

# Gold Price Prediction Based on Long Short-Term Memory Network

Ming Yan<sup>1\*</sup>, Dongxi Li<sup>2</sup>

<sup>1</sup>School of Mathematics, Taiyuan University of Technology, Jinzhong Shanxi

<sup>2</sup>School of Big Data, Taiyuan University of Technology, Jinzhong Shanxi

Email: \*mingyan0357@163.com

Received: May 28<sup>th</sup>, 2020; accepted: Jun. 12<sup>th</sup>, 2020; published: Jun. 19<sup>th</sup>, 2020

---

## Abstract

Accurate prediction of gold price is helpful for investors to understand the gold market and provide scientific and effective reference for them to make correct investment decisions. Therefore, it is particularly important to improve the prediction accuracy of gold price. This paper presents a gold price forecasting method based on Long Short-term memory network (LSTM) model. This method combines the gold price data, uses LSTM model to generate training model, and finally realizes the prediction of gold price. The results show that this method is feasible and effective, and has higher prediction accuracy than BP neural network and SVR intelligent prediction.

## Keywords

Gold Price, Long Short-Term Memory Network (LSTM), Short-Term Prediction, Long-Term Prediction

---

# 基于长短期记忆网络的黄金价格预测

闫 铭<sup>1\*</sup>, 李东喜<sup>2</sup>

<sup>1</sup>太原理工大学数学学院, 山西 晋中

<sup>2</sup>太原理工大学大数据学院, 山西 晋中

Email: \*mingyan0357@163.com

收稿日期: 2020年5月28日; 录用日期: 2020年6月12日; 发布日期: 2020年6月19日

---

## 摘 要

精确的预测黄金价格, 有助于投资者了解黄金市场的行情, 并对他们做出正确的投资决策提供了科学有效的参考, 因此, 提高黄金价格的预测精度显得尤为重要。本文提出了一种基于长短期记忆网络(LSTM)

---

\*通讯作者。

模型的黄金价格预测方法。该方法结合黄金价格数据, 利用长短期记忆网络生成训练模型, 最终实现对黄金价格的预测。结果表明, 本文的方法可行有效, 较BP神经网络和SVR智能预测拥有更高的预测精度。

## 关键词

黄金价格, 长短期记忆网络(LSTM), 短期预测, 长期预测

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

在如今的国际金融体系中, 由于黄金的非货币化, 各国逐渐放松了对黄金的管制, 于是全球黄金市场得到了极大的发展, 已经成为与股票市场、期货市场、债券市场、外汇市场等并列的金融投资市场。作为一种金融产品, 黄金价格的动态变化, 是经济行为主体在金融市场中投资决策过程的具体反映, 黄金价格的预测更是众多投资者和学者讨论的热点话题[1] [2] [3]。但在实际中, 黄金价格受到供求关系、美元汇率、石油价格、通货膨胀率、国际形势、股票、债券等因素的影响, 形成一种复杂的非线性过程, 很难得到精确的预测。国内外学者在黄金价格预测方面做出了众多成果, 使用了不同的方法: 文献[4]使用 BP 神经网络对黄金价格进行预测和分析, 得出了精度较高的预测结果。然而 BP 神经网络的滚动优化算法的收敛速度较慢, 且多目标优化能力不强, 难以摆脱决策过程中的随机性和模糊性。文献[5]引入 SVR 智能预测方法对黄金价格进行预测研究, 并通过引入网络搜索法对 SVR 模型的最优参数进行寻找, 从而构建了最优的 SVR 智能预测模型, 结果表明该模型具有优越的学习性能与泛化推广性能, 能够较准确地预测黄金价格。但是这些传统的方法只是对黄金价格进行小数据量的分析, 导致预测精度低、速度慢且效率低下。随着大数据的深入研究, 将长短期记忆网络(LSTM)用于价格指数的预测, 能满足用户对实时数据处理的需求。

黄金价格不但具有非线性, 而且具有动态特性, 即系统的输出不仅与当前时刻的输入有关, 还与过去的输入有关。作为一种特殊的循环神经网络, 长短期记忆网络(LSTM)不但具有“记忆性”, 适合处理带“序列”性质的数据, 如时间序列数据、每天的股价走势情况等, 同时采用特殊的结构取代原始网络中的隐藏层单元, 采用“累加”的形式, 避免了循环神经网络中的梯度消失和梯度爆炸[6] [7]。目前, LSTM 已经被应用于声音识别[8], 视频分类[9]、语音理解[10]等诸多前沿领域中。

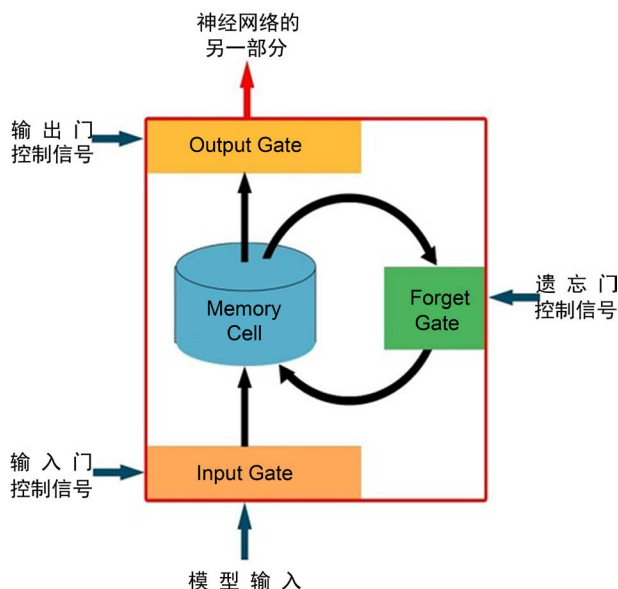
## 2. 长短期记忆网络

长短期记忆网络是一种时间递归神经网络, 适合于处理和预测时间序列中间隔和延迟相对较长的问题。目前, 长短期记忆网络已经在科技领域有了诸多应用: 基于长短期记忆网络可以进行语言翻译、控制机器人、图像分析、语音识别、手写识别。2015年, 谷歌通过基于 CTC 训练的长短期记忆网络程序大幅提升了安卓手机和其他设备中语音识别的能力; 苹果的 iPhone 在 QuickType 和 Siri 中使用了长短期记忆网络; 微软不仅将长短期记忆网络用于语音识别, 还将这一技术用于虚拟对话形象生成和编写程序代码等[11] [12] [13]。

### 2.1. 长短期记忆网络模型结构及网络原理

长短期记忆网络(Long Short-term Memory, LSTM)是一种特殊的循环神经网络。之所以说它是特殊的 RNN, 是因为它本身不是一个独立存在的网络结构, 而是由 LSTM 特殊的结构单元取代循环神经网络的

隐藏层单元部分。在 LSTM 中, 不同时间点之间的网络存在连接, 而不是单个时间点处的网络存在前馈或反馈。

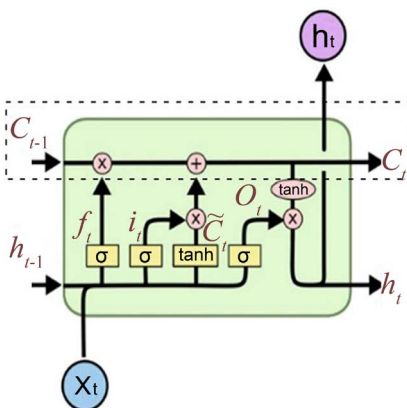


**Figure 1.** Unit structure diagram of LSTM  
**图 1.** 长短期记忆网络单元结构图

从图 1 可以看出: 一个 LSTM 单元包含了一个细胞状态(Memory Cell)、一个模型输入和三个门信号。其中, 细胞状态是 LSTM 的核心思想, 是整个模型的记忆空间。LSTM 模型共有四个输入, 分别是: 模型输入(Input)、输入门(Input gate)、输出门(Output gate)和遗忘门(Forget gate), 而普通神经网络只有一个输入, 因此 LSTM 的参数量是它们的 4 倍。其中 Input gate, Output gate, Forget gate 是三个门信号, 用来选择性处理数据, 通过 Sigmoid 函数和点乘操作实现, Sigmoid 的取值介于 0 和 1 之间, 点乘则决定了传递的信息量。因此当门信号的取值为 0 时代表完全关闭, 1 代表完全打开。

## 2.2. 长短期记忆网络模型的训练算法

长短期记忆网络模型的训练算法包括信号的正向传播和误差的反向传递两部分。



**Figure 2.** Structure chart of LSTM  
**图 2.** 长短期记忆网络结构图

### 2.2.1. 信号的正向传播

根据图 2, 我们可以清楚地看到输入节点和门信号的功能:

1) 输入节点: 接受当前时刻的样本输入和上一时刻隐层单元的输出。

2) 遗忘门: 用来决定前一时刻 Memory Cell 中的信息是否会被保留下来。当遗忘门的输出为 1 时, 前一时刻 Memory Cell 中的信息会被保留下来; 当遗忘门的输出为 0 时, 前一时刻 Memory Cell 中的信息就会被遗忘。简单地说, 遗忘门的作用是刷新 Memory Cell 的状态, 控制 Memory Cell 的上一状态对当前状态的影响。公式为:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

其中:  $h_{t-1}$  是  $(t-1)$  时刻的输出,  $x_t$  是  $t$  时刻本层的输入,  $W_f$  是各个变量的权重,  $b_f$  是截距项,  $\sigma$  是遗忘门的 Sigmoid 函数,  $f_t$  是 0~1 之间的数。

3) 输入门: 更新细胞状态中的存储信息。具体分为 3 步计算: 第 1 步决定输入节点的激活值有多少会被保留下来, 可以看到输入门的 Sigmoid 函数会和输入节点的激活值相乘, 组成 LSTM 中 Memory Cell 的一部分。我们知道, 在序列输入中, 并不是每个信息都同等的重要, 当输入的信息有用时, 输入门的输出为 1, 输入节点的值可以全部输入进去; 当输入的信息完全无用时, 输入门的输出为 0, 此时输入节点的信息对 internal state 没有任何影响。第 2 步根据 tanh 函数创建新向量, 添加到细胞状态中。第 3 步用旧的细胞状态乘以遗忘门  $f_t$ , 然后再加上新的输入信息, 更新当前的细胞状态。公式为:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (3)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (4)$$

其中: tanh 为正切激励函数,  $\tilde{C}_t$  表示从  $t$  时刻输入的信息中提取出来要记住的信息,  $C_{t-1}$  表示  $(t-1)$  时刻的细胞状态值,  $C_t$  表示更新后的细胞状态值,  $i_t$  是 0~1 之间的数。

4) 输出门: 先使用 Sigmoid 函数决定输出多少信息量, 然后使用 tanh 函数决定当前 Memory Cell 中有多少信息会被输出, 可以看到输出门会和 Memory Cell 的值相乘, 当输出门的输出为 1 时, Memory Cell 的值会被完全输出; 当输出门的输出为 0 时, Memory Cell 的值就不会被输出。公式为:

$$O_t = \sigma(W_o \cdot [H_{t-1}, x_t] + b_o) \quad (5)$$

$$h_t = O_t * \tanh(C_t) \quad (6)$$

其中:  $O_t$  是 0~1 之间的数。

### 2.2.2. 误差的反向传递

对模型进行训练的关键在于权重矩阵  $W$  和截距项  $b$  的不断调整, 目的是使得网络的输出误差在选定的损失函数下越小越好, 而参数的更新主要通过反向传播的链式求导得到。因此反向传播的目的就是求预测误差关于所有参数的梯度, 在 LSTM 模型中, 共有五个部分。其类似于 BPTT 算法[14]。总梯度为:

$$\frac{\partial E}{\partial W} = \sum_{t=0}^T \frac{\partial E^t}{\partial W} \quad (6)$$

## 3. 实证研究和结果分析

### 3.1. 实验数据

为了方便比较和分析, 从英为财经官网下载每日黄金期货(GCZ8)数据。把这些数据先从影响因素上

分为单因子(最高价)和多因子(开盘价、收盘价、最高价、最低价)两部分, 再从时间上分为短期数据(2008/10/22~2018/10/22)和长期数据(1990/06/22~2018/10/22), 从而建立 4 个数据集, 表 1 列出了这些数据集的特征。

**Table 1.** Summary of experimental data

**表 1.** 实验数据汇总

序号	数据集	样本个数	时间范围	变量数目
1	Short_1	2586	2008/10/22~2018/10/22	1 个(最高价)
2	Short_1	2586	2008/10/22~2018/10/22	4 个(开盘价、收盘价、最高价、最低价)
3	Long_1	7189	1990/06/22~2018/10/22	1 个(最高价)
2	Long_2	7189	1990/06/22~2018/10/22	4 个(开盘价、收盘价、最高价、最低价)

### 3.2. 长短期记忆网络模型单因子预测

#### 3.2.1. 实验步骤

##### 1) 数据预处理

读入黄金期货数据, 对其进行标准化处理, 消除量纲带来的影响。一般数据标准化的步骤为: 先求出各变量的算术平均值(数学期望)  $x_i$  和标准差  $s_i$ ; 再进行标准化处理:  $z_{ij} = \frac{x_{ij} - x_i}{s_i}$ , 其中  $z_{ij}$  为标准化后的变量值,  $x_{ij}$  为实际变量值。

接下来定义模型训练时调节的主要参数, 主要包括: 批大小(batch\_size)、隐层节点数(rnn\_unit)、模型的时间步长(time\_step)等。

① 批大小是指进行梯度下降时每一批包含的样本数, 在训练的时候每批的样本会被计算一次梯度下降, 使目标函数被优化一步。它可以决定梯度下降的方向, 比如一个极端情况为 batch\_size 等于 1, 每次参数更新的方向都是各自样本梯度下降的方向, 目标函数的优化就难以收敛。

② 隐层节点数(rnn\_unit)反映隐藏层内部映射和输出维度。隐层神经元个数的选择对于模型预测效果有着非常重要的影响, 如果隐层神经元的个数太少, 那么模型就无法很好地提取数据特征, 其预测效果也很差; 如果神经元个数太多, 就会造成模型在训练集表现很好, 但在测试集上反而表现很差, 即过拟合。

③ 模型的时间步长反映的是模型最多可以记忆多长的时间, 如果 LSTM 的时间步长为 10 天, 即 LSTM 最多可以记忆 10 天的时间, 那么第 11 天的黄金期货价格与前 10 天的价格有关。

本文定义时间步为 20, 每批次训练 60 个样例, 隐层节点数为 10, 输入为 1 维数据, 输出为 1 维数据。

##### 2) 长短期记忆网络的定义与训练

在定义 LSTM 模型之前, 先对神经网络的变量进行定义, 如定义输入权重, 输入偏差等。然后通过之前定义的输入维度, 输出维度定义 LSTM 神经网络, 再对其进行训练。训练的目的在于不断调整权重和偏差项, 使得最终的均方误差降到最低。最后对模型分别训练 1000 次、2000 次、3000 次、4000 次和 5000 次, 比较最终的结果。

##### 3) 黄金期货价格的预测

使用训练好的 LSTM 模型, 得到下一日黄金价格的预测值并与实际值比较。

### 3.2.2. 短期预测结果

使用数据集 Short\_1 生成训练模型, 得到单因子短期预测结果, 如表 2 所示。可以看出, 只使用最高价一个因子, 且有较少数据时, 模型的均方误差介于 0.001 和 0.003 之间, 随着训练次数的增加, 均方误差先减小后增大, 这是因为当训练次数过多时, 会出现过拟合的情况, 即训练出来的模型对训练集来说误差很小, 但并不适用于测试集。在单因子短期预测中, 下一日黄金价格的真实值为 1243.00, 其标准化结果为-0.268350911。本文的 LSTM 模型在训练次数为 2000 次时, 得到预测结果为-0.29719335, 均方误差为 0.0014775329, 说明用 LSTM 模型预测黄金价格可行有效。

Table 2. Single factor short-term forecast results

表 2. 单因子短期预测结果

训练次数	均方误差	预测值
1000	0.0016045270	-0.27769384
2000	0.0014775329	-0.29719335
3000	0.0015098210	-0.28685120
4000	0.0015384698	-0.30222230
5000	0.0028385934	-0.29912508

### 3.2.3. 长期预测结果

使用数据集 Long\_1 生成训练模型, 得到单因子长期预测结果, 如表 3 所示。可以看出, 使用最高价一个因子, 且有较多数据时, 模型的均方误差介于 0.0005 和 0.003 之间, 相比单因子短期预测, 同样是单因子, 长期预测比短期预测的均方误差有所降低。但随着训练次数的增加, 仍会出现过拟合的情况。在单因子长期预测中, 下一日黄金价格的真实值为 1243.00, 其标准化结果为 1.10。本文的 LSTM 模型在训练次数为 4000 次时, 得到预测结果为 1.0620697, 均方误差为 0.000524908, 说明 LSTM 模型在预测黄金价格时表现较好。

Table 3. Single factor long-term forecast results

表 3. 单因子长期预测结果

训练次数	均方误差	预测值
1000	0.002784248	1.0873673
2000	0.001507329	1.0839670
3000	0.001201251	1.0795054
4000	0.000524908	1.0620697
5000	0.002959938	1.0629648

## 3.3. 长短期记忆网络模型多因子预测

### 3.3.1. 实验步骤

#### 1) 数据预处理

读入黄金期货数据, 并将其分为两部分。对于短期预测, 将前 2300 组数据作为训练集, 剩下的 286 组数据作为测试集; 对于长期预测, 将前 6000 组数据作为训练集, 剩下的 1189 组数据作为测试集。再对训练集和测试集数据进行标准化处理, 消除量纲带来的影响。最后定义 LSTM 有两个隐藏层, 每个隐

藏层有 10 个神经元, 输入为收盘价、开盘价和最低价 3 维数据, 输出为最高价 1 维数据等。

2) 长短期记忆网络的定义与训练

在定义 LSTM 模型之前, 先对神经网络的变量进行定义, 如定义输入权重, 输入偏差等。再通过之前定义的输入维度, 输出维度定义 LSTM 神经网络, 最后对其进行训练。训练的目的是不断调整权重和偏差项, 使得最终的均方误差降到最低。最后对模型分别训练 1000 次、2000 次、3000 次、4000 次和 5000 次, 比较最终结果。

3) 黄金期货价格的预测

利用生成的训练模型, 对测试集上的黄金期货数据进行预测, 并计算偏差, 即:  $acc = \frac{y_{预} - y_{真}}{y_{真}}$ 。

3.3.2. 短期预测

使用数据集 Short\_2 生成训练模型, 得到多因子短期预测结果, 如表 4 所示。可以看出, 使用最高价, 最低价, 开盘价, 收盘价四个因子, 且数据较少时, 模型的均方误差介于 0.0001 和 0.0004 之间, 相较单因子短期预测, 同样数据较少, 多因子比单因子预测的均方误差更小。但随着训练次数的增加, 仍会出现过拟合的情况。当训练次数为 4000 次时, 模型的均方误差最小, 拟合效果最好。

Table 4. Multi factor short-term forecast results

表 4. 多因子短期预测结果

训练次数	均方误差	预测值
1000	0.000325662	0.001845699
2000	0.000260878	0.001912991
3000	0.000198966	0.001884724
4000	0.000149761	0.001960919
5000	0.000219345	0.001902263

使用得到的多因子短期预测模型, 设置训练次数为 4000 次, 对测试集的黄金价格进行预测, 如图 3 所示。蓝色为实际值, 红色为预测值, 可以看出预测值和实际值很接近, 说明模型的拟合效果比较好。

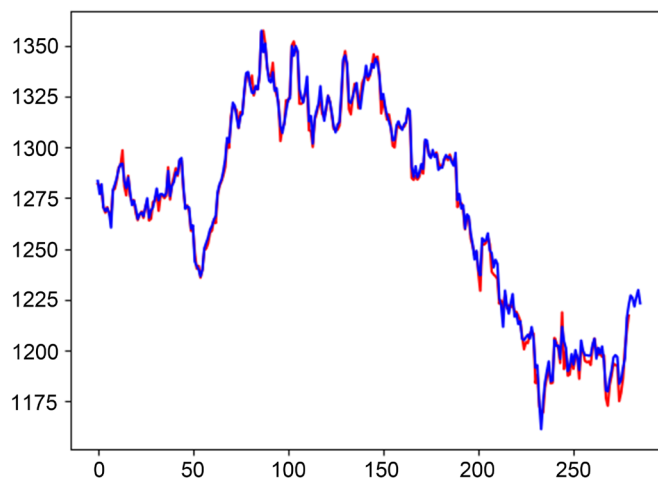


Figure 3. Multi factor short-term forecast chart

图 3. 多因子短期预测图

### 3.3.3. 长期预测

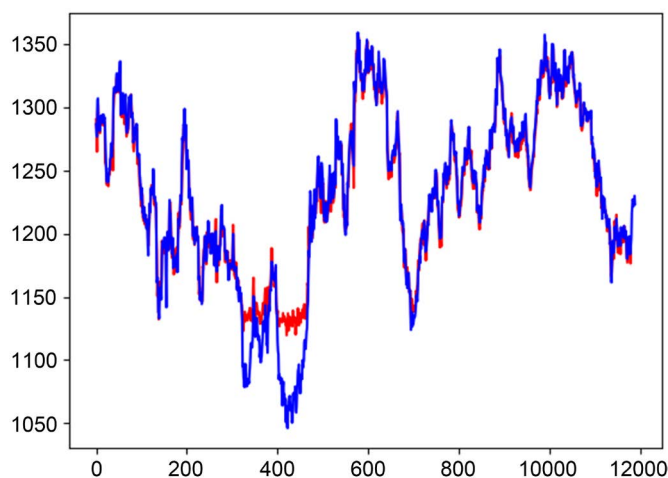
使用数据集 Long\_2 生成训练模型, 得到多因子长期预测结果, 如表 5 所示。可以看出, 使用最高价, 最低价, 开盘价和收盘价四个因子, 且数据较多时, 模型的均方误差介于  $0.00005$  和  $0.0005$  之间, 相较于多因子短期预测, 同样是多因子, 长期预测比短期预测的均方误差更小。当训练次数为 5000 次时, 模型的均方误差最小, 拟合效果较短期预测也更好。

**Table 5.** Multi factor long-term forecast results

**表 5.** 多因子长期预测结果

训练次数	均方误差	预测值
1000	$4.006 \times 10^{-4}$	0.007742414644
2000	$3.289 \times 10^{-4}$	0.008534511771
3000	$6.627 \times 10^{-5}$	0.007894591309
4000	$1.360 \times 10^{-4}$	0.008153175614
5000	$5.861 \times 10^{-5}$	0.007131973184

使用得到的多因子长期预测模型, 设置训练次数为 5000 次, 对测试集的黄金价格进行预测, 如图 4 所示。蓝色为实际值, 红色为预测值, 可以看出预测值的趋势和实际值吻合, 且在大部分数据点上, 预测值和实际值很接近, 说明用本文的 LSTM 模型预测黄金价格的效果比较好。



**Figure 4.** Multi factor long-term forecast chart

**图 4.** 多因子长期预测图

### 3.4. 预测结果分析

从以上单因子预测和多因子预测结果来看, 无论是有一个影响因素, 还是有多个影响因素, 长期预测相较于短期预测的均方误差更小, 预测更接近真实值, 这个结果指导我们在用 LSTM 模型做数据预测时, 应尽可能多的收集数据(表 6)。另一方面, 无论是短期预测还是长期预测, 多因子预测较单因子预测的均方误差更小(表 7), 这一结果说明在用 LSTM 模型做黄金价格预测时, 应更全面地考虑影响黄金价格的因素, 预测结果才能越精确。



**Table 6.** Statistics of prediction results of LSTM  
**表 6.** 长短期记忆网络的预测结果统计

影响因素	时间范围	训练次数	均方误差
单因子(最高价)	短期预测 (过去 10 年数据)	2000	$1.478 \times 10^{-3}$
	长期预测 (过去 28 年数据)	4000	$5.249 \times 10^{-4}$
多因子 (开盘价、收盘价 最高价、最低价)	短期预测 (过去 10 年数据)	4000	$1.498 \times 10^{-4}$
	长期预测 (过去 28 年数据)	5000	$5.861 \times 10^{-5}$

#### 4. 结束语

为了对黄金价格进行准确的预测, 本文提出了基于长短期记忆网络的预测方法。长短期记忆网络的主要思想体现在用特殊的结构单元取代原始网络的隐藏层部分, 用“累加”的形式, 避免了梯度消失和梯度爆炸, 同时具有记忆性, 适合处理间隔和延迟较长的时间序列数据。本文通过对原始数据集中抽样构造多个小数据集, 在每个小数据集上用长短期记忆网络进行预测, 结果表明多因子长期预测具有更小的均方误差, 而长短期记忆网络与文献中的方法相比具有很大的优势。BP 神经网络[4]预测的均方误差为 0.0053104, SVR 智能预测方法[5]的均方误差为 0.0002, 本文采用长短期记忆网络预测的均方误差为  $5.861 \times 10^{-5}$ , 所以长短期记忆网络在预测黄金价格上有更大的适用性。

**Table 7.** Comparison of mean square error of each model  
**表 7.** 各模型均方误差的比较

模型	均方误差
BP 神经网络[4]	$5.310 \times 10^{-3}$
SVR 智能预测方法[5]	$2.000 \times 10^{-4}$
长短期记忆网络(本文)	$5.861 \times 10^{-5}$

由于国际黄金市场同时受供求关系、美元汇率、石油价格、通货膨胀率、国际形势、股票、债券等因素的影响, 内在规律很复杂, 长短期记忆网络虽然可以在一定程度上帮助人们进行预测, 但训练次数较多, 耗费时间过长, 想要使预测达到更精确的程度, 需要进一步研究与挖掘。

#### 基金项目

山西省自然科学基金资助项目(11571009)。

#### 参考文献

- [1] 谢睿航, 申慧涛, 段绪晨. 基于 LSTM 网络的地铁线路客流量预测[J/OL]. 中国战略新兴: 1-2[2018-11-19].
- [2] 曹博, 高茂庭. 基于 LSTM 的短时交通流预测研究[J]. 现代计算机(专业版), 2018(25): 3-7.
- [3] 任君, 王建华, 王传美, 王建祥. 基于正则化 LSTM 模型的股票指数预测[J]. 计算机应用与软件, 2018, 35(4): 44-48+108.
- [4] 凌晨, 张骅月. 基于 BP 神经网络的黄金价格预测分析[J]. 天津科技, 2014, 41(1): 68-73.
- [5] 杨潇. 中国黄金期货价格的 SVR 智能预测研究[J]. 会计之友, 2017(17): 50-53.

- [6] Hochreiter, S., Bengio, Y., Frasconi, P., *et al.* (2001) Gradient Flow in Recurrent Nets: The Difficulty of Learning Long-Term Dependencies. In: *A Field Guide to Dynamical Recurrent Networks*, Wiley-IEEE Press, Munchen, 237-243.
- [7] Bengio, Y., Simard, P. and Frasconi, P. (1994) Learning Long-Term Dependencies with Gradient Descent Is Difficult. *IEEE Transactions on Neural Networks*, **5**, 157-166. <https://doi.org/10.1109/72.279181>
- [8] 陈超学. 基于循环神经网络的声学车型识别研究[D]: [硕士学位论文]. 西安: 西安电子科技大学, 2015.
- [9] 张磊. 基于深度学习的视频分析系统[D]: [硕士学位论文]. 南京: 南京邮电大学, 2017.
- [10] Song, E., Soong, F.K. and Kang, H.G. (2017) Effective Spectral and Excitation Modeling Techniques for LSTM-RNN-Based Speech Synthesis Systems. *IEEE/ACM Transactions on Audio, Speech and Language Processing*, **25**, 2152-2161. <https://doi.org/10.1109/TASLP.2017.2746264>
- [11] Cortez, B., Carrera, B., Kim, Y.-J. and Jung, J.-Y. (2018) An Architecture for Emergency Event Prediction Using LSTM Recurrent Neural Networks. *Expert Systems with Applications*, **97**, 315-324. <https://doi.org/10.1016/j.eswa.2017.12.037>
- [12] Li, Y.F. and Cao, H. (2018) Prediction for Tourism Flow Based on LSTM Neural Network. *Procedia Computer Science*, **129**, 277-283. <https://doi.org/10.1016/j.procs.2018.03.076>
- [13] Li, Z.B., Peng, F., Niu, B.S., Li, G.Y., Wu, J. and Miao, Z. (2018) Water Quality Prediction Model Combining Sparse Auto-Encoder and LSTM Network. *IFAC Papers Online*, **51**, 831-836. <https://doi.org/10.1016/j.ifacol.2018.08.091>
- [14] Sivakumar, S.C., Robertson, W. and Phillips, W.J. (1997) A Modified BPTT Algorithm for Trajectory Learning in Block-Diagonal Recurrent Neural Networks. *IEEE 1997 Canadian Conference on Electrical and Computer Engineering Innovation: Voyage of Discovery*, Saint Johns, 25-28 May 1997, 297-300. <https://doi.org/10.1109/CCECE.1997.614848>