

基于深度残差网络的国际金价预测研究

The Study of International Gold Price Forecasting Based on Deep Residual Network

汤吉 121143325

指导老师: 张鸿燕 博士

中国民航大学
中欧航空工程师学院

2016 年 06 月 06 日



目录

- 1 课题研究背景
- 2 国际金价可预测性分析
- 3 深度残差网络
- 4 专家系统
- 5 总结



课题研究背景

国际黄金价格

经济全球化：复杂性，不稳定性；

传统经济学理论的局限性；

有效市场假说、分型市场假说；

.....

深度学习网络

层数越高，对数据的“理解”越“深”；

梯度发散和精度下降；

深度残差网络；

.....



R/S 分析法

Hurst 指数是用来量化时间序列的自相关能力的。设有一个长度为 n 的时间序列，则其 Hurst 值 H 的原始定义为：

$$E\left[\frac{R(n)}{S(n)}\right] \sim Cn^H \quad n \rightarrow \infty \quad (1)$$

其中 $R(n)$ 为这 n 个值的极差， $S(n)$ 是它们的标准差， E 表示求期望， C 是一个常数。

Hurst 指数的含义

- ① $0.5 < H \leq 1$ - 序列具有一定的保持趋势的能力；
- ② $0 \leq H < 0.5$ - 序列相邻的值会不断在高值和低值之间切换。

R/S 分析法

对照组设置

- ① 我们选择 2001 年 1 月 2 日至 2014 年 12 月 31 日时间段的国际金价走势，称为实际金价走势。
- ② 实际金价进行小于 30 交易日滤波之后加上均值为 0，方差为当日金价 3% 的高斯白噪声，得到随机模拟金价。

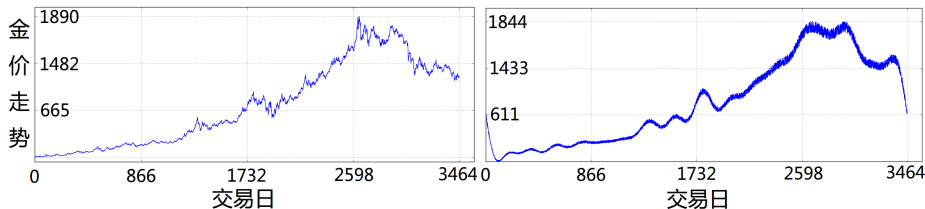


图 1: 实际金价走势（左）与随机模拟金价走势（右）



R/S 分析法

我们以实际金价（蓝色）与随机模拟金价（绿色）以 10 交易日为基础，计算每交易日对应的 Hurst 指数（H）。

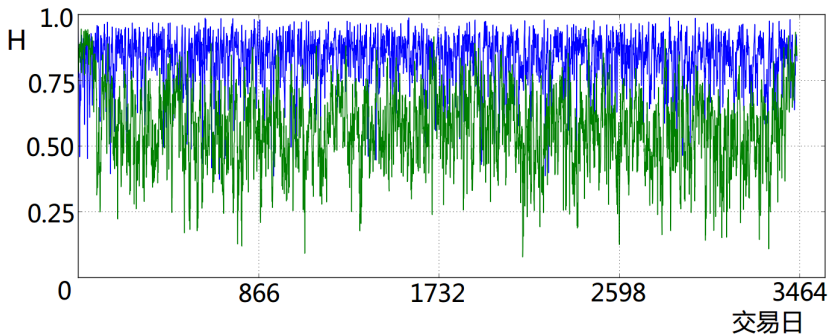


图 2: 实际金价走势（蓝）与随机模拟金价走势（绿）Hurst 指数对比

国际金价预测流程

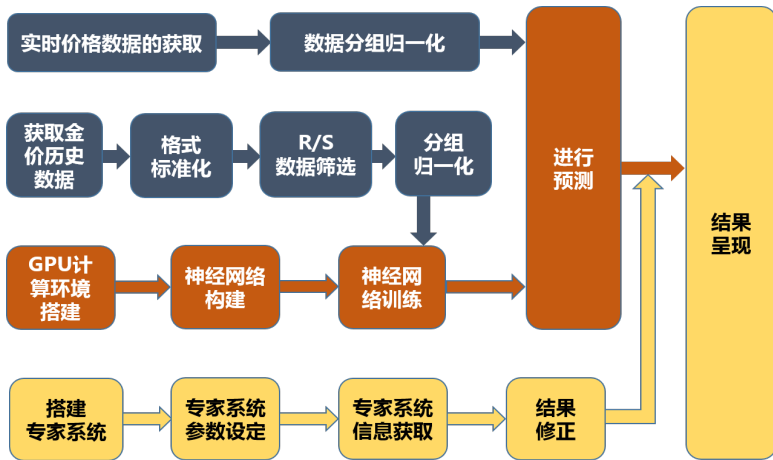


图 3: 国际金价预测流程图

深度学习

假设有一个输入 $Input$ 经过系统 $\{S_1, \dots, S_n\}$ 得到输出 $Output$ ，可以将这个过程 $Input \rightarrow S_1 \rightarrow S_2 \rightarrow \dots \rightarrow S_n \rightarrow Output$ 表示为：

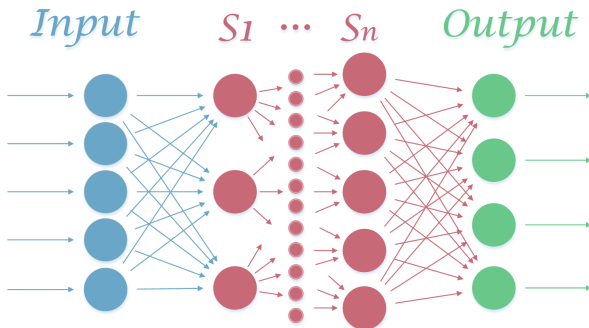


图 4: 深度学习数据处理流程

深度学习自编码器

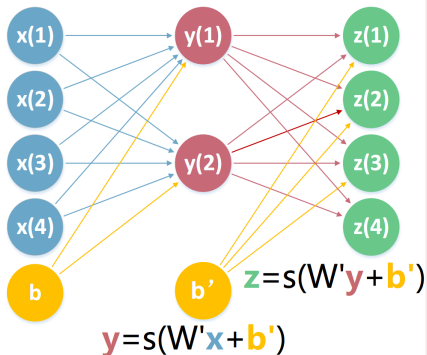


图 5: 自编码器值传递流程

- 1 s 是 Sigmoid 函数

$$s(x) = \frac{1}{1 + e^{-x}} \quad (2)$$

- 2 W 是权重矩阵
- 3 W' 是逆映射权重矩阵

这个模型的目标便是优化得出平均重构误差最小的映射。

$$\text{Argmin}[L(xz)] = \|x - z\|^2 \quad (3)$$

深度残差网络

深度残差网络解决了深度学习中精度下降的问题。

假设深度神经网络的基础映射为 $H(x)$ ，然后构造一个映射：

$$F(x) = H(x) - x \quad (4)$$

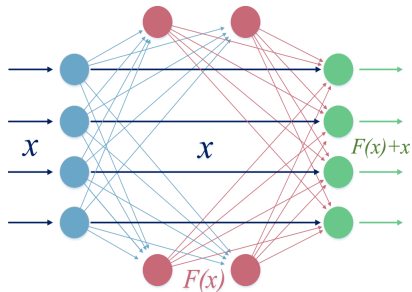


图 6: 深度残差网络值传递过程

国际金价历史数据的获取

表 1: 原始数据（部分）的储存形式

<TICKER>	<DTYYYYMMDD>	<TIME>	<OPEN>	<HIGH>	<LOW>	<CLOSE>	<VOL>
XAUUSD	20010102	230900	268.8	268.8	268.8	268.8	4
XAUUSD	20010102	231200	268.8	268.8	268.8	268.8	4
XAUUSD	20010102	231600	268.9	268.9	268.9	268.9	4
XAUUSD	20010102	233800	268.8	268.8	268.8	268.8	4

通过 **Forextester** 网站可免费
下载国际金价的历史数据。其
数据以纯文本格式储存，提取
其中的有用数据（黄色）并将
时间转换为 **UTC** 格式。

表 2: 时间格式转换后的数据

<DATE>	<UTC>	<CLOSE>
20010102230900	978448140	268.8
20010102231200	978448320	268.8
20010102231600	978448560	268.9
20010102233800	978449880	268.8



历史数据的预处理

设 $[i]$ 表示第 i 天对应的价格，且

P_{Mean} : 均值

LR : 10 日平均对数收益率

P_{Max} : 最大值

XLR : 最大值的 10 日平均对数收益率

P_{Min} : 最小值

$NALR$: 最小值的 10 日平均对数收益率

表 3: 深度残差网络数据输入格式

第一列	第二列	第三列	第四列	第五列	第六列	第七列
$P_{Mean}[i]$	$ALR[i - 5]$	$LR[i - 4]$	$LR[i - 3]$	$LR[i - 2]$	$LR[i - 1]$	$LR[i]$
$P_{Max}[i]$	$XALR[i - 5]$	$XLR[i - 4]$	$XLR[i - 3]$	$XLR[i - 2]$	$XLR[i - 1]$	$XLR[i]$
$P_{Min}[i]$	$NALR[i - 5]$	$NLR[i - 4]$	$NLR[i - 3]$	$NLR[i - 2]$	$NLR[i - 1]$	$NLR[i]$

深度残差网络的搭建环境

在本实验的硬件条件下，深度残差网络总共训练 161,000 次，耗时约 134 小时。

表 4: 深度残差网络的搭建环境

系统版本	Linux Ubuntu 14.04 LTS 64-bit
内存	7.8 GB
硬盘	483.7 GB
CPU	Intel Core i5-2400 @ 3.10GHz * 4
GPU	NVIDIA GeForce GTX 950/PCIe/SSE2
编程语言	Python
开源库	“numpy”、“theano”、“lasagne”等
源代码	lasagne_residual_network(Github)



深度残差网络的训练过程

判断预测准确性的规定

- 1 若某一日的预测与实际情况同为上涨或下跌，称为预测正确，否则称为预测不正确；
- 2 若某一日的预测与实际情况同为大涨跌幅或小涨跌幅，称为涨跌幅预测正确，否则称为涨跌幅预测不正确。

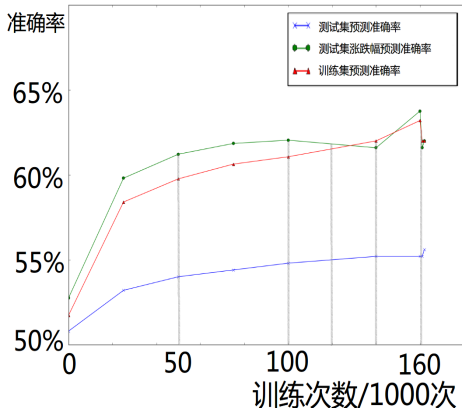


图 7: 训练的准确率变化曲线

深度残差网络的预测结果及分析

利用训练完成的深度残差网络进行 25 交易日的金价走势。

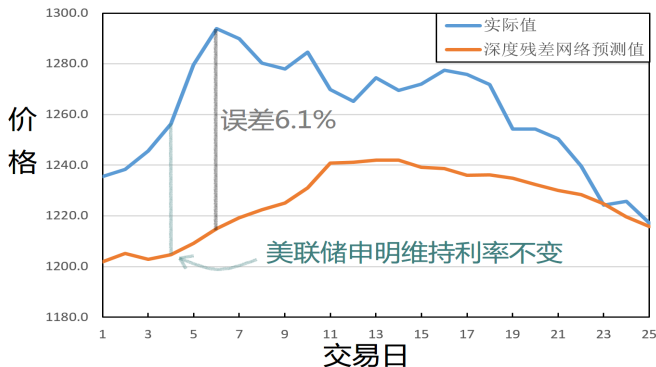


图 8: 2016 年 4 月 23 日至 5 月 27 日预测结果

专家系统知识库的数据获取

专家系统的数据来源

出于技术条件的限制，暂时使用“今日头条”门户网站的信息作为专家系统知识库的数据来源。

爬虫软件将检测出“今日头条”搜索页面中的新闻网址链接并进行访问，然后将所有与“黄金”和“美元”相关的新闻备份至本地知识库。

今日头条 你关心的 才是头条 首页 动态

Q 国际黄金 搜索

搜索结果：

黄金下午茶_国际黄金首饰价格还涨不涨？（6月2日）

恒信贵金属 17:22

今日国际黄金价格

中金网 09:40

国际黄金ETF持仓增加4.46吨 白银ETF增持48.52吨

中金网

图 9: 今日头条搜索结果

专家系统推理库的设计

专家系统对知识库中的信息进行分析和关键词提取。

Julius Baer(宝盛)的大宗商品研究分析师Carsten Menke在一份报告中称,“美元的反弹和期货市场情绪的冷却**推动金价跌向每盎司1200美元。**”

“我们对黄金保持一个中立的观点,但认为保险起见投资者应该在他们的投资组合中**继续考虑黄金**,这是鉴于当前经济和金融市场的普遍存在风险的背景。”

近期美联储的高官们纷纷暗示在未来几个月加息是适当的,黄金因此一直承受较大的卖盘压力。上周,美联储主席耶伦表示,在未来几个月如果经济数据改善加息是可能的。此番言论更是进一步增加**黄金持续下降**的压力。

SignalPro的首席市场策略师桑迪Sandy Jadeja周二向媒体表示,他**短期内看好黄金**,尤其是在当前市场整体不确定的环境下。

他补充道:“我**看涨黄金至1300美元**每盎司。最近黄金跌破了50日均线,这是交易员在中期都会很关注的一个技术指标,在这会形成支撑。”他们上调黄金的**目标价至每盎司1420美元**。不过同时他也警告称,一旦黄金达到这一水平将再次面临挑战。“接下来会发生什么?这一切都要看美联储的脸色!”

图 10: 推理库对关键词搜索结果



专家系统推理库的设计

根据最大后验概率原理，确定权重系数 k ：

$$k = \frac{(u - d)^2}{u + d} \quad (5)$$

其中 u 和 d 分别表示预测金价上涨和下跌的关键词的出现次数。

然后计算得到修正后的金价 CP ：

$$CP = \frac{kD + P}{1 + k} \quad (6)$$

其中 D 表示推理库的预测值， P 表示深度残差网络的预测值。

表 5: 专家系统 5 月 24 日至 5 月 26 日知识库数据

日期	实际值	深度残差网络预测值	推理库预测值	上涨关键词数	下跌关键词数	修正预测值
20160524	1239.7	1228.4	1241	66	119	1240.2
20160525	1224.2	1224.8	1243	22	93	1242.5
20160526	1225.8	1219.6	1224	46	106	1223.9



专家系统的修正结果与分析

我们利用专家系统推理库对 25 交易日的金价预测结果进行修正。

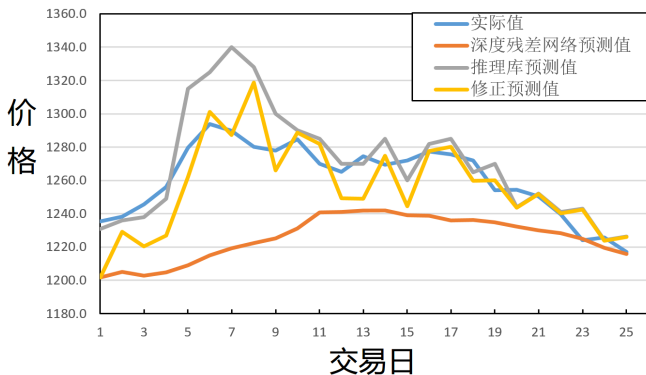


图 11: 推理库修正预测值与深度残差网络预测值对比

专家系统的修正结果与分析

结果分析

深度残差网络预测均方误差（误差平方的均值）为 **1682.7**，专家系统推理库预测均方误差为 **370.6**，深度残差网络加权修正预测均方误差为 **292.8**。

深度残差网络预测了国际金价市场的内部规律，专家系统提供对于国际金融事件的应变能力，两者互补得到预测均方误差最小的加权修正预测值。



总结

- 1 在规定时间内完成了课题所要求的全部任务。
- 2 总共编写超过 2000 行 Python 代码，编程能力有了一定的突破，对 Python 编程语言有了更深的理解。
- 3 查阅深度学习、国际金融学大量书籍，对于计算机科学和经济学有了更深的认识和学习兴趣。
- 4 知识整合能力得到提升，在整个课题中，运用不同平台下各种工具综合解决问题。



展望

- 1 利用纯粹的深度算法进行金融市场的预测研究不建议继续进行。
- 2 专家系统的预测准确性极高，值得进一步的研究。可以尝试从增加专家系统搜索关键词的数量、增加门户网站的搜索数量等方面，提高专家系统知识库的覆盖面。
- 3 可以考虑利用深度学习结合专家系统的爬虫软件，进行“智能”的数据挖掘。
- 4 专家系统可以考虑支持多语言的数据处理。



致谢

感谢各位的聆听! 请各位老师给予批评和指导!

