

# 轆铁进展

夏商



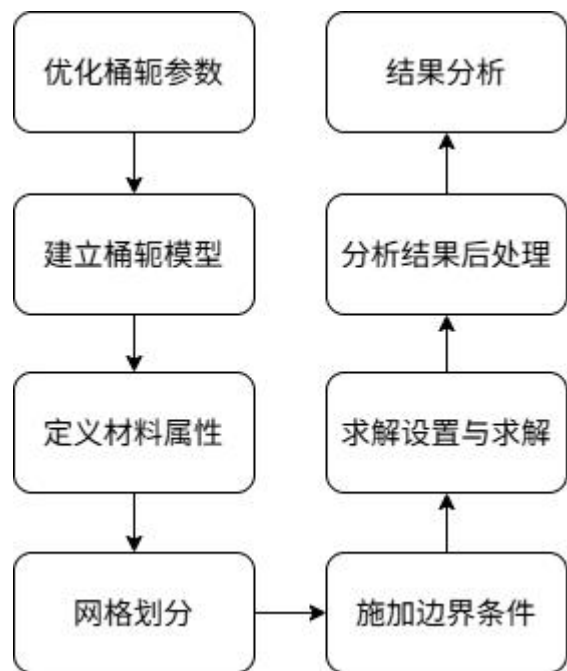
中国科学院高能物理研究所  
*Institute of High Energy Physics*  
*Chinese Academy of Sciences*

# 目录

- 研究背景与意义
- 模型选择与介绍
- 横向对比
- 代理模型驱动的结构优化
- 总结

# 研究背景与意义

## 桶辘FEA流程



### 优点:

1.分析精度高

### 缺点:

1.优化流程耗时, 不断重复性分析流程, 效率低

2.已有分析结果无法有效利用

3.多参数优化难度大

## 基于代理模型的桶辘优化流程



### 优点:

1.代理模型替代FEA分析, 计算效率高

2.充分利用已有的分析数据

3.便于多参数优化

### 缺点:

1.预测精度会略低于有限元分析

2.依赖数据集的质量

**意义:** 用1500组的有限元计算数据 (输入: 端部法兰厚度、宽度; 输出: 桶辘最大变形量、最大应力值), 训练高效近似预测模型替代有限元分析, 在保证一定精度的前提下, 降低计算耗时、提升优化效率。

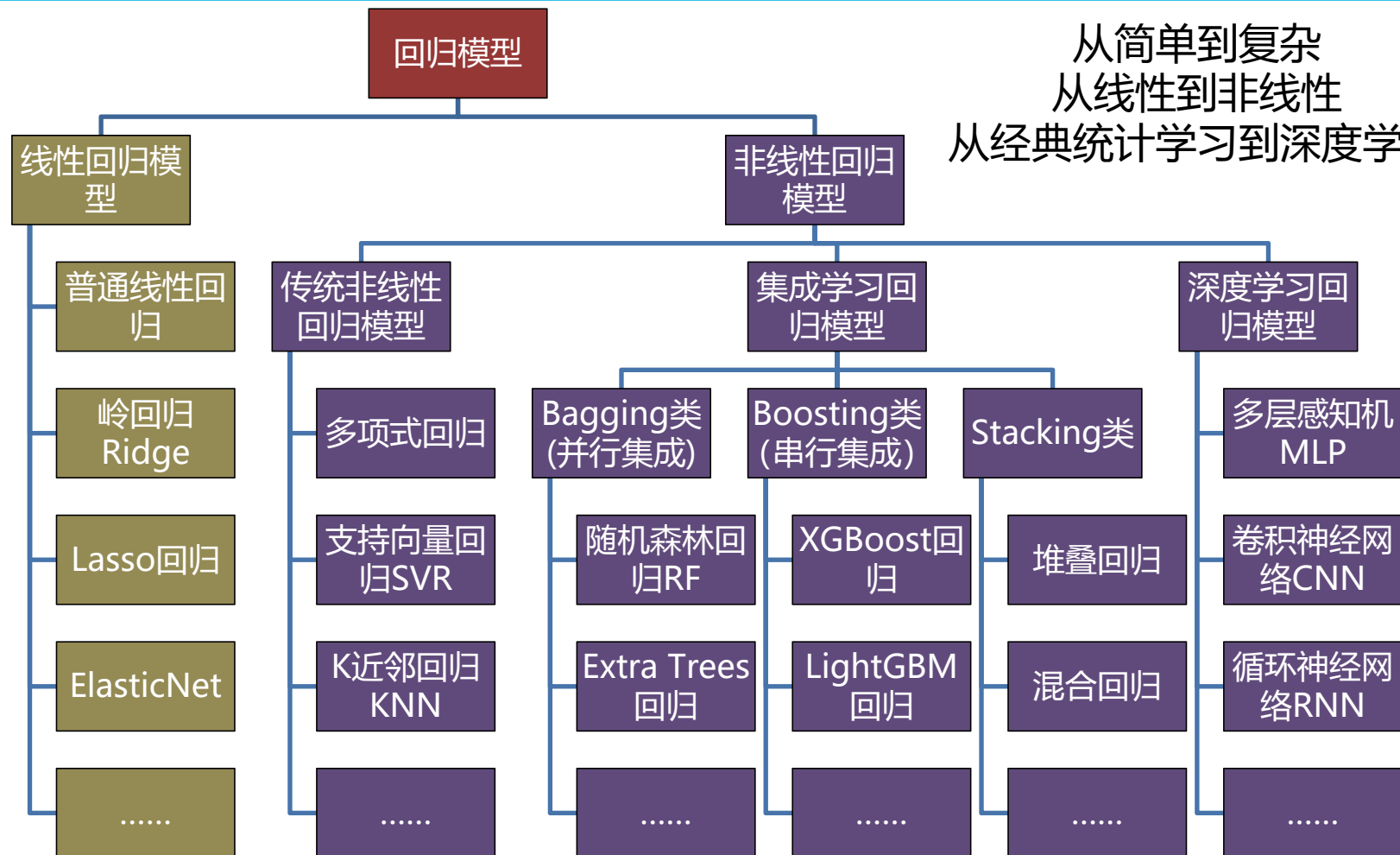
# 模型选择

## 机器学习类别(按任务类型)

1. 分类 Classification  
——判断一个样本类别
2. **回归 Regression**  
——**预测连续的数值变量**
3. 聚类 Clustering  
——将相似的数据点归为一组
4. 降维 Dimensionality Reduction  
——将高维数据映射到低维空间

## 选取的回归模型:

1. 岭回归(Ridge)模型  
——传统线性回归 (基准模型)
2. 支持向量机回归(SVR)模型  
——传统非线性回归
3. 极端梯度提升(XGBoost)模型  
——集成学习回归
4. 多层感知机(MLP)模型  
——深度学习回归



从简单到复杂  
从线性到非线性  
从经典统计学习到深度学习

## 核心场景:

小样本低维有限元数据  
代理模型加速仿真

## 数据集: 1500 条小样本仿真数据

2个低维输入特征(端部法兰厚度/支撑宽度)  
双目标预测(变形/应力)(平滑+强非线性)

# 模型选择

## 核心场景:

小样本低维有限元数据  
代理模型加速仿真

## 数据集: 1500 条小样本仿真数据

2个低维输入特征(端部法兰厚度/支撑宽度)  
双目标预测(变形/应力)(平滑+强非线性)

## 选取的回归模型:

- 1.岭回归(Ridge)模型  
——传统线性回归 (基准模型)
- 2.支持向量机回归(SVR)模型  
——传统非线性回归
- 3.极端梯度提升(XGBoost)模型  
——集成学习回归
- 4.多层感知机(MLP)模型  
——深度学习回归

## 核心目的

### 1.贴合研究场景

建立基准线: 用最简单的线性模型 (Ridge回归) 定一个效果下限, 后续所有非线性模型, 必须超过这个基准才算合格。

### 2.全面对比+择优适配

覆盖不同的非线性拟合逻辑, 找到最适配我们小样本 (1500条)、低维度 (2个输入特征)、物理非线性有限元数据集的最优模型。

### 3.验证不同模型的场景适配性

同时测试4类模型对变形 (低非线性的平滑响应) 和应力 (高非线性的离散响应) 的拟合能力。

# 岭回归(Ridge)模型

## 预测原理

线性回归的一种改进，主要是线性回归求得的系数可能会过大。因此，岭回归在线性回归的基础上，在线性回归模型的损失函数中加入二范正则项，以惩罚过大的系数带

## 预测表达式

$$y = \omega_1 x_1 + \omega_2 x_2 + \dots + \omega_n x_n + b$$

## 损失函数

$$L(W) = \sum_{i=1}^N (y - \hat{y})^2 + \alpha \sum_{j=1}^P (\omega_j)^2$$

残差平方和 + L2正则化惩罚项

超参数	意义
$\alpha$	惩罚系数

找到一组最优的权重系数，拟合出公式，让预测结果和实际结果的总误差最小，同时加了一个安全惩罚项，比普通线性回归更稳，不会学偏。

## 优势

- 1.结构简单，可解释性强
- 2.低维小样本，无过拟合
- 3.可作为基准方案

## 劣势

- 1.只能拟合线性关系，无法捕捉结构力学中的非线性规律，能力有限

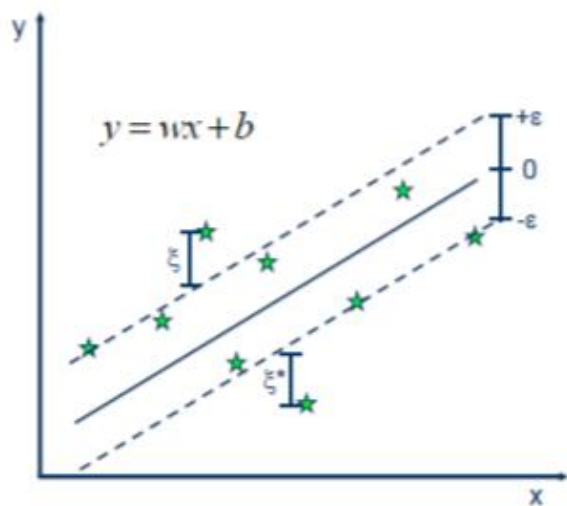
## 场景适配性

适用于本研究作为线性基准模型，用于界定预测精度下限、验证问题的非线性本质，为后续非线性模型对比提供参照。

# 支持向量机回归(SVR)模型

## 预测原理

在线性函数的两侧构建宽度为 $2\varepsilon$ 的“间隔带”，间隔带内的样本点视为预测正确（损失为零），间隔带外的样本点通过支持向量计算损失，最终通过最小化总损失与最大化间隔的平衡实现模型优化。针对非线性问题，SVR采用核函数将输入空间映射到特征空间进行线性回归。



**核函数：**  
径向基函数RBF

超参数	意义
C	惩罚系数
gamma	控制核函数的范围
epsilon	误差容忍带

## 预测表达式

$$f(x) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) K(x_i, x) + b$$

## 优势

1. 在小样本、低维、强非线性场景下泛化能力突出，适配数据集
2. 对连续、平滑的物理场响应（如结构变形）拟合精度极高
3. 抗噪声干扰能力强，模型稳定性好

## 劣势

1. 在高维特征与大规模数据场景下训练效率较低。（本研究场景下无）

## 场景适配性

与本研究1500条小样本、2维低维输入、变形场平滑连续的特点高度匹配。

# 极端梯度提升树(XGBoost)模型

## 预测原理

第一棵小树先用厚度、宽度做一个初步预测，算出和有限元真实值的误差；第二棵小树专门拟合这个误差；第三棵再补第二棵的误差，以此类推，最后把所有小树的结果加起来，就是最终的预测值。

## 预测表达式

$$\hat{y}_i = \phi(x_i) = \sum_{k=1}^K f_k(x_i), f_k \in \mathcal{F}$$

## 损失函数

$$L(\phi) = \sum_{i=1}^N l(\hat{y}_i, y_i) + \sum_{k=1}^K \Omega(f_k)$$

$$\Omega(f_k) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j^2$$

L1正则项 + L2正则项

超参数	意义
max depth	树的最大深度
learning rate	学习率
colsample bytree	列采样比例
n estimators	树的数量
reg alpha	L1 正则化权重
reg lambda	L2 正则化权重
subsample	行采样比例

## 优势

1. 非线性强、局部突变明显的力学响应（如应力集中、应力突变）拟合能力强
2. 可以特征交互，可输出特征重要性
3. 可解释性优于深度学习模型

## 劣势

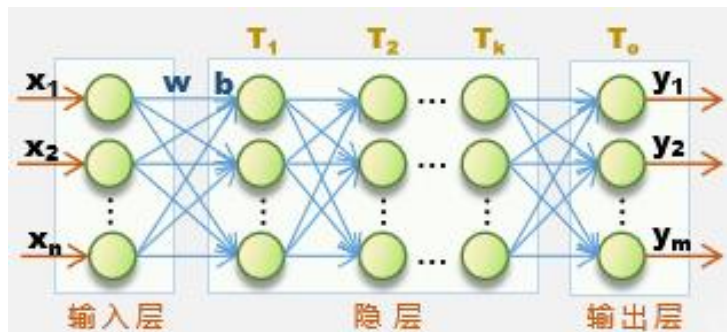
1. 小样本条件下存在一定过拟合风险

## 场景适配性

适用于本研究中非线性、突变性更强的应力预测，可作为备选方案，与SVR形成互补。

# 多层感知机(MLP)模型

## 预测原理



MLP由输入层、隐层、输出层组成(其中隐层可以是有多层的),各层则由神经元构成。MLP以前馈方式进行计算,即每层的输出作为下层的输入,不断套娃,直到输出层。

每层的计算公式:  $y = T(WX + B)$

其中,  $T$ 为该层的激活函数

## 损失函数

$$E(W, b) = \frac{1}{m} \sum_{i=1}^m \frac{1}{k} \sum_{j=1}^k (\hat{y}_{ij} - y_{ij})^2$$

## 训练过程:

多层神经网络通过“加权求和做线性组合+激活函数引入非线性”,一层一层叠加,最终可以逼近任意复杂的非线性函数。

## 优势

- 1.数据足够,能拟合任何复杂关系
- 2.具备良好的可扩展性,可方便扩展至多输入特征、多物理场耦合等复杂场景

## 劣势

- 1.依赖数据集
- 2.调参复杂
- 3.可解释性差

## 场景适配性

用于验证深度学习方法在本仿真数据集上的适用性,兼顾当前双目标预测需求与未来多特征、多场耦合仿真代理模型的扩展潜力。

# 横向对比

## 回归模型:

- 1.岭回归(Ridge)模型
- 2.支持向量机回归(SVR)模型
- 3.极端梯度提升(XGBoost)模型
- 4.多层感知机(MLP)模型

## 验证方法:

通过训练集训练好的模型，用测试集的输入值来预测桶轭的最大响应值，然后将预测的最大响应值和实际值进行比较。

常用  
评估  
指标

指标	公式	含义	取值范围	评价标准
决定系数 ( $R^2$ )	$R^2 = 1 - \frac{\sum_{k=1}^m (P_k - Q_k)^2}{\sum_{k=1}^m (\bar{Q}_k - Q_k)^2}$	衡量模型对数据变异的解释程度，反映预测值与真实值的拟合优度	(0,1)	越接近 1，模型拟合效果越好；越接近 0，拟合效果越差
平均绝对误差 (MAE)	$MAE = \frac{1}{m} \sum_{k=1}^m  P_k - Q_k $	预测值与真实值绝对误差的平均值，反映平均预测偏差	$[0, +\infty)$	值越小，预测精度越高
均方根误差 (RMSE)	$RMSE = \sqrt{\frac{1}{m} \sum_{k=1}^m (P_k - Q_k)^2}$	均方误差的算术平方根，衡量预测值与真实值的离散程度	$[0, +\infty)$	值越小，模型拟合精度越高

# 横向对比

模型	核心拟合逻辑	最大变形预测 (测试集)	最大应力预测 (测试集)	综合效果排名
Ridge 回归	线性超平面拟合	R <sup>2</sup> =0.9648 MAE=0.569 RMSE=0.0719	R <sup>2</sup> =0.8935 MAE=4.5221 RMSE=5.7887	4 (基准垫底)
SVR	支持向量 + 核 函数安全管道拟合	R <sup>2</sup> =0.9998 MAE=0.046 RMSE=0.0056	R <sup>2</sup> =0.9651 MAE=1.8492 RMSE=3.3134	1 (综合最优)
XGBoost	多棵决策树迭代 补误差	R <sup>2</sup> =0.9924 MAE=0.0251 RMSE=0.0334	R <sup>2</sup> =0.9569 MAE=2.5838 RMSE=3.6828	3
MLP	多层神经网络非 线性映射	R <sup>2</sup> =0.9933 MAE=0.0233 RMSE=0.0234	R <sup>2</sup> =0.9661 MAE=2.1556 RMSE=3.2662	2

## 1. 线性模型无法满足强非线性力学响应的预测需求

Ridge回归在变形预测里还能拿到0.96的R<sup>2</sup>，但在非线性更强的应力预测里，R<sup>2</sup>直接掉到0.89，和非线性模型差距显著，验证了必须采用非线性模型的判断。

## 2. SVR是该场景的最优解

在变形预测里，它对连续平滑的物理规律拟合能力碾压其他模型；在应力预测里，它仍保持微弱优势，综合表现最稳定、精度最高，适配小样本、低维度的有限元加速场景。

## 3. XGBoost和MLP是优秀的备选方案

在非线性更强的应力预测里，两者和SVR的差距极小，满足需求。其中XGBoost通用性强、可解释性好；MLP扩展性强，适合后续更复杂的研究拓展。

# 横向对比

## 最大变形量预测

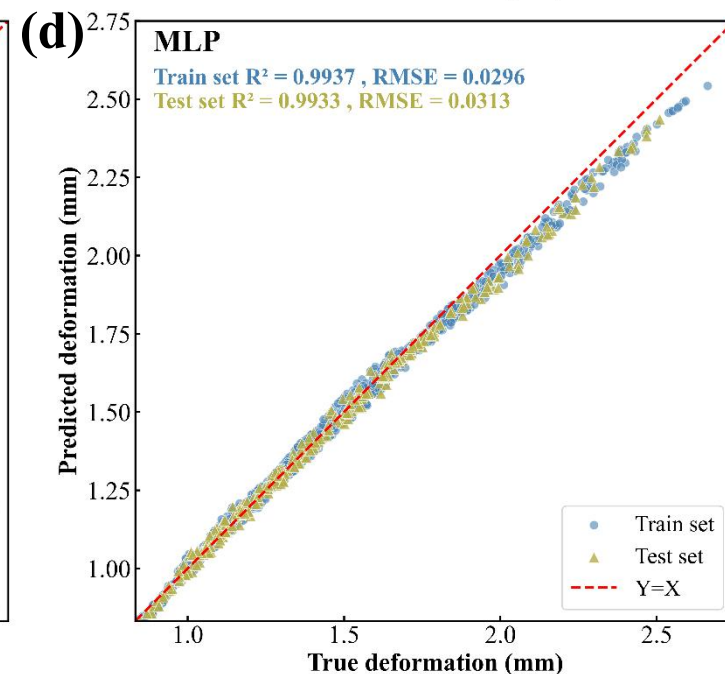
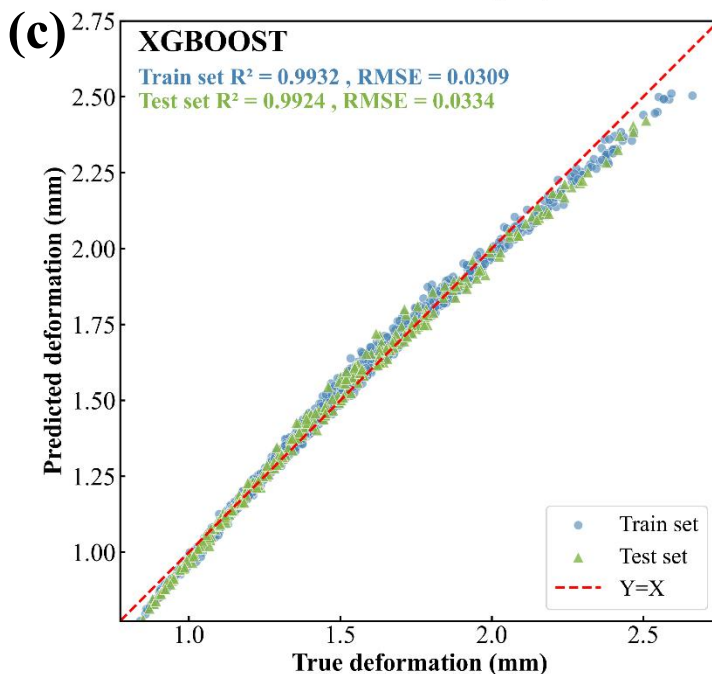
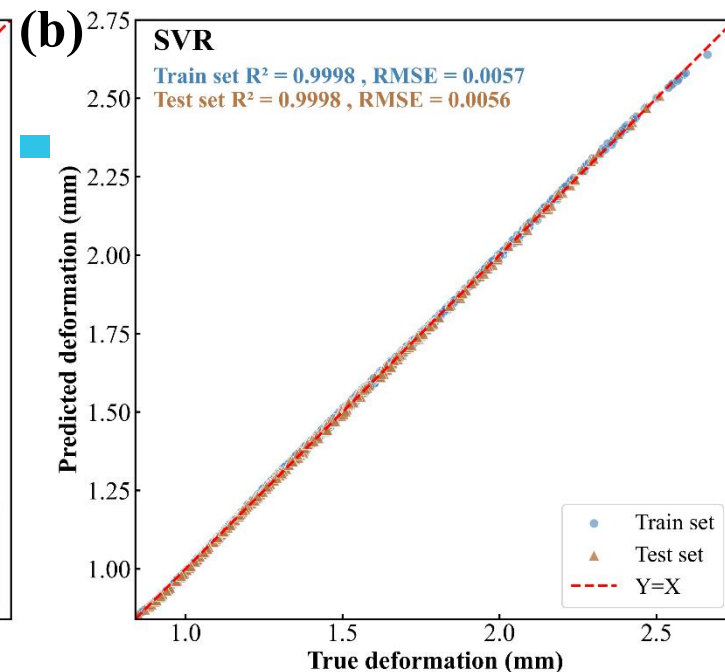
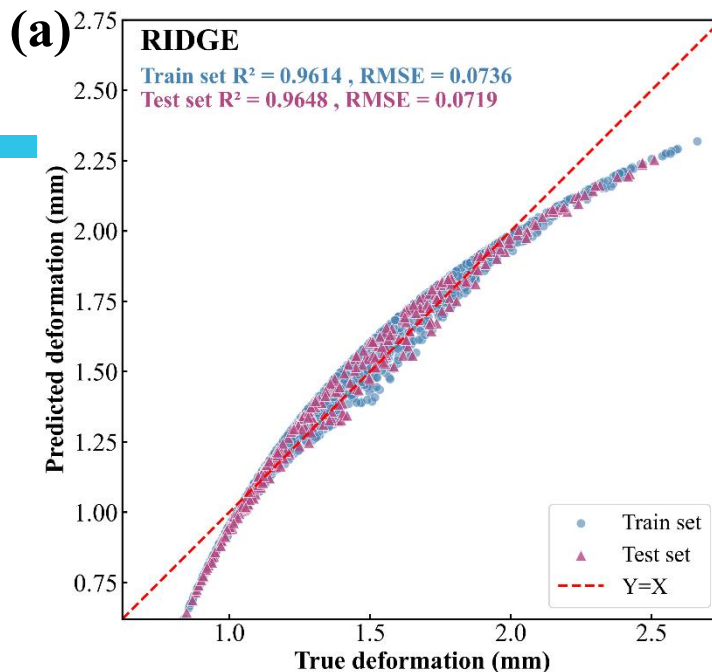
$Y=X$  为理想预测基准线  
预测值与真实值完全一致

### 结论:

1. 所有非线性模型 (SVR、MLP、XGBoost) 的预测精度均优于线性基准模型Ridge回归, 直接验证了结构变形与设计参数 (法兰厚度、支撑宽度) 之间存在显著的非线性映射关系, 必须采用非线性模型实现高精度预测。

2. SVR模型性能最优, 全变形区间拟合精度、泛化能力均为4个模型最佳; MLP、XGBoost处于第二梯队, 精度远优于线性基准, 但与SVR存在差距。

3. 所有模型的训练集与测试集指标无显著差异, 无明显过拟合现象, 1500条小样本数据可充分支撑各模型的训练, 模型泛化能力满足工程应用要求。



# 横向对比

## 最大应力值预测

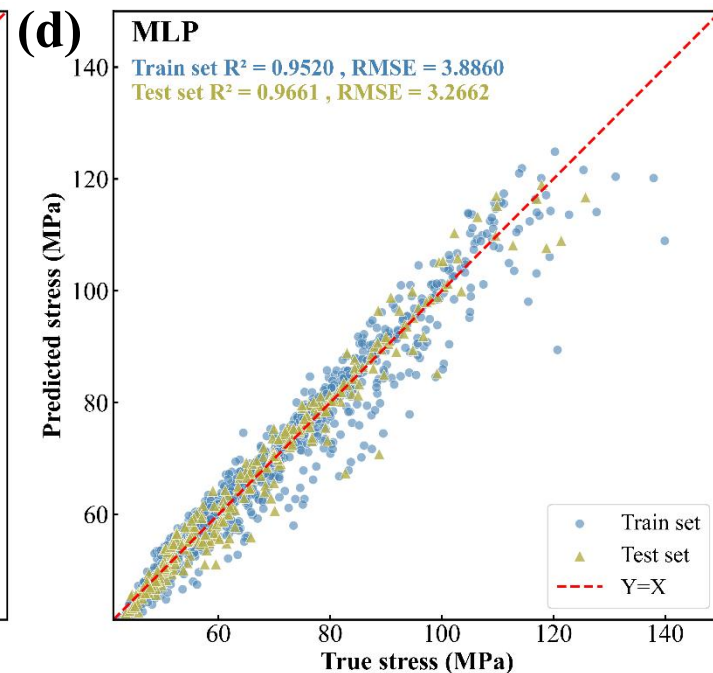
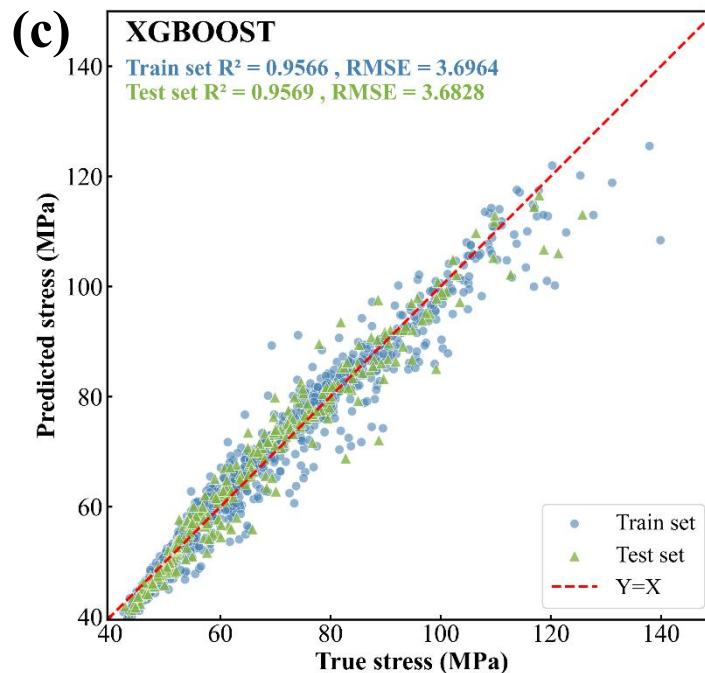
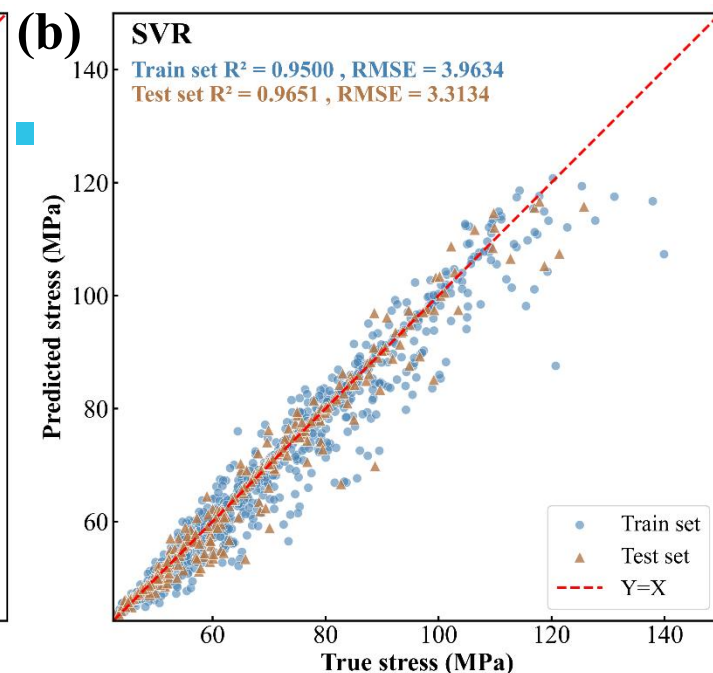
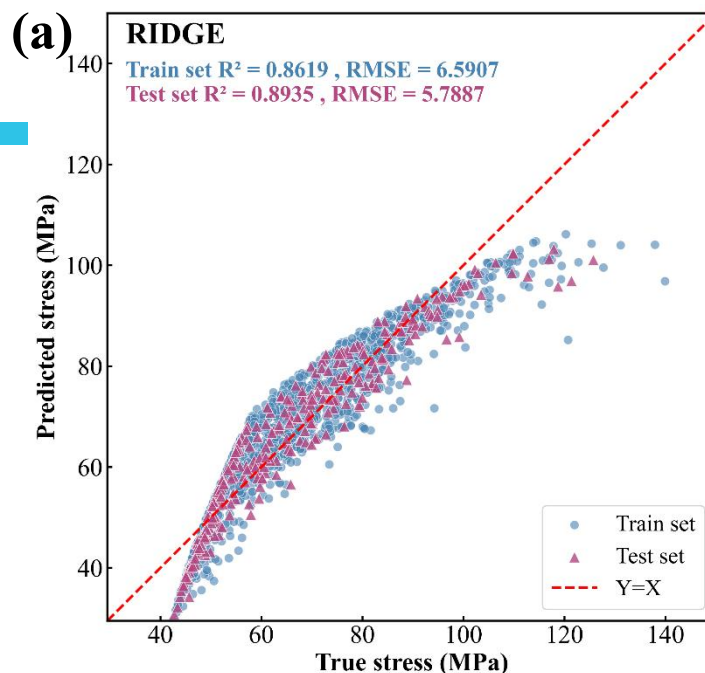
### 结论:

1. 应力预测的非线性特征远强于变形预测，线性基准模型Ridge回归的精度出现大幅下跌，与非线性模型差距显著，直接验证了结构应力与设计参数（法兰厚度、支撑宽度）之间存在强非线性映射关系，必须采用非线性模型预测。

2. 与变形预测中SVR领先不同，应力预测中SVR、MLP、XGBoost三个非线性模型处于同一性能梯队，测试集R2差距均在0.01以内，RMSE差距不足0.4MPa，性能差异极小。

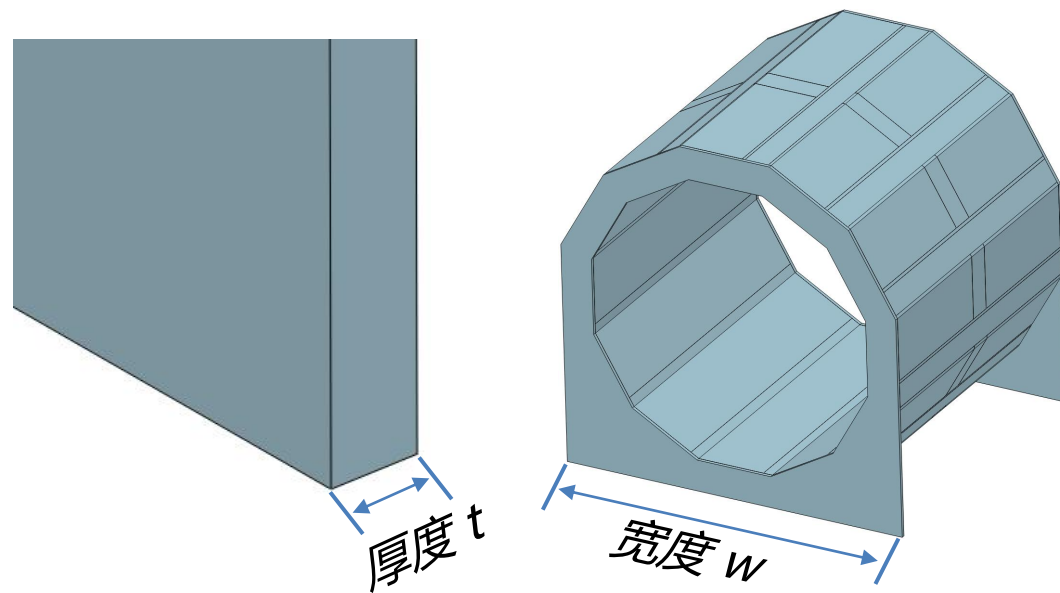
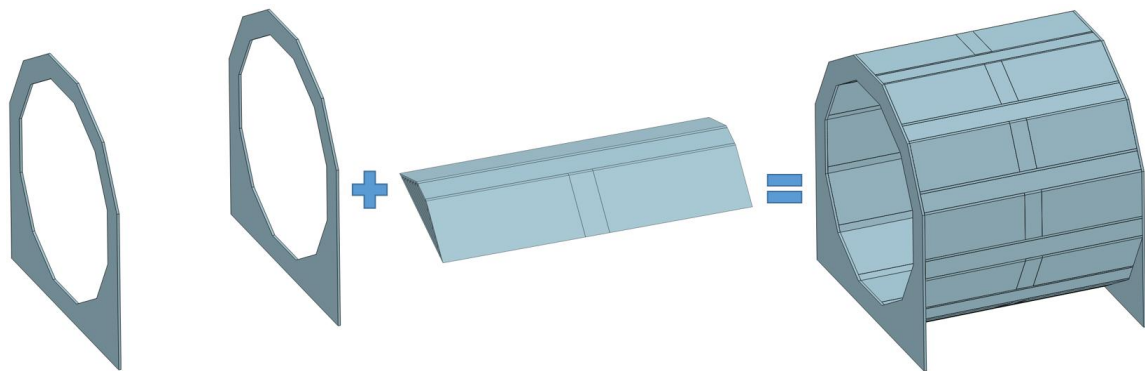
3. 所有非线性模型的测试集指标均优于或持平训练集，无任何过拟合现象，证明数据集的样本覆盖范围、数据质量可充分支撑模型训练，所选模型适配小样本仿真代理模型的场景。

4. 所有模型均呈现“低应力区拟合精度高、高应力区离散性大”的规律，符合结构力学特性——高应力区多伴随应力集中，非线性、突变性更强，拟合难度显著提升。



# 代理模型驱动的结构优化

## 端部法兰优化



### 设计要求:

1. 桶箍最大变形量 $\leq 2\text{mm}$
2. 桶箍最大应力值 $\leq 80\text{MPa}$
3. 在满足上述条件下，端部法兰质量越小越好

端部法兰质量与厚度 $t$ 和底部宽度 $w$ 相关

用训练好的 SVR 模型代替有限元仿真；再用寻优算法，在所有可能的 $t$ 和 $w$ 里，快速找到满足所有要求、质量最小的最优解。

**SVR模型:** svr\_deformation.pkl和svr\_stress.pkl

**寻优算法:** scipy.optimize

加载模型

设定优化  
目标约束

优化算法  
寻优

# 总结

- 4个模型，覆盖了从线性到非线性、从传统机器学习到深度学习的主流回归算法，既建立了基准线，也完成了横向对比。
- 所有模型的对比，都做了严格的控制变量：统一8:2划分训练/测试集、统一评价指标 ( $R^2$ /MAE/RMSE)、每个模型都进行了超参数寻优，保证对比公平。
- 综合考虑预测精度和稳定性的要求，SVR模型综合效果最优，变形预测 $R^2$ 达0.9998，应力预测 $R^2$ 达0.9651，满足工程设计的精度要求；同时预测速度是传统有限元分析的数倍，实现了加速有限元分析、提升设计迭代效率的核心目标。