LSTM（Long Short－time Memory Networks，LSTM）是一种特殊的循环神经网络（RNN），可以进行长期依赖信息的学习。RNN神经网络模型一直被广泛用于语言识别和文本分类等多个研究领域。与RNN模型相比，LSTM模型最明显的改进是采用了精心设计的称作“门”的结构，来削弱或者增强信息到达细胞状态的能力。LSTM模型增加了１个细胞状态Ｃ和３个阀门。３个阀门分别是遗忘门ｆ、输出门o和输入门i。在传统的RNN模型中，由于权重在反向传播的过程中进行了连续相乘的操作，当需要预测的信息与相关的信息的时间跨度较大时，会出现vanishing gradient and exploding gradient的问题。然而，在 LSTM模型误差反向传播校正权重时，有些误差可以直接通过输入门传递给下一层神经元，有些误差则可以通过遗忘门丢弃，这样就解决了vanishing gradient and exploding gradient的难题，即有效地降低了历史数据中相关信息的冗余度。

LSTM的工作原理

单向的LSTM一般按照时间序列从前向后运行，学习数据内部特征，通过分类器得到最终结果；而双向LSTM是按照时间顺序正向和反向运行的，通过分类器的内部特征由两个单向LSTM 共同决定最后输出的结果。因此，算法可以更好地学习数据内部特征，提高了精确度。因为题目要求我们设计比特币-黄金购买策略模型是一种时间序列问题，而且每天的购买策略受比特币和黄金历史价格的影响，所以我们选用LSTM模型来设计购买策略。而且我们在做出交易决策时，需要同时考虑前一天的数据和预测出的后一天的数据, 双向LSTM由前向和反向层共同决定最后输出结果，更适用于解决我们的问题，因此我们选用了双向的LSTM模型。

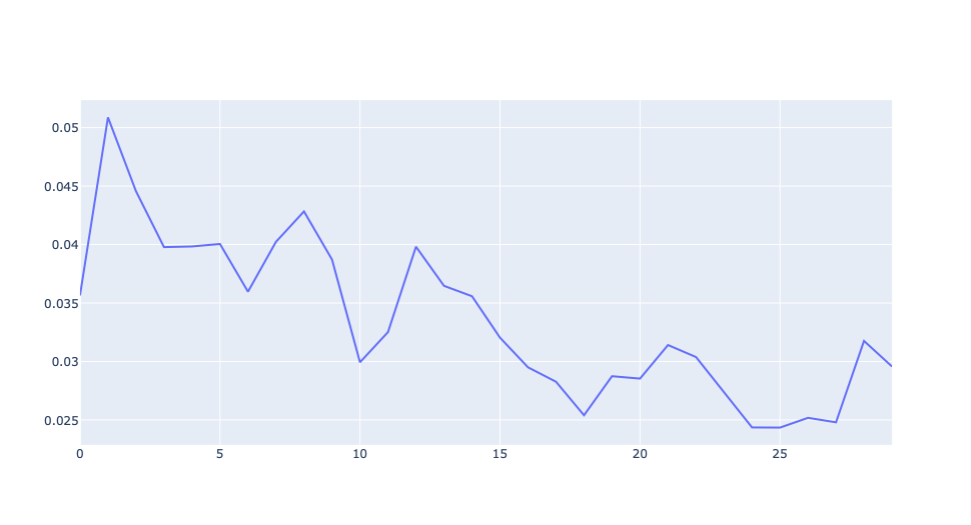
LSTM的训练过程:

使用Pytorch中的LSTM模型来进行训练，首先把data表中的数据分别按time\_step为18、30、60、90进行数据的筛选，分别挑选出以三周、一个月、两个月、一个季度的三种情况，接着把数据按照8:2的比例进行training set和prediction set的分割，接着开始构建Iterator，设置batch\_size分别为128、256、512进行一个epoch的sample训练，并设置num\_workers为4进行多线程读取数据，接着开始构建LSTM模型，设置input\_size与time\_step相等、hidden\_size个数为6、num\_layers为2、bidirectional为True，构建一个含有两层hidden layer、第一层hidden layer为6的hidden neurons的双向LSTM模型，接着开始进行epoch为30的训练，误差分别使用了MAE、MSE，optimizer分别使用了Adam、SGD，Learning Rate 设为0.01，训练完成以后不断调优直至误差值达到最小的时刻，并使用训练好的LSTM来进行预测。参数设置在Tab1中。

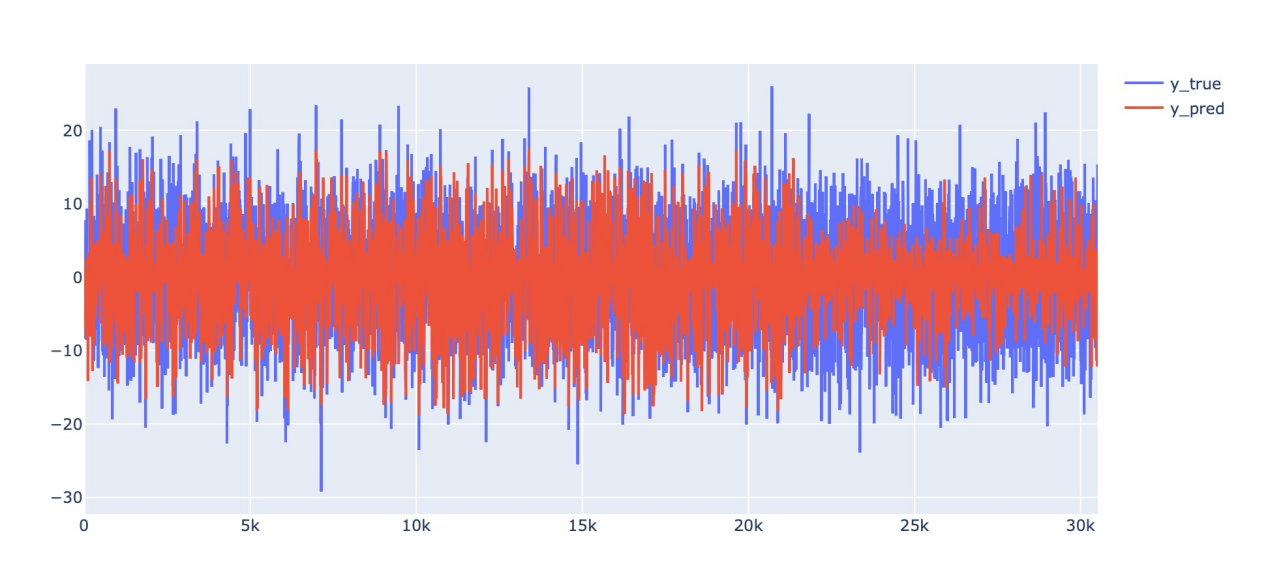
Tab1：The setting of long-short time memory neural network parameters

|  |  |
| --- | --- |
| 设置项 | 参数值 |
| hiddern\_size | 6 |
| num\_layers | 2 |
| num\_workers | 4 |
| learning rate | 0.01 |
| bidirectional | True |

通过模型训练及不断地调优后得到的误差结果在Pic1中显示，LSTM预测值和真实值的情况在Pic2中。



Pic1 Mean Squared Error



Pic2 预测值和真实值

LSTM (long short-time memory networks, LSTM) is a special type of RNN（Recurrent Neural Networks), which can learn long-term dependent information. RNN model has been widely used in many research fields, such as language recognition and text classification. The most obvious improvement of the LSTM model over the RNN model is the carefully designed structure called “gates” to weaken or enhance the ability of information to reach cellular states. The LSTM model adds one cell state C and three valves, which are forgetting gate F, output gate O and input gate I respectively. In the traditional RNN model, because the weights operate in continuous multiplication during the backpropagation, the problem of gradient explosion or disappearance occurs when the time span of the predicted information with the relevant information is large. However, when the LSTM model error backpropagation correction weights, some errors can be passed directly through the input gate to the next layer of neurons, and some errors can be discarded through the forgetting gate, thus solving the problem of gradient explosion and disappearance, that is, effectively reducing the redundancy of relevant information in the historical data.

**How the LSTM model works.**

The one-way LSTM model generally runs back to the time series, learning the internal characteristics of the data and obtains the final results through the classifier, while the two-directional LSTM is run in the sequential forward and reverse order, and the final output results are jointly determined by two one-way LSTM models. Therefore, the algorithm can better learn the internal characteristics of the data and improve the accuracy. The design of bitcoin-gold purchase strategy model required by the topic is a time-series problem, and the daily purchase strategy is affected by the historical price of Bitcoin and gold. Therefore, we chose the LSTM model to design the purchase strategy. Moreover, when we make transaction decisions, we need to consider both the data of the previous day and the predicted data of the next day. The two-way LSTM determines the final output results jointly by the forward and reverse layers, which is more suitable to solve our problem. Therefore, we chose the two-way LSTM model.