1. 使用模型总结

1.1 单一模型

XGBoost

随机森林

CatBoost

LightGBM

AutoGluon

K最近邻回归（KNN）

神经网络（MLP）

支持向量机（SVC）

ExtraTrees

线性回归

1.2 复合模型

方法一：XGBoost + 随机森林

方法二：XGBoost + LightGBM + CatBoost

方法三：AutoGluon + XGBoost + 神经网络

方法四：ExtraTrees + XGBoost + LightGBM + CatBoost

方法五：随机森林回归 + 概率校准（Isotonic Regression）

方法六：AutoGluon + Optuna优化

2. 模型简短介绍

XGBoost：基于梯度提升的集成学习算法，通过构建多棵决策树优化目标函数，广泛应用于分类和回归任务，具有高效处理大规模数据和防止过拟合的能力。

随机森林：一种集成学习方法，通过构建多个决策树并综合其预测结果来提高模型的稳定性和准确性，适用于分类和回归任务，对数据中的噪声具有较强的鲁棒性。

CatBoost：基于梯度提升的机器学习框架，能够自动处理分类特征，无需手动编码，适用于分类和回归任务，对缺失值和过拟合有较好的处理能力。

LightGBM：基于梯度提升框架的高效机器学习模型，通过优化算法和树结构实现快速训练和高精度预测，适合处理大规模数据集。

AutoGluon：自动化机器学习工具，能够自动选择最优的机器学习算法和超参数，适用于分类和回归任务，通过集成多种算法提供高效且准确的预测结果。

K最近邻回归（KNN）：基于最近邻思想的机器学习模型，通过计算目标点与最近邻点的距离，对目标点的值进行加权平均，从而实现回归预测。

神经网络（MLP）：模拟人脑神经元结构的计算模型，通过多层非线性变换学习数据中的复杂模式，适用于分类和回归任务，具有强大的拟合能力。

支持向量机（SVC）：基于最大间隔原理的分类模型，通过寻找最优分割超平面实现分类，适用于线性和非线性分类任务，对高维数据有较好的处理能力。

ExtraTrees：基于随机森林的扩展模型，通过构建多个决策树并引入随机性提高模型的泛化能力，适用于分类和回归任务。

线性回归：一种简单的预测模型，通过线性组合特征来预测目标变量，适用于特征与目标之间存在线性关系的场景。