

# 机器学习与私人专家系统优化下的国际黄金价格预测系统开发

**The development of International Gold Price forecasting optimized by  
Machine Learning and Private Expert System**

汤 吉  
ABS-TEMP-2015A

中国民航大学  
中欧航空工程师学院

1 April, 2016

- 1 | 课题实现流程
- 2 | 黄金市场可预测性分析
- 3 | 国际金价历史数据的获取与预处理
- 4 | 神经网络的选择与构建
- 5 | 总结



# 1. 课题实现流程简述

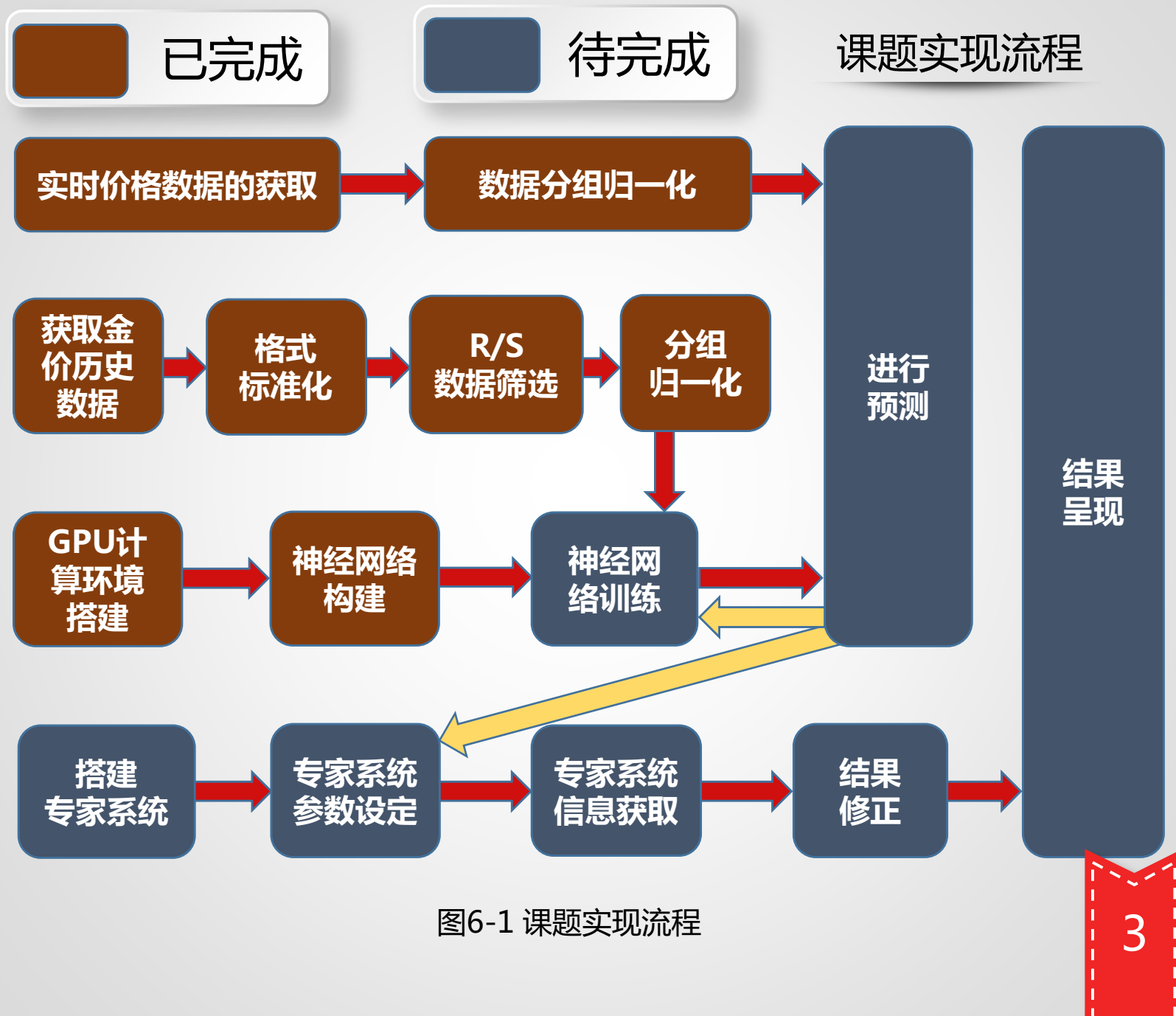


图6-1 课题实现流程



## 2. 黄金市场可预测性分析

# 市场的分类

Fama根据投资者可以获得的信息种类，将有效市场分成了三个层次[1]：

- **弱式市场(2000-2010年)**
- **半强市场(2010年之后)**  
依靠企业的财务报表等公开信息进行的基础分析法是无效的。
- **强式市场(未知时间)**



图2-1 国际金价历史走势

考虑到认知能力的差异，不管社会多么发达，教育水平有多高，对于相同的信息集总是有人先理解，有人后理解，有人不理解。

## R/S分析法

基于重标极差(R/S)分析方法基础上的Hurst指数(H) 作为判断时间序列数据遵从随机游走还是有偏的随机游走过程的指标。[2]

$$E\left[\frac{R(n)}{S(n)}\right] = Cn^H \quad \text{当 } n \rightarrow \infty$$

R(n)为前n个值的标准化量变化范围，S(n)为均方差，H即为Hurst指数

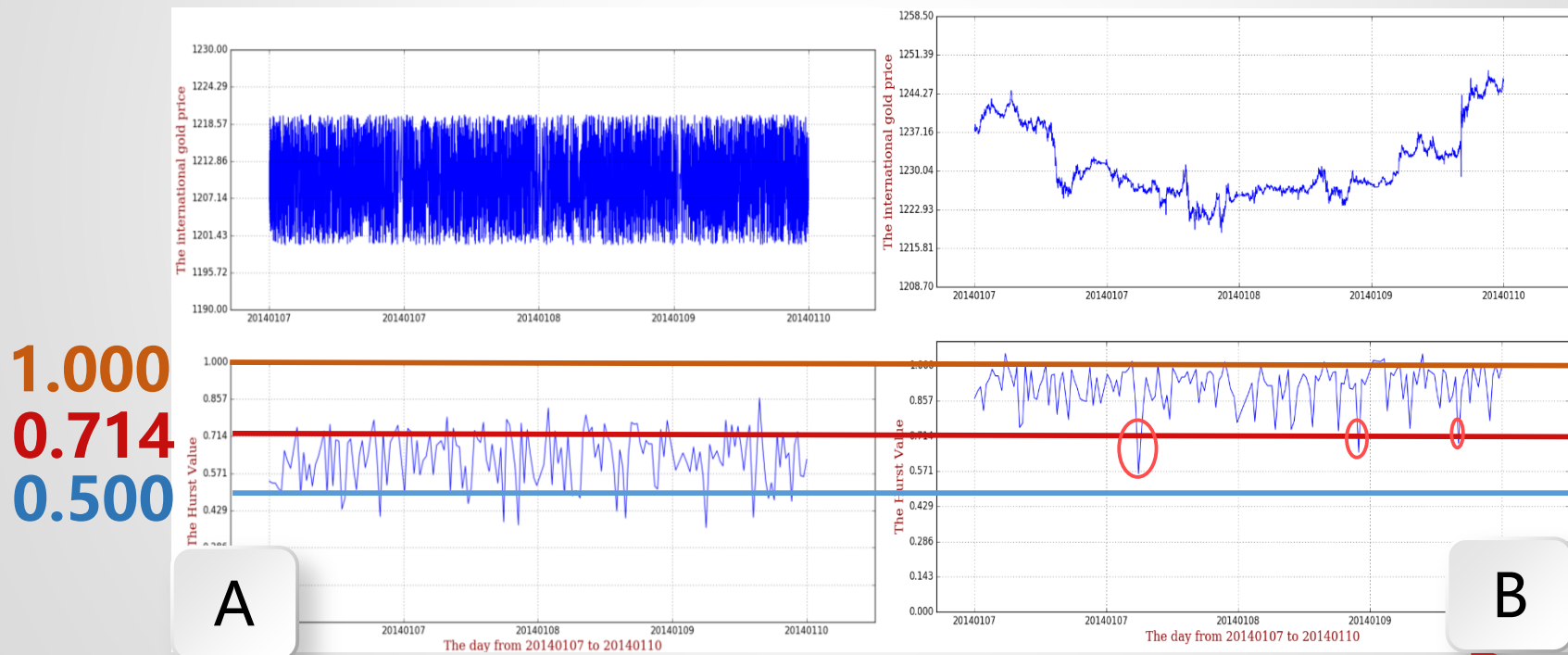


图2-2 随机游走与实际金价的(R/S)分析对比

### 3. 国际金价历史数据的 获取与预处理



## 数据的获取途径与格式

Forextester是美国的一家外汇交易网站，免费提供各种外汇交易的国际均价的历史数据(从2001年至今)。我们可以在<http://www.forextester.com/data/datasources>免费下载到近15年的国际黄金价格（XAU/USD）的历史数据。

下载到的数据以.scv的形式（纯文本格式）储存，我们可以用简单的程序实现对所有数据的读取，并将时间转换为更实用的UTC（国际标准时）格式。

```
20030101, 230100, 346. 5000, 346. 5000, 346. 5000, 346. 5000, 4
20030101, 231700, 346. 5000, 346. 5000, 346. 5000, 346. 5000, 4
20030101, 233400, 346. 5000, 346. 5000, 346. 5000, 346. 5000, 4
20030101, 235100, 346. 5000, 346. 5000, 346. 5000, 346. 5000, 4
20030102, 000300, 346. 0000, 346. 0000, 346. 0000, 346. 0000, 4
20030102, 002000, 346. 0000, 346. 0000, 346. 0000, 346. 0000, 4
20030102, 002500, 345. 9000, 345. 9000, 345. 9000, 345. 9000, 4
20030102, 003200, 345. 8000, 345. 8000, 345. 8000, 345. 8000, 4
20030102, 003400, 345. 6000, 345. 6000, 345. 4000, 345. 4000, 4
20030102, 003600, 345. 5000, 345. 5000, 345. 5000, 345. 5000, 4
20030102, 004100, 345. 6000, 345. 6000, 345. 6000, 345. 6000, 4
20030102, 004400, 345. 8000, 345. 8000, 345. 8000, 345. 8000, 4
20030102, 005100, 346. 0000, 346. 0000, 346. 0000, 346. 0000, 4
20030102, 005400, 346. 1000, 346. 1000, 346. 1000, 346. 1000, 4
20030102, 011100, 346. 1000, 346. 1000, 346. 1000, 346. 1000, 4
20030102, 011600, 345. 8000, 345. 8000, 345. 8000, 345. 8000, 4
20030102, 013200, 345. 8000, 345. 8000, 345. 8000, 345. 8000, 4
20030102, 014900, 345. 8000, 345. 8000, 345. 8000, 345. 8000, 4
20030102, 015600, 346. 0000, 346. 0000, 346. 0000, 346. 0000, 4
```

图3-1 原始数据储存形式

```
20030101230100, 1041433260. 0, 346. 5
20030101231700, 1041434220. 0, 346. 5
20030101233400, 1041435240. 0, 346. 5
20030101235100, 1041436260. 0, 346. 5
20030102000300, 1041436980. 0, 346. 0
20030102002000, 1041438000. 0, 346. 0
20030102002500, 1041438000. 0, 346. 0
20030102003200, 1041438720. 0, 345. 8
20030102003400, 1041438840. 0, 345. 6
20030102003600, 1041438960. 0, 345. 5
20030102004100, 1041439260. 0, 345. 6
20030102004400, 1041439440. 0, 345. 8
20030102005100, 1041439860. 0, 346. 0
20030102005400, 1041440040. 0, 346. 1
20030102011100, 1041441060. 0, 346. 1
20030102011600, 1041441360. 0, 345. 8
```

图3-2 数据预处理储存形式

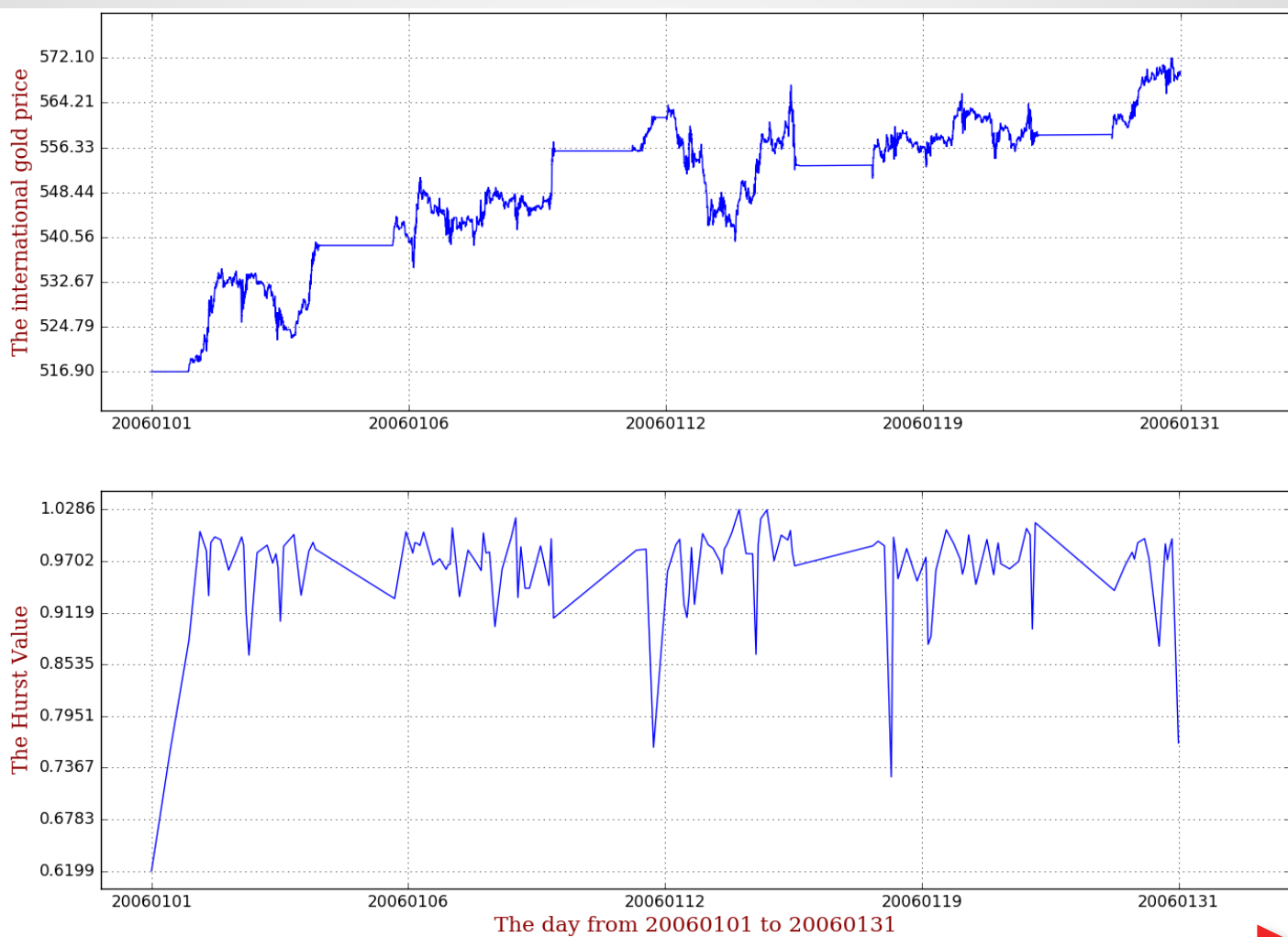


图3-3 非交易日对数据的影响



图3-4 数据预处理方案



## 4. 神经网络的选择与构建

机器学习是用数据或以往的经验，以此优化计算机程序的性能标准。[3]

--Alpaydin

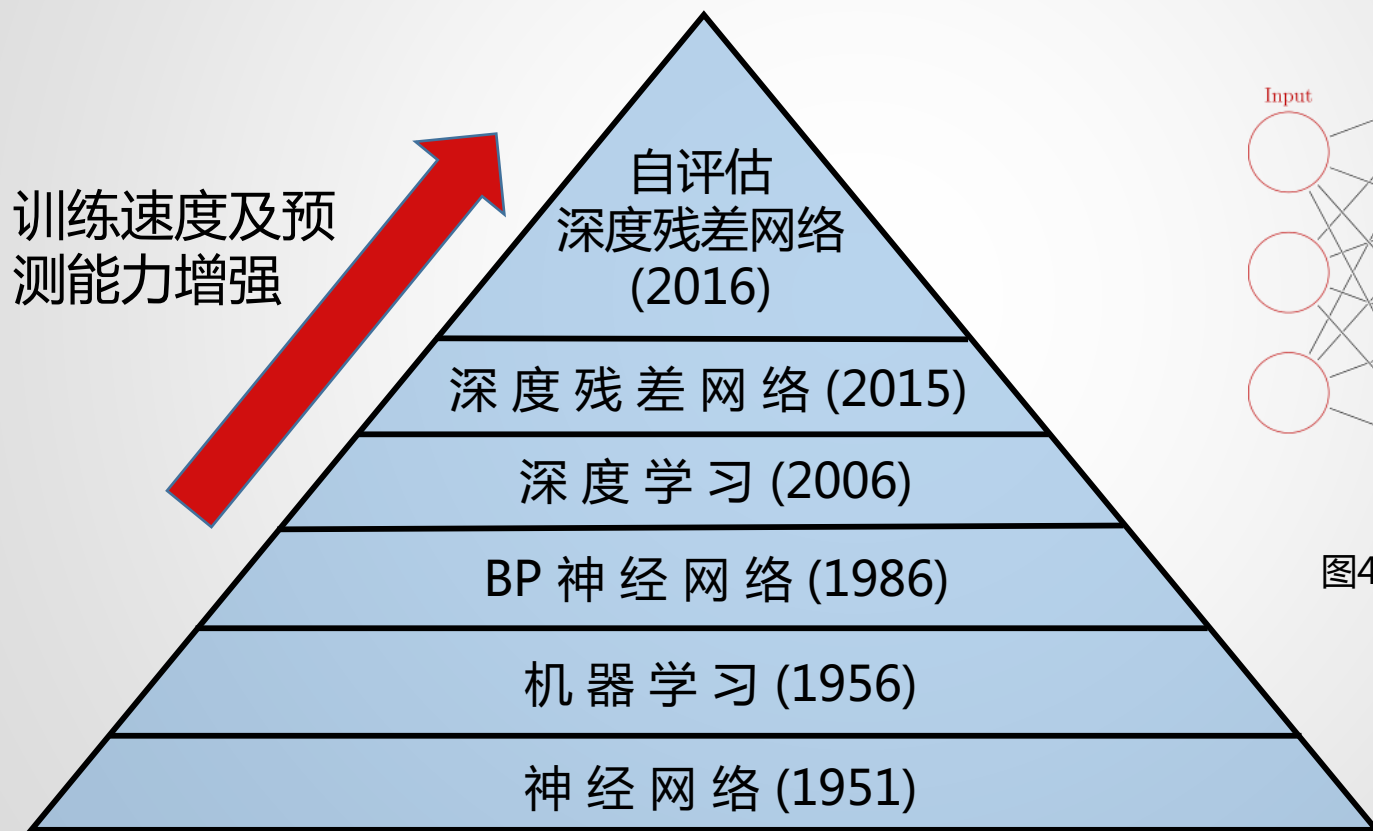


图4-1 机器学习算法发展历史

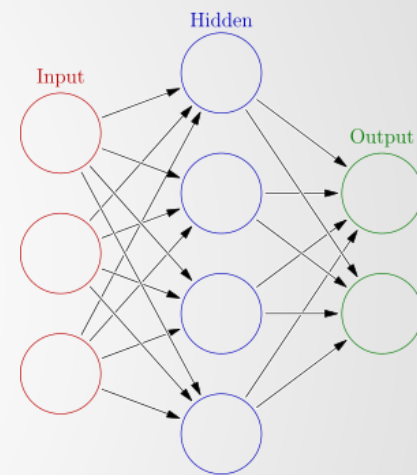


图4-2 神经网络

## 神经网络的优势

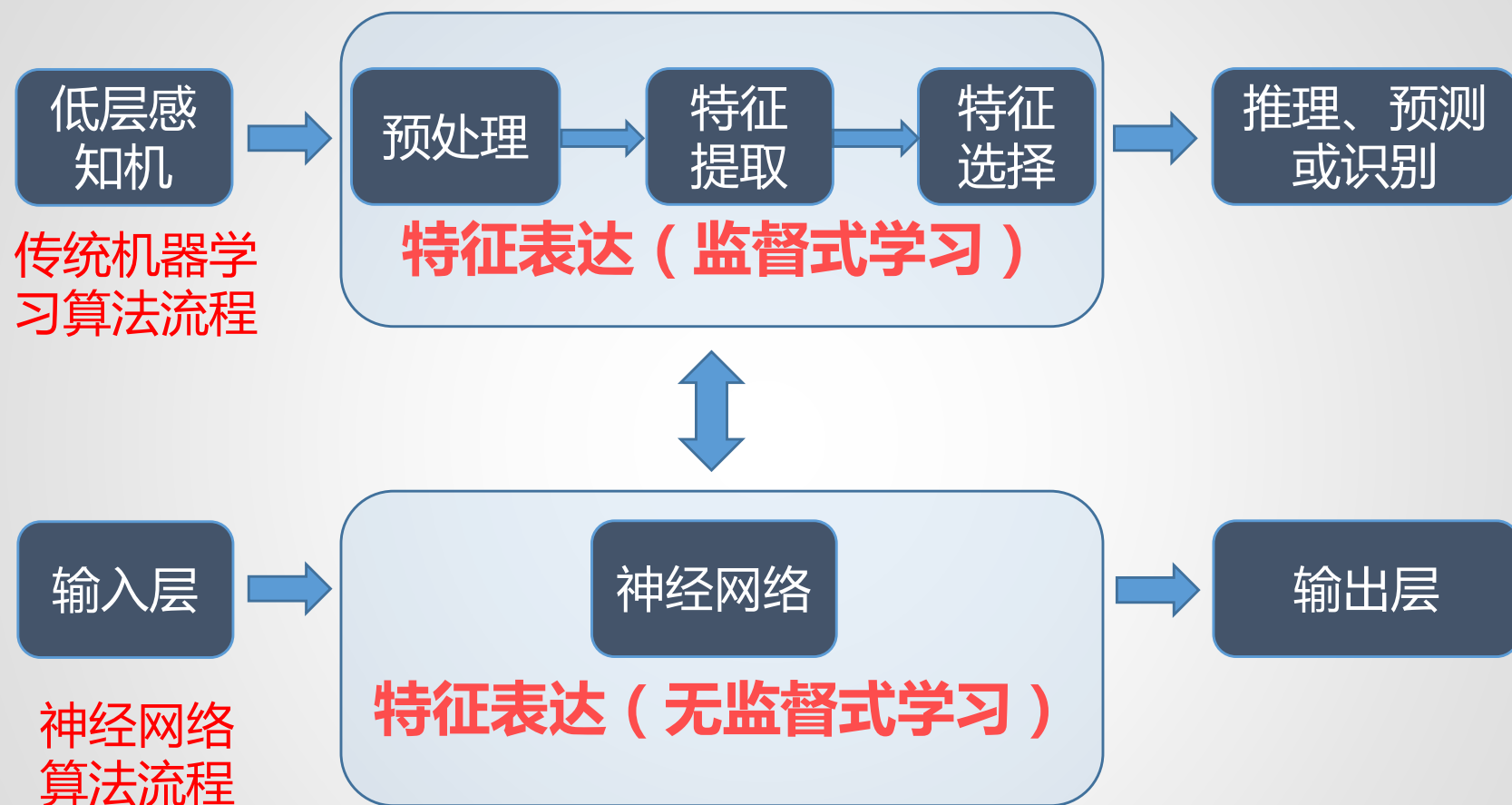
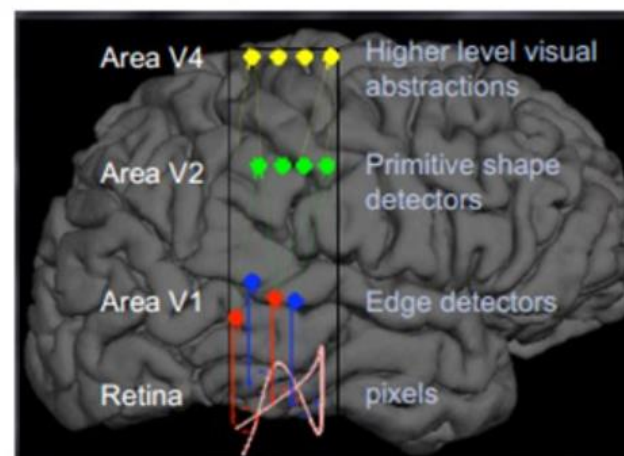
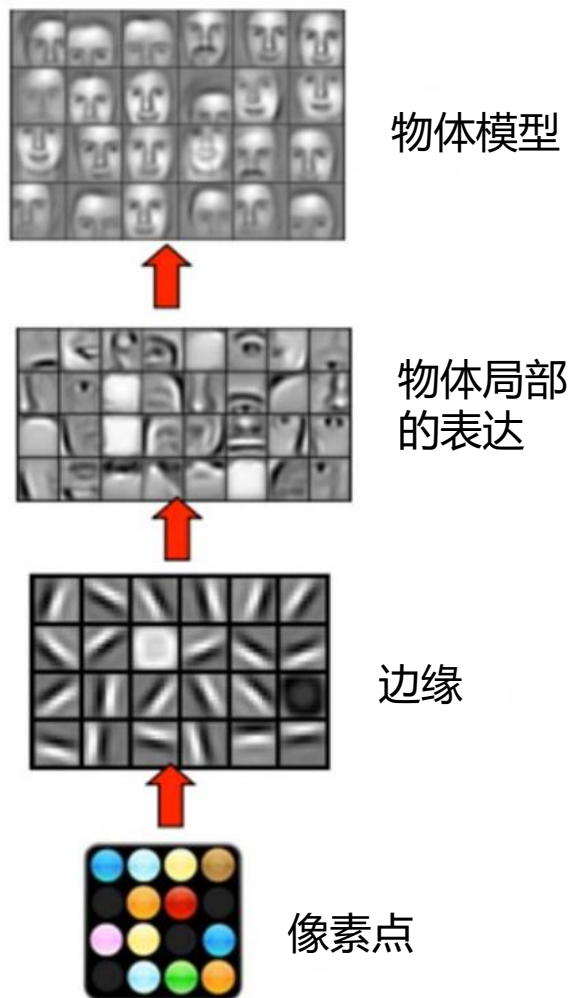


图4-2 传统机器学习算法与深度学习算法对比



深度学习强调了模型结构的深度，通常有5层、6层，甚至10多层的隐层节点。

图4-3 深度学习

## 自评估深度残差网络

深度残差网络的训练中，我们首先求取残差映射  $F(x) := H(x) - x$ ，那么原先的映射就是  $F(x) + x$ 。它俩的学习难度是不一样的。其隐层深度达到了152层。[4]

自评估深度神经网络相对于传统人工神经网络多出了一个潜意识分区。这将极大地提高神经网络的鲁棒性[5]。

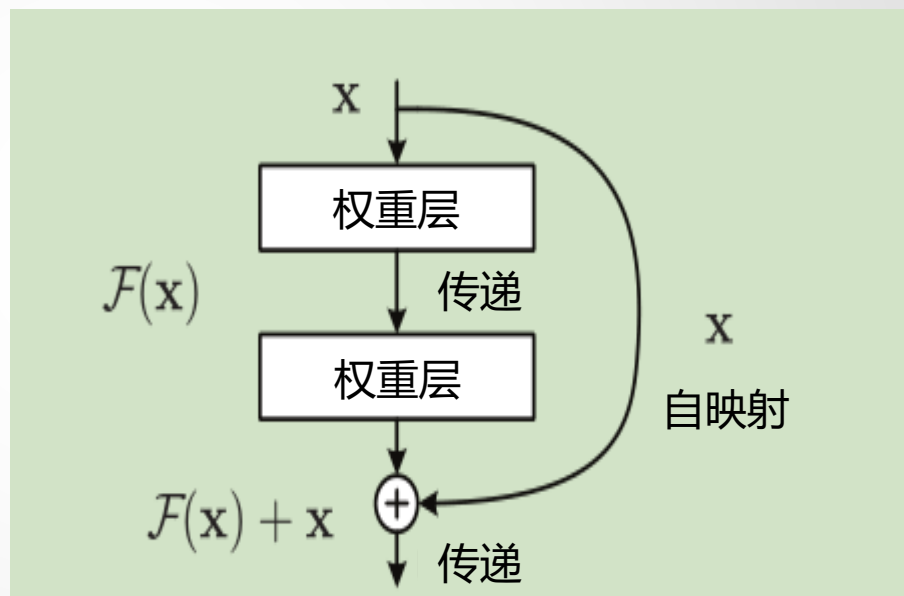


图4-3 深度残差网络的训练





## 5. 总结

本课题已经实现了：

- 黄金价格的可预测性分析
- 国际黄金价格的历史数据获取、筛选与预处理
- 深度残差网络的构建与改进
- GPU数据处理环境的配置

在接下来的工作中将要：

- 确定神经网络合适的权重和传递率
- 获取与分析实时的黄金价格信息以及专家预测信息
- 完成专家系统并对自评估深度残差网络的预测结果进行修正与表达

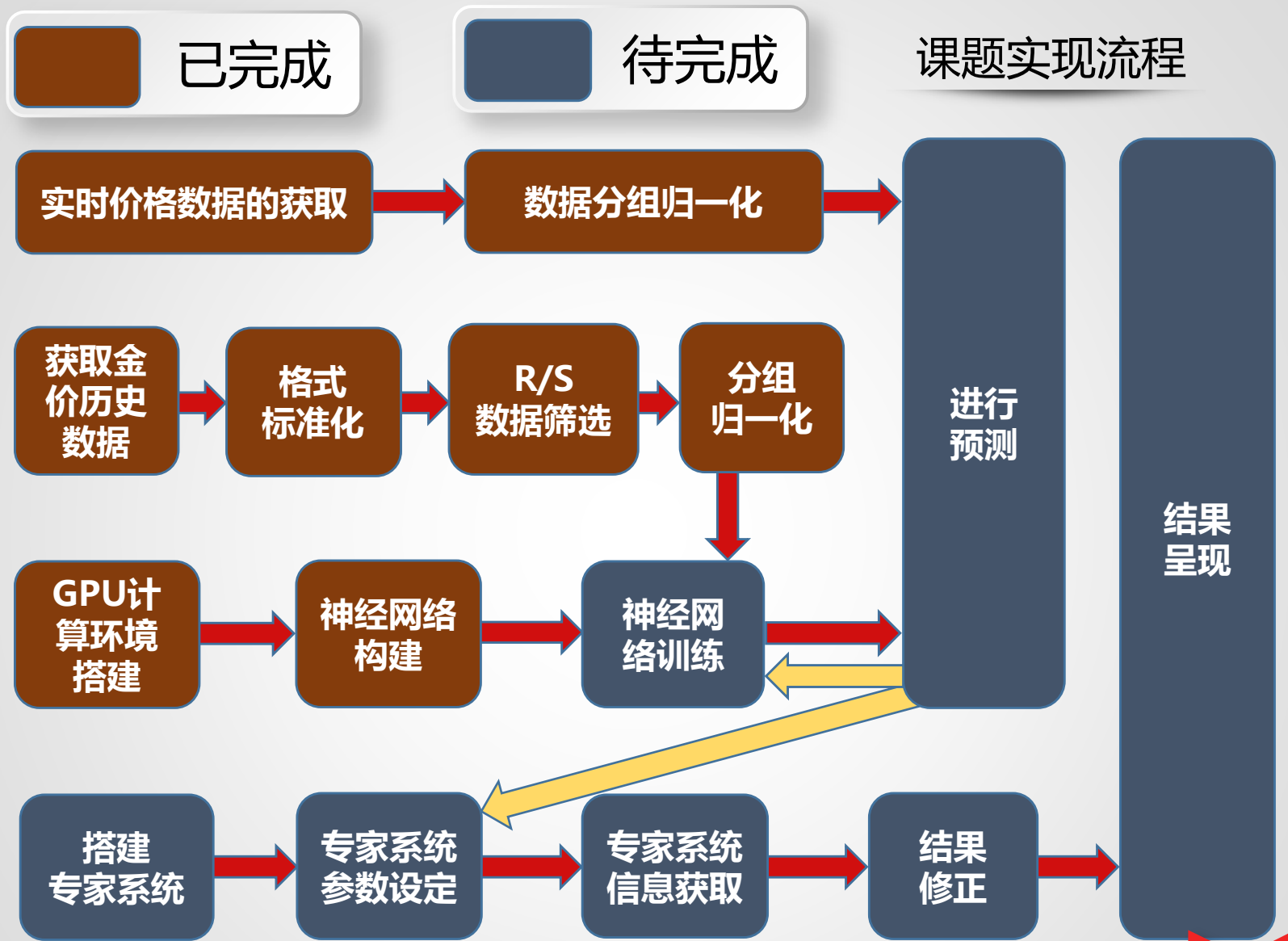


图6-1 课题实现流程

## 参考文献

### References

- [1] Fama E F, French K R. Common risk factors in the returns on stocks and bonds[J]. Journal of Financial Economics, 2010, 33(93):3-56.
- [2] Hurst H E. Long Term Storage Capacity of Reservoirs[J]. Transactions of the American Society of Civil Engineers, 1951, 116(12):776-808.
- [3] Alpaydin E. Introduction to Machine Learning (Adaptive Computation and Machine Learning)[J]. Cdn.intechopen.com, 2004, 5(8):28.
- [4] He. K. and Zhang. X. and Ren. S. and Sun. J. Deep Residual Learning for Image Recognition. ArXiv e-prints, 2015, 1512.03385.
- [5] Wieland A. Dealing with supply chain risks: Linking risk management practices and strategies to performance[J]. International Journal of Physical Distribution & Logistics Management, 2012, 42(10):887-905.

20

谢谢大家

Thank you for your attention!