Q2:

建立模型后，我们需要验证模型的最优性。我们采用拟合优度、MSE、MAE三个指标来作为模型的评估指标，它们的计算公式如下。

越接近1，MSE、MAE越小，则模型拟合效果越好。

我们分别采取了三个特征组合，分别用LSTM模型和随机森林模型进行训练，得出、MSE、MAE，并进行比较分析。特征组合如下：

①

|  |  |  |  |
| --- | --- | --- | --- |
| 特征名 | 定义 | 类型 |  |
| Value', | 价格 | float64 | 当天对应价格 |
| 'roi\_x', | 投资回报率 | int | 今天价格是否大于昨天，若大于为1，反正为-1，相等则为0 |
| 'roi\_Value\_x', | 价格变化量 | float64 | 今天价格与昨天价格的差值 |
| roi\_Value\_percent\_x', | 价格变化量 | float64 | 今天价格与昨天价格的差值除以昨天的价格 |
| month\_mean\_Value\_x', | 每月平均价格 | float64 | 按每个月分类，求每个月的均值 |
| month\_roi\_x' | 月份投资回报率 | int | 当天价格与本月均值比较，若当天价格大于本月均值为1，反之为-1，相等则为0 |
| USD (PM)' | 价格 | float64 | 当天对应价格 |
| 'roi', | 投资回报率 | int | 今天价格是否大于昨天，若大于为1，反正为-1，相等则为0 |
| roi\_Value\_y', | 价格变化量 | float64 | 今天价格与昨天价格的差值 |
| roi\_Value\_percent\_y', | 价格变化量 | float64 | 今天价格与昨天价格的差值除以昨天的价格 |
| month\_mean\_Value\_y', | 每月平均价格 | float64 | 按每个月分类，求每个月的均值 |
| month\_roi\_y' | 月份投资回报率 | int | 当天价格与本月均值比较，若当天价格大于本月均值为1，反之为-1，相等则为0 |
| buy\_Value | 去掉交易成本的比特币购买价格 | float64 | 购买比特币的真实价格 |
| sale\_Value | 去掉交易成本的比特币出售价格 | float64 | 出售比特币的真实价格 |
| buy\_USD (PM) | 去掉交易成本的黄金购买价格 | float64 | 购买黄金的真实价格 |
| sale\_USD (PM) | 去掉交易成本的黄金出售价格 | float64 | 出售黄金的真实价格 |
| new\_roi\_bitcoin | 去掉交易成本的比特币回报率 | float64 | 比特币的真实回报率 |
| new\_roi\_gold | 去掉交易成本的黄金回报率 | float64 | 黄金的真实回报率 |

②

|  |  |  |  |
| --- | --- | --- | --- |
| 特征名 | 定义 | 类型 |  |
| Value | 价格 | float64 | 当天对应价格 |
| roi\_x | 投资回报率 | int | 今天价格是否大于昨天，若大于为1，反正为-1，相等则为0 |
| roi\_Value\_x | 价格变化量 | float64 | 今天价格与昨天价格的差值 |
| roi\_Value\_percent\_x | 价格变化量 | float64 | 今天价格与昨天价格的差值除以昨天的价格 |
| month\_mean\_Value\_x | 每月平均价格 | float64 | 按每个月分类，求每个月的均值 |
| month\_roi\_x | 月份投资回报率 | int | 当天价格与本月均值比较，若当天价格大于本月均值为1，反之为-1，相等则为0 |
| USD (PM) | 价格 | float64 | 当天对应价格 |
| roi | 投资回报率 | int | 今天价格是否大于昨天，若大于为1，反正为-1，相等则为0 |
| roi\_Value\_y | 价格变化量 | float64 | 今天价格与昨天价格的差值 |
| roi\_Value\_percent\_y | 价格变化量 | float64 | 今天价格与昨天价格的差值除以昨天的价格 |
| month\_mean\_Value\_y | 每月平均价格 | float64 | 按每个月分类，求每个月的均值 |
| month\_roi\_y | 月份投资回报率 | int | 当天价格与本月均值比较，若当天价格大于本月均值为1，反之为-1，相等则为0 |

③

|  |  |  |  |
| --- | --- | --- | --- |
| 特征名 | 定义 | 类型 |  |
| Value | 价格 | float64 | 当天对应价格 |
| USD (PM) | 价格 | float64 | 当天对应价格 |

用LSTM模型得出的结果为：

用随机森林模型得出的结果为：

After building the model, we need to verify the optimality of the model. We use three indicators of goodness of fit 、MSE, and MAE as the evaluation indicators of the model, and their calculation formulas are as follows.

We took three feature combinations respectively, and trained them with LSTM model and random forest model, respectively, and obtained 、MSE, and MAE, and compared and analyzed them. The combination of features is as follows:

①

②

③

The result obtained with the LSTM model is:

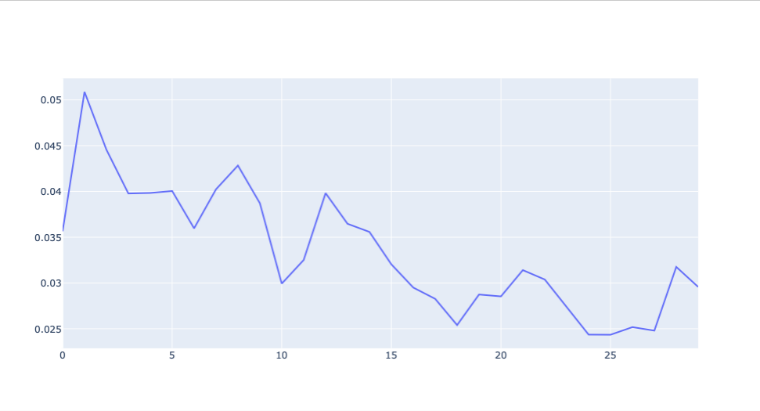
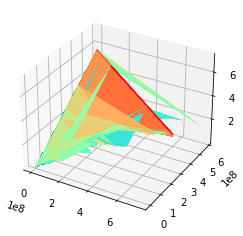
Table : The average value of various evaluation index values using different feature combinations—LSTM

|  |  |  |  |
| --- | --- | --- | --- |
| Feature combination | R2 | MAE | MSE |
| ① | \* | 0.01416 | 0.0026 |
| ② | \* | 1.2789 | 2.8675 |
| ③ | \* | 1.6784 | 3.9857 |

The result obtained with the RF model is:

Table : The average value of various evaluation index values using different feature combinations—RF

|  |  |  |  |
| --- | --- | --- | --- |
| Feature combination | R2 | MAE | MSE |
| ① | 0.9545 | 0.3567 | 0.0146 |
| ② | 0.8598 | 1.3245 | 2.9985 |
| ③ | 0.8056 | 1.7895 | 4.0934 |

The MSEs of the LSTM and random forest models are shown above, respectively. It can be seen from the data in the table that the effect of the LSTM model is better than that of the random forest model, and Feature combination① is the best among all Feature combinations. From this, it can be proved that the model we obtained is the optimal model. Therefore, the LSTM model we trained with Feature combination① can help us calculate the optimal policy.

LSTM和随机森林模型的MSE分别如上图所示。由表中数据可知，LSTM模型的效果比随机森林模型的效果更优，并且Feature combination①是所有Feature combination中效果最好的。由此可以证明，我们得出的模型是最优模型。因此，我们用Feature combination①训练出的LSTM模型可以帮助我们计算出最优策略。