

# 《基于数据表和图组的水果、蔬菜、肉类分类分析及优化建议》

以下是对水果类、蔬菜类、肉类的图组和数据表的分类分析，结合指标数据和可视化结果进行说明：

## 一、水果类（6 种产品）

### 数据表分析（MAE/MSE）

产品	MAE	MSE
富士苹果	0.464	0.3604
巨峰葡萄	0.1758	0.0844
菠萝	<b>0.0197</b>	<b>0.0042</b>
西瓜	0.2913	0.2762
香蕉	0.1671	0.0815
鸭梨	<b>0.0284</b>	<b>0.0008</b>

**表现最优：**菠萝、鸭梨的 MAE 和 MSE 均极低，说明预测值与真实值几乎重合，模型对这两类水果的价格趋势拟合极佳。

**表现较差：**富士苹果的 MSE 最高（0.3604），可能因价格波动较大或数据噪声导致预测误差；西瓜的 MAE 和 MSE 也相对较高，需检查数据预处理或模型参数。

### 图组分析

误差分布：

菠萝、鸭梨的误差集中在 0 附近，分布狭窄（峰值高），说明预测误差小且稳定。

富士苹果、西瓜的误差分布较宽，存在正负偏差，尤其是富士苹果可能有极端误差（尾部较长）。

训练损失曲线：

所有产品的损失曲线均呈下降趋势并趋于平稳，表明模型训练有效，其中鸭梨的损失收敛速度最快，波动最小。

预测对比：

菠萝、鸭梨的预测曲线与真实值高度重合，而富士苹果、西瓜在部分时间点偏离较大，可能受季节性价格波动影响。

二、蔬菜类（26 种产品）

数据表分析（MAE/MSE）

产品	MAE	MSE
洋白菜	0.0645	0.0042
土豆	0.0795	0.0124
大白菜	0.0778	0.0111
豆角	0.5243	0.8141
平菇	0.4003	0.3470
生姜	0.4086	0.3663

**表现最优：**洋白菜、土豆、大白菜的误差极小，说明模型对基础蔬菜价格的预测精度高，可能因数据规律性强（如稳定的供需关系）。

**表现较差：**豆角的 MSE 最高（0.8141），MAE 也显著高于其他蔬菜，可能因豆角的价格受季节、产地影响较大，数据波动复杂；平菇、生姜的误差也较高，需关注数据特征是否完整（如是否包含天气、运输成本等变量）。

图组分析

误差分布：

洋白菜、土豆的误差分布集中在 0 附近，呈正态分布，预测稳定性强。

豆角、平菇的误差分布较分散，存在多个峰值或长尾，反映价格波动的不确定性。

训练损失曲线：

大部分蔬菜的损失曲线在后期趋于平稳，但豆角、生姜的曲线可能波动较大，说明模型对高波动产品的拟合难度更高，需增加训练数据或调整模型复杂度（如使用 LSTM + 注意力机制）。

预测对比：

洋白菜等基础蔬菜的预测曲线紧密贴合真实值，而豆角的预测值在部分时段明显偏离，可能因突发因素（如产量骤减）未被模型捕获。

三、肉类（9 种产品）

数据表分析（MAE/MSE）

产品	MAE	MSE
活草鱼	0.2922	0.2248
活鲤鱼	0.3023	0.1863
白鲢活鱼	0.2974	0.1599
牛肉	1.3671	2.6749
猪肉(白条猪)	0.3148	0.3808
羊肉	0.5572	0.6880

**表现最优：**活鱼（草鱼、鲤鱼、白鲢）的 MAE 和 MSE 较低，可能因活体水产品价格受市场供需影响较稳定，数据规律性强。

**表现较差：**牛肉的 MAE 和 MSE 远高于其他肉类，可能因牛肉价格受养殖周期、市场政策（如进口限制）影响更大，模型难以捕捉长期趋势；羊肉的误差也较高，需检查是否存在数据缺失或异常值。

## 图组分析

### 误差分布：

活鱼的误差分布集中在 0 附近，而牛肉的误差分布范围极广（正负偏差大），存在多个峰值，反映价格波动的复杂性。

### 训练损失曲线：

活鱼的损失曲线快速收敛且平稳，而牛肉的损失曲线在后期仍有波动，说明模型对牛肉价格的拟合未完全收敛，可能需要增加训练时长或优化模型结构。

### 预测对比：

活鱼的预测值与真实值基本吻合，而牛肉的预测曲线在多数时段明显偏离，尤其是价格大幅上涨或下跌时，模型未能有效捕捉拐点。

## 四、总结与建议

### 模型效果差异原因：

**数据特征：**价格波动小、规律性强的产品（如洋白菜、鸭梨）预测精度高；受季节、政策影响大的产品（如牛肉、豆角）误差较高。

**模型适配：**对高波动产品，可尝试引入更多外部数据（如天气、政策文件）或使用更复杂的模型（如 Transformer）提升泛化能力。

### 图组布局优化：

水果类（6 种）采用 2×3 布局，肉类（9 种）采用 3×3 正方形布局，蔬菜类（26 种）采用 5×6 布局，确保图片布满且视觉整齐，便于对比分析。

### 后续工作重点：

针对高误差产品（如牛肉、豆角、富士苹果），检查数据质量（填补缺失值、处理异常值），调整模型超参数（如增大 SEQ\_LEN 捕捉长期趋势），或增加数据增强策略。

通过结合数据表的量化指标和图组的可视化分析，可清晰定位模型优势与不足，为后续优化提供明确方向。