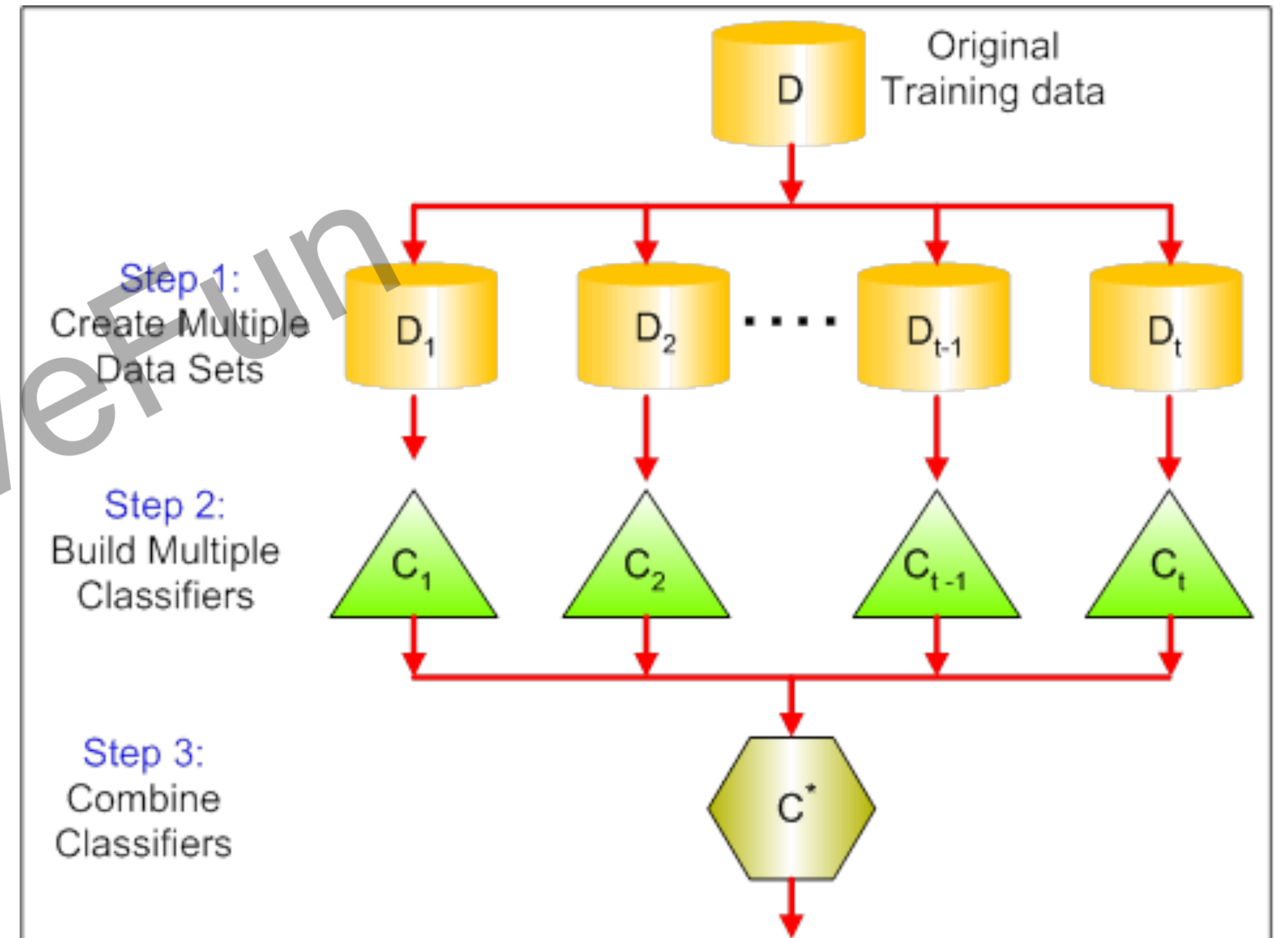


- 平均的方法：从训练集进行子抽样组成每个基模型所需要的子训练集，对所有基模型预测的结果进行综合产生最终的预测结果



# Bagging袋装法

- 作用：减少过拟合
- 适用于强大和复杂的模型
- 过程：在原始训练集的随机子集上构建黑盒估计器的多个实例，然后聚合它们的各自的预测以形成最终预测。
- 被用作减少基本估计器（例如决策树）方差的一种方法，将随机化引入其构建过程，然后将其组合起来。简单的方法改进单个模型，无需调整底层。



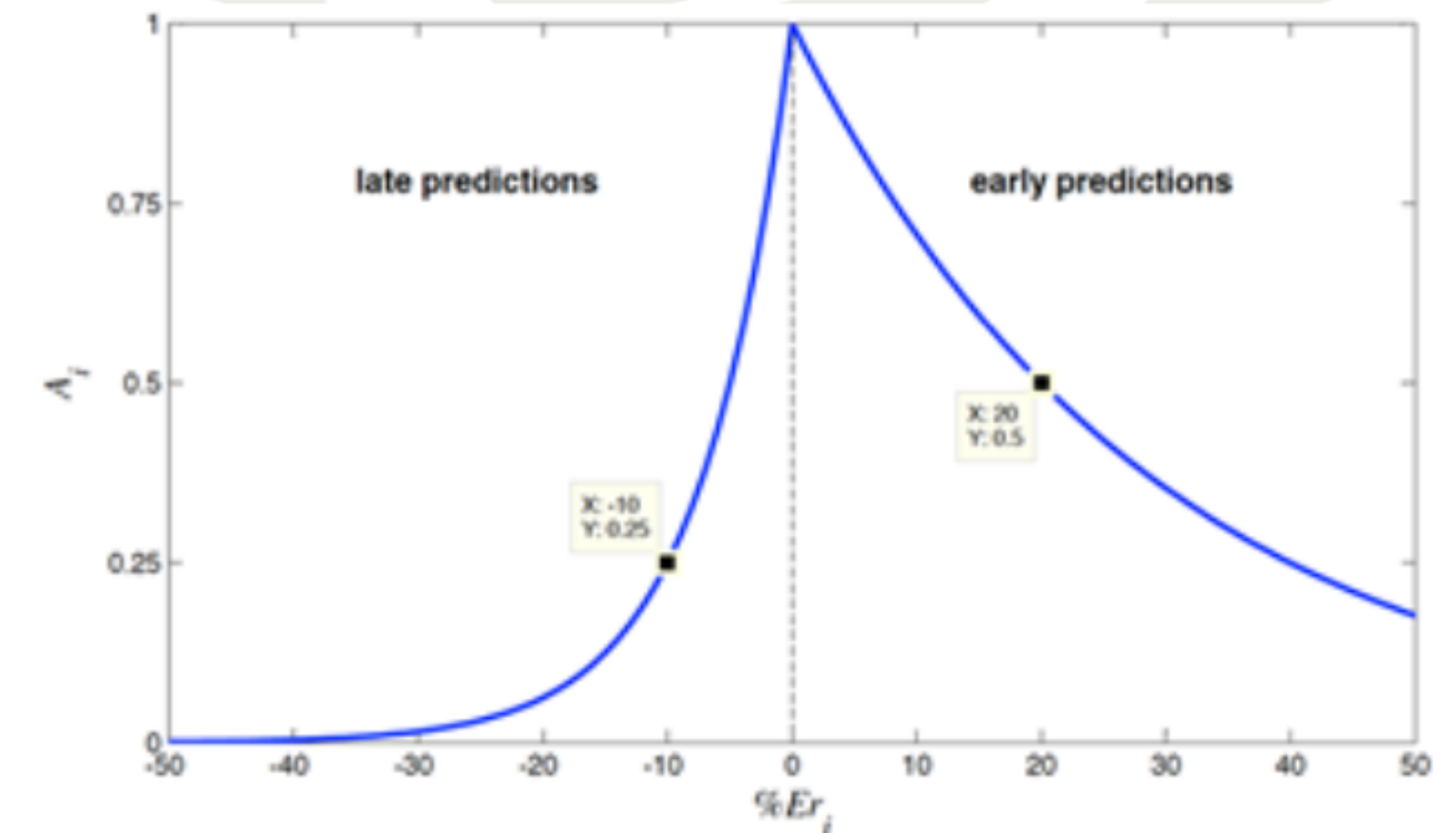
# 5 模型评价

赛题评分方法如下：

$$A_i = \begin{cases} \exp^{-\ln(0.5) \cdot \left(\frac{Er_i}{5}\right)} & , \text{if } Er_i \leq 0 \\ \exp^{+\ln(0.5) \cdot \left(\frac{Er_i}{20}\right)} & , \text{if } Er_i > 0 \end{cases}$$

以上计算公式中：

1.  $i$  是测试的第  $i$  把刀；
2.  $Er_i$  = 真实值（第  $i$  把刀） - 预测值（第  $i$  把刀）。当预测值  $>$  真实值时，为过晚预测（late prediction）；当预测值  $<$  真实值时，为过早预测（Early prediction）；
3. 每个预测的分值以百分制计算。例如：当  $A$  为 1 时，得分为 100； $A$  为 0.51 时，得分为 51；
4. 竞赛评分系统给出选手得分分值为所有测试组得分的平均值。



图片来源：[http://www.aii-alliance.org/upload/202009/0903\\_084520\\_197.pdf](http://www.aii-alliance.org/upload/202009/0903_084520_197.pdf)

# 模型预测结果

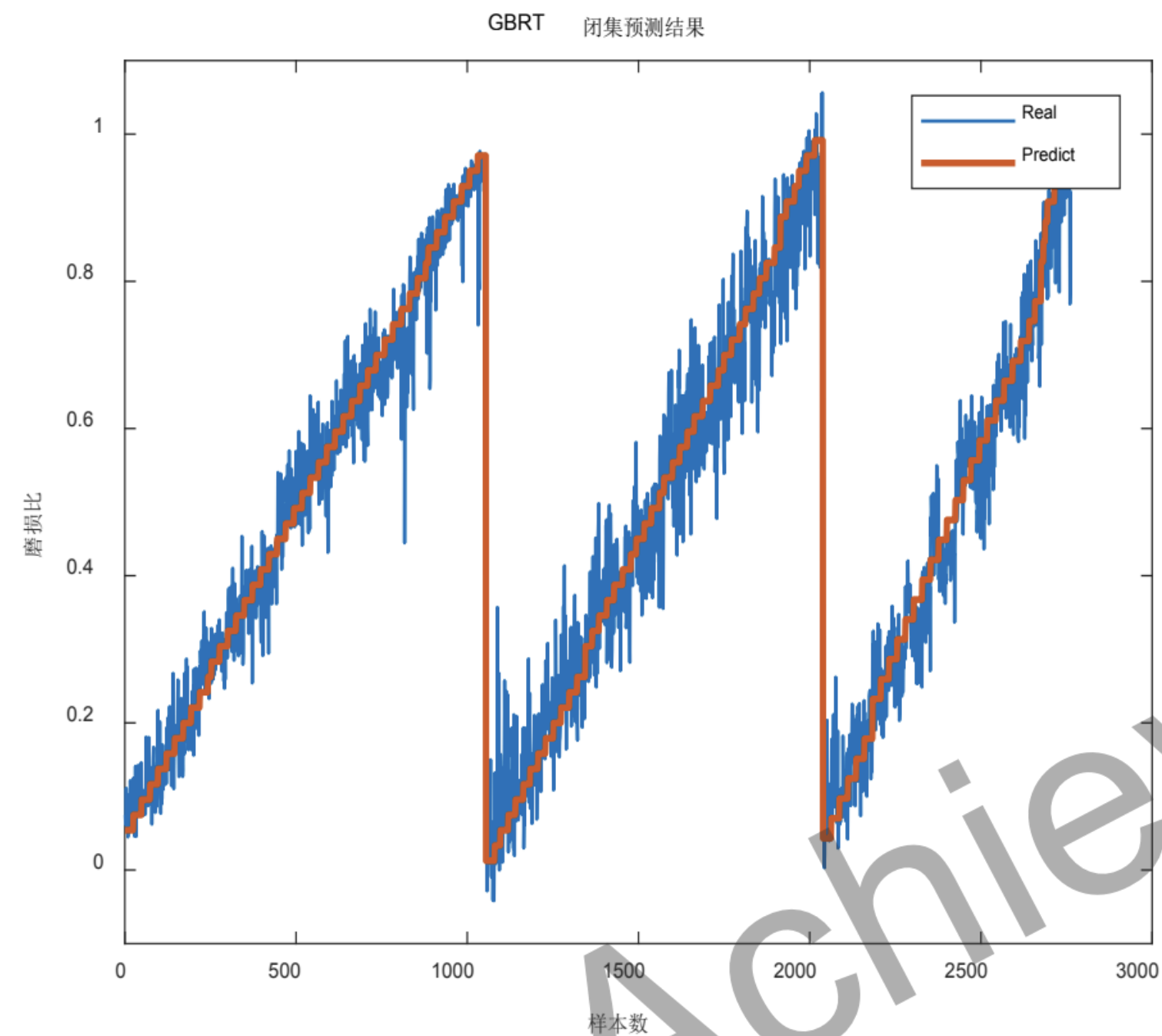


图 1-3-18 GBRT 算法预测结果

将数据集等间距采样分为 5 份，4 份做训练，1 份为测试。

以 GBRT 算法为基线，各决策树最大深度为 5，最多 500 棵树，学习率 0.1

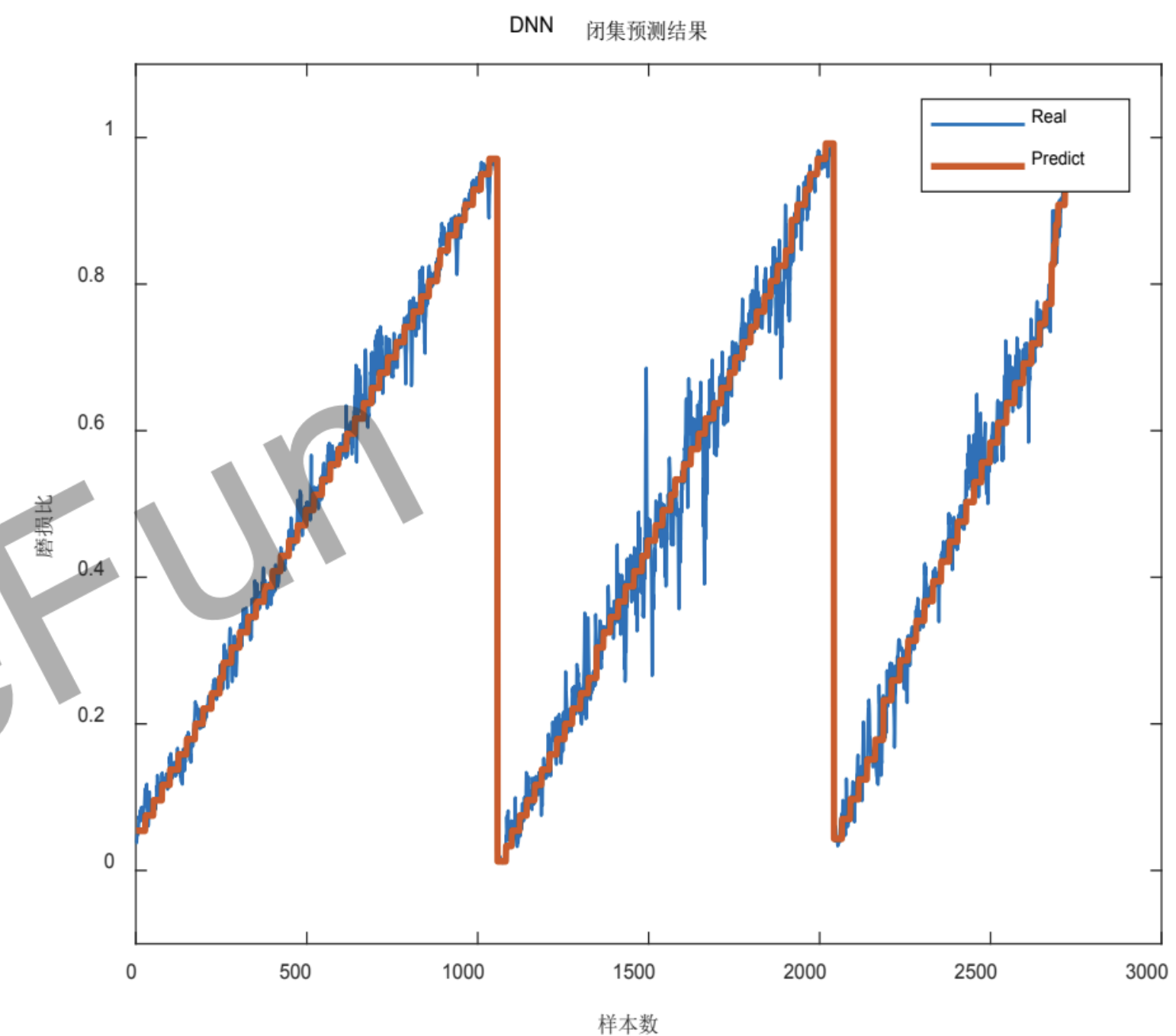


图 1-3-19 DNN 网络预测结果

利用4 层 DNN 模型进行训练，各层节点数分别为 2048，2048，1024，1024；采用 L2 正则化，基于 MAE 指标，利用 Adam 方法进行优化。

图片来源：[http://www.aii-alliance.org/upload/202009/0903\\_084520\\_197.pdf](http://www.aii-alliance.org/upload/202009/0903_084520_197.pdf)



# 增强算法泛化能力的模型预测结果

- 基于Bagging方法，排列组合已知标签的三把刀的训练数据，训练出多个预测模型，分别对决赛刀具进行预测后，采用**求均值**的方式减小随机噪声对预测的影响。
- 进一步优化方法：采用中值或均值滤波等操作，去除噪声，使预测更加准确。

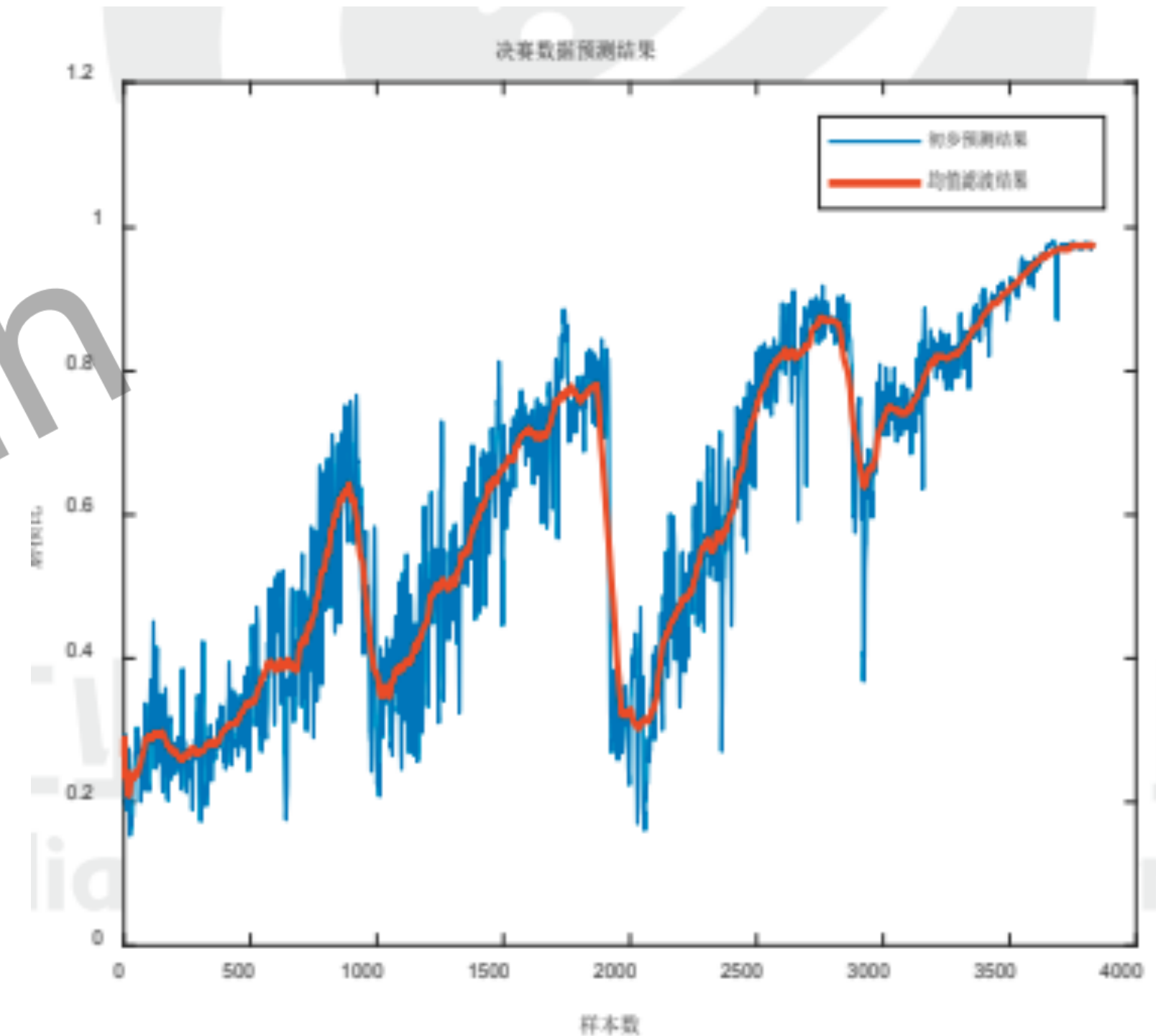


图 1-3-20 模型对决赛数据的预测结果

# 运算结果

功能	所用硬件	总耗时	每个样本耗时
数据预处理与特征提取	普通 i7 处理器	4分半	200ms
网络模型训练与预测	容天 Multi GPU SCS4450 超算服务器	预测耗时5-6S	几个ms

# 6 模型部署

- 模型部署的自动化：提高效率
- 实施和运行中的问题和解决方法：自动修正模型？
- 部署后的持续优化：提高精度、扩大适用范围以及增加知识。



# 方案总结

- 目标：明确为解决利用含噪标签进行 0-1 之间“磨损比”预测的回归问题。
- 步骤：
  - 从分析原始数据出发，有针对性的对数据进行了清洗、分帧等预处理。
  - 对预处理后的数据进行各种时域、频域和时频域特征提取，**将刀具磨损类比于结构损伤**，引入**时间序列分析技术**，并合理地利用控制器数据中的主轴负载信息，将其均值加入到特征集中，以弥补不确定的工况带来的影响。
  - 深入探讨特征，试图找出各类特征与磨损量之间的映射关系。
  - 在最开始进行模型训练时得到的结果与选定的 Baseline 也就是 GBRT 算法相比并不理想，又引入多帧联合与 MIX UP 数据增强模块，提升了系统的预测性能与泛化能力。
- 改进部分：
  - 数据预处理阶段时人为介入了静音段的删除工作，可能会对模型的泛化能力产生一定影响；
  - 特征集中的部分维度可能并不包含有用信息，导致了一定冗余；
  - 后续可以使用适合处理序列信息的**Transformer**。

# 方法2:基于 DTW 退化特征和 CNN+SVR 网络的刀具剩余寿命预测

# 方案核心

- 查看 [http://www.aii-alliance.org/upload/202009/0903\\_084520\\_197.pdf](http://www.aii-alliance.org/upload/202009/0903_084520_197.pdf) p54-p59页
- 提取符合刀具退化过程的退化指标。  
DWT: 动态时间规整算法 (DWT: 比较不同长度的两个数组或时间序列之间的相似性或计算两者间的距离) 计算数据之间的相似度, 以提取以 5 分钟为单位的退化指标。
- 使用卷积神经网络学习训练集的信号 - 退化指标之间的映射关系, 训练模型, 预测测试集信号的退化指标。

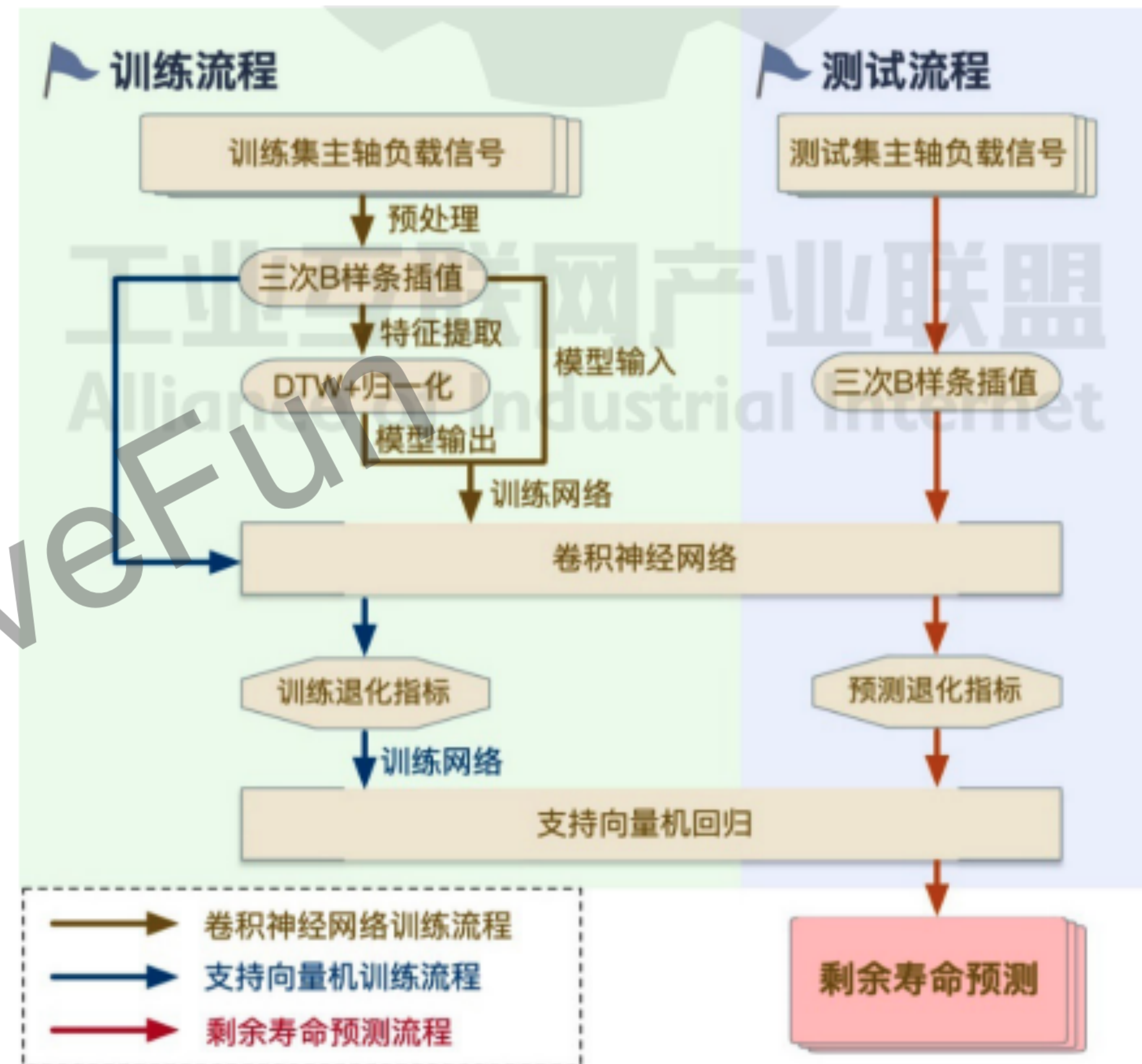


图 1-3-30: 算法流程图

图片来源: [http://www.aii-alliance.org/upload/202009/0903\\_084520\\_197.pdf](http://www.aii-alliance.org/upload/202009/0903_084520_197.pdf)



