**1. 图**

1. **线性回归**用于价值预测

• **图示**：残差图

• **用途**：检查残差是否均匀分布，以验证模型拟合质量。

2. **逻辑回归**用于BS分类（Over/Under）

• **图示**：ROC曲线

• **用途**：展示模型在不同阈值下的分类表现。

3. \*\*K近邻（KNN）\*\*用于价值预测

• **图示**：误差率与邻居数量的关系图

• **用途**：识别最佳邻居数量，以尽量减少预测误差。

4. **决策树回归**用于价值预测

• **图示**：决策树可视化

• **用途**：显示基于预测变量的决策路径，提供对模型决策的洞察。

5. **随机森林**用于BS分类

• **图示**：特征重要性图

• **用途**：揭示哪些预测变量对分类结果影响最大。

6. \*\*支持向量机（SVM）\*\*用于BS分类

• **图示**：决策边界图

• **用途**：展示SVM所实现的类别分隔及验证决策边界。

7. **神经网络**用于价值预测

• **图示**：损失曲线（训练 vs 验证）

• **用途**：监控收敛情况，防止过拟合或欠拟合。

8. **梯度提升**用于BS分类

• **图示**：混淆矩阵

• **用途**：评估分类准确率和误分类率。

9. \*\*主成分分析（PCA）\*\*用于降维与特征分析

• **图示**：Scree图

• **用途**：评估所需的主成分数量，以捕获主要方差。

10. **XGBoost**用于价值预测

• **图示**：Shapley值总结图

• **用途**：解释模型预测并识别影响因素。

**2. PPT**

**幻灯片1：介绍 -** **问题的发现**

• **内容**：介绍准确预测期权价格和将其分类为“Over”或“Under”的挑战。提到传统模型如Black-Scholes的鲁棒性，但缺乏灵活性，从而促使我们探索数据驱动的解决方案。

**幻灯片2：****解决方案的探索 - 多种模型**

• **内容**：解释由于有多种预测变量（如资产价值、执行价、利率等），测试了不同的统计和机器学习技术，观察哪种方法的表现最佳。

• **图示**：模型选择流程图。

**幻灯片3-12：****模型细节与实验（每个模型1张幻灯片）**

每张幻灯片包含：

1. **模型介绍**：简要说明该模型及选择的原因。

2. **实验图示**：插入建议的特定图表。

3. **解释**：讨论每张图的洞察：

• **示例**：“线性回归的残差图显示误差大部分是随机分布的，说明模型能很好地拟合线性模式。”

**幻灯片13：****模型比较**

• **内容**：对比模型表现（如R平方值、分类器的AUC值等）。使用表格或柱状图展示。

**幻灯片14：****选择的最终方案**

• **内容**：基于在价值预测和BS分类中的表现，展示所选择的最终模型。

**幻灯片15：总结 - 故事的收尾**

• **内容**：总结从发现传统模型的局限性，到探索多种方法，最终选出最佳方案的整个过程。