### 分析流程 数据源： TOPSIS\_副本(3).xlsx 算法配置： 算法： XGBoost回归 变量： 变量X:{set\_no，game\_no，server，p1\_ace，p1\_net\_pt，p1\_break\_pt，p1\_break\_pt\_won，p1\_break\_pt\_missed，is\_tie\_breakers，runs\_won\_past，runs\_of\_fails，rate\_score\_past，rate\_of\_point}；变量Y:{mark\_wave}收起 分析结果： XGBoost回归基于MSE、RMSE、MAE、MAPE、R²指标对模型进行评价，请看详细结论。

### 分析步骤 1. 通过训练集数据来建立XGBoost回归模型。 2. 通过建立的XGBoost来计算特征重要性。 3. 将建立的XGBoost回归模型应用到训练、测试数据，得到模型评估结果。 4. 由于XGBoost具有随机性，每次运算的结果不一样，若保存本次训练模型，后续可以直接上传数据代入到本次训练模型进行计算预测。 5. 注：XGBoost无法像传统模型一样得到确定的方程，通常通过测试数据预测精度来对模型进行评价。

### 详细结论

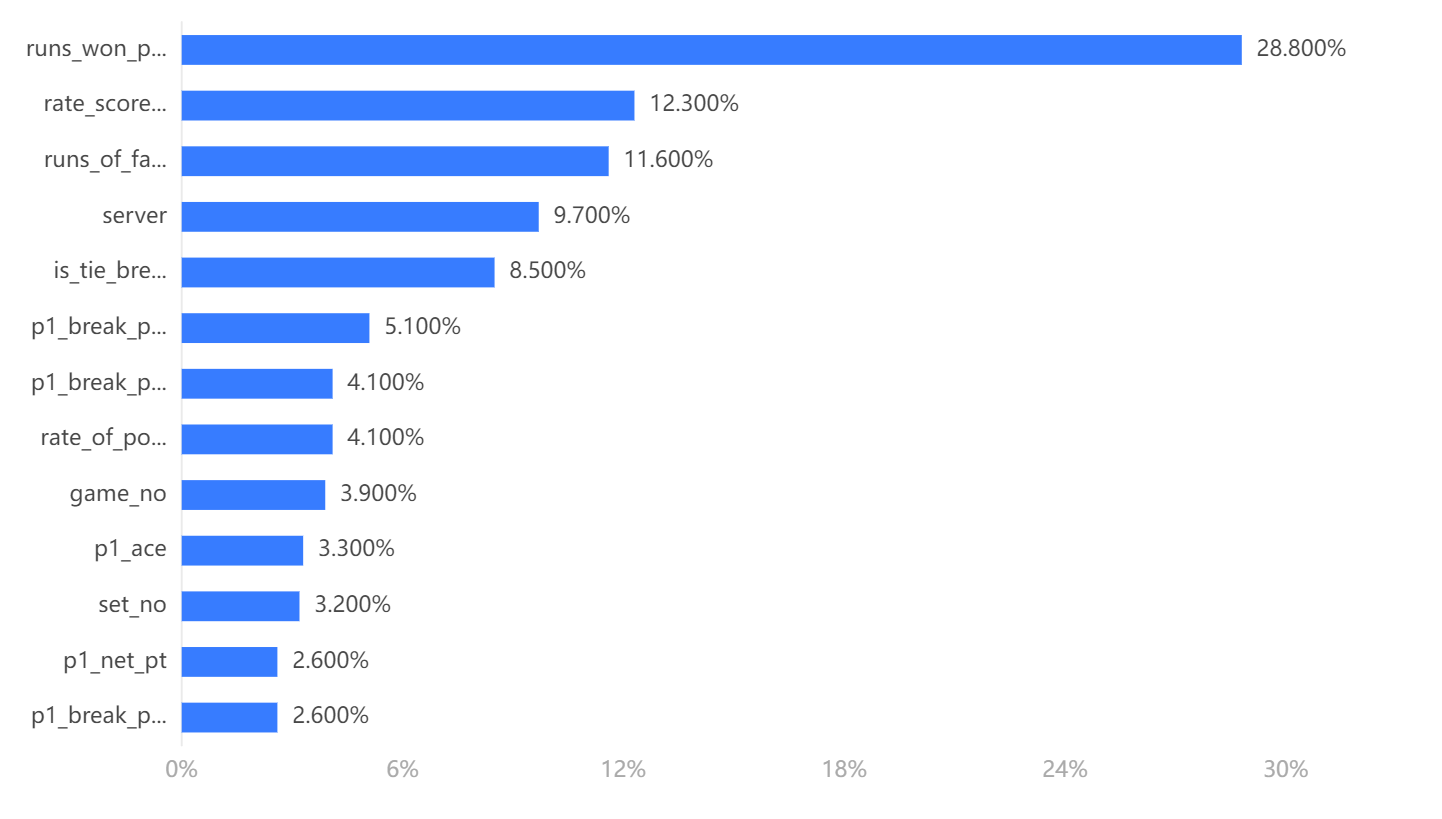
**输出结果1：模型参数**

|  |  |
| --- | --- |
| 参数名 | 参数值 |
| 训练用时 | 23.14s |
| 数据切分 | 0.8 |
| 数据洗牌 | 是 |
| 交叉验证 | 否 |
| 基学习器 | gbtree |
| 基学习器数量 | 1000 |
| 学习率 | 0.01 |
| L1正则项 | 10 |
| L2正则项 | 5 |
| 样本征采样率 | 0.6 |
| 树特征采样率 | 0.7 |
| 节点特征采样率 | 0.7 |
| 叶子节点中样本的最小权重 | 0.2 |
| 树的最大深度 | 100 |

**图表说明：**

上表展示了模型各项参数配置以及模型训练时长。

**输出结果2：特征重要性**



**图表说明：**

上柱形图或表格展示了各特征（自变量）的重要性比例。

**输出结果3：模型评估结果**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | MSE | RMSE | MAE | MAPE | R² |
| 训练集 | 10.336 | 3.215 | 2.478 | 567.061 | 0.742 |
| 测试集 | 23.314 | 4.828 | 3.673 | 571.315 | 0.411 |

**图表说明：**

上表中展示了交叉验证集、训练集和测试集的预测评价指标，通过量化指标来衡量XGBoost的预测效果。其中，通过交叉验证集的评价指标可以不断调整超参数，以得到可靠稳定的模型。  
● MSE（均方误差）： 预测值与实际值之差平方的期望值。取值越小，模型准确度越高。  
● RMSE（均方根误差）：为MSE的平方根，取值越小，模型准确度越高。  
● MAE（平均绝对误差）： 绝对误差的平均值，能反映预测值误差的实际情况。取值越小，模型准确度越高。  
● MAPE（平均绝对百分比误差）： 是 MAE 的变形，它是一个百分比值。取值越小，模型准确度越高。  
● R²： 将预测值跟只使用均值的情况下相比，结果越靠近 1 模型准确度越高。

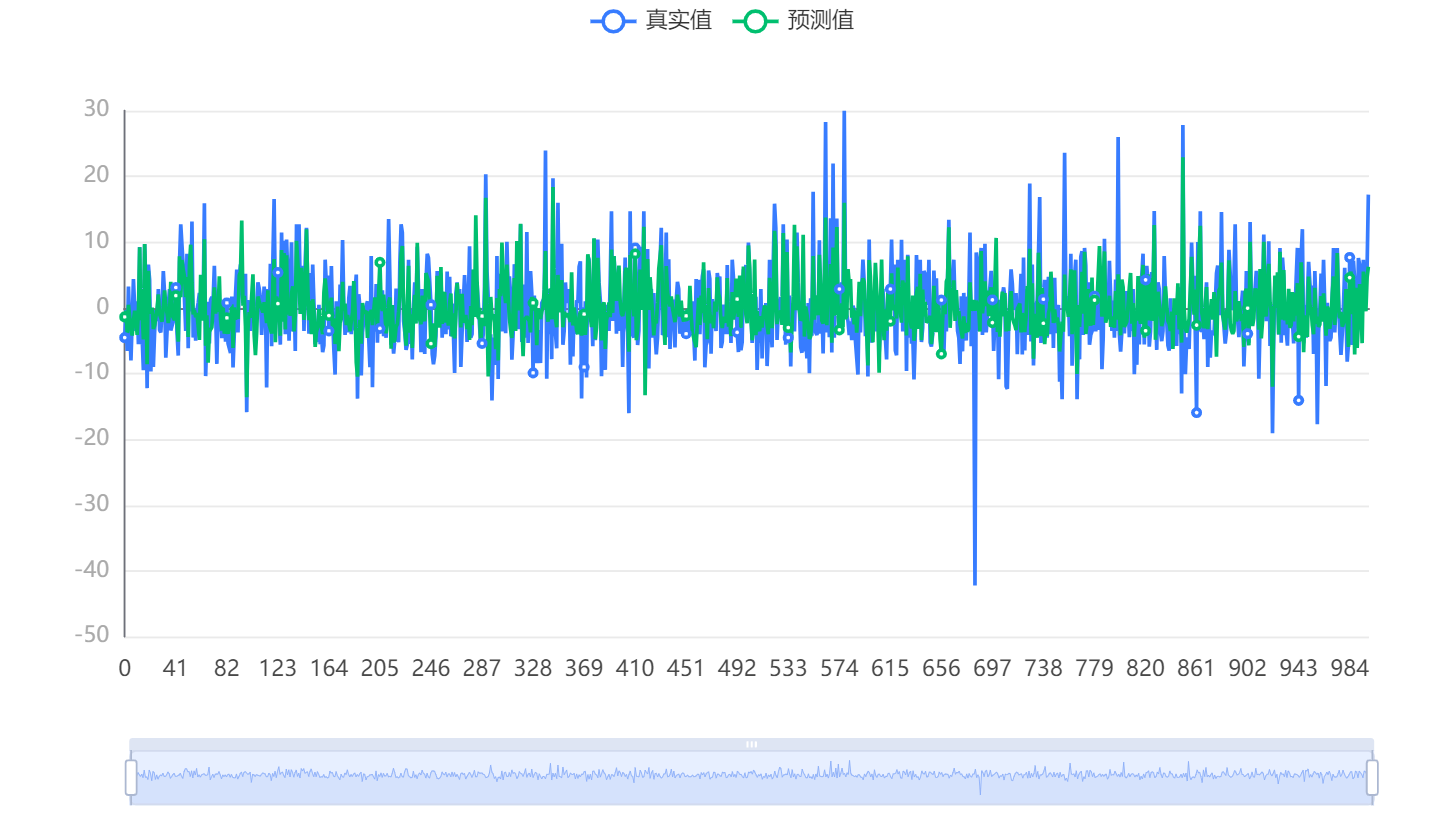
**输出结果4：预测结果**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 预测测试集结果Y | mark\_wave | set\_no | game\_no | server | p1\_ace | p1\_net\_pt | p1\_break\_pt | p1\_break\_pt\_won | p1\_break\_pt\_missed | is\_tie\_breakers | runs\_won\_past | runs\_of\_fails | rate\_score\_past | rate\_of\_point |
| -1.3826687335968018 | -4.520821433 | 4 | 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0.501858736 |
| -0.46506986021995544 | -0.323468774 | 3 | 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0.5 | 0.475903614 |
| -3.8458380699157715 | -6.56664229 | 4 | 3 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0.5 | 0.573033708 |
| 0.4699242413043976 | 3.232788066 | 2 | 13 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0.5 | 0.46835443 |
| 0.7171541452407837 | -2.669732949 | 1 | 7 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0.459459459 |
| -3.3097782135009766 | -7.989230463 | 4 | 5 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0.507389163 |
| -4.198944568634033 | -1.984430819 | 4 | 5 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0.527638191 |
| -1.9014095067977905 | 4.395940418 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0.5 | 0.5 |
| -0.44979554414749146 | 0.093335109 | 2 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0.422535211 |
| -1.7202898263931274 | -3.507220773 | 2 | 7 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 1 | 0.49382716 |
| -4.096675395965576 | 0.990463834 | 4 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0.476190476 |
| -0.6208010315895081 | -5.524460625 | 1 | 8 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0.476190476 |
| 9.231284141540527 | 9.242084766 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0.578947368 | 0.555555556 |
| -1.8367998600006104 | -5.524460625 | 5 | 9 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0.478395062 |
| -0.8340600728988647 | -4.520821433 | 1 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0.416666667 |

**图表说明：**

上表格为预览结果，只显示部分数据，全部数据请点击下载按钮导出。  
上表展示了XGBoost对测试数据的预测情况。

**输出结果5：测试数据预测图**



**图表说明：**

上图中展示了XGBoost对测试数据的预测情况。预测图最多只展现测试集前1000个样本预测信息，若需要绘制完整的预测图，可在预测结果中导出数据重新绘图。

**输出结果6：模型预测与应用**

**图表说明：**

如果预测的数据超过15条，系统会自动展示前15条，其余数据请点击预览数据右上方下载按钮查看全部。

### 参考文献 [1] Scientific Platform Serving for Statistics Professional 2021. SPSSPRO. (Version 1.0.11)[Online Application Software]. Retrieved from https://www.spsspro.com. [2] Chen T , Guestrin C . XGBoost: A Scalable Tree Boosting System[J]. ACM, 2016.