

**本科毕业设计（论文）**



基于深度残差网络的国际金价预测研究

**学生姓名： 汤 吉**

**专 业： 飞行器动力工程**

**学 号： 121143325**

**指导教师： 张鸿燕**

**所属学院：中欧航空工程师学院**

**二〇一六年六月**

中国民航大学

本科毕业设计（论文）

**基于深度残差网络的国际金价预测研究**

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

**学生姓名：汤吉**

**专 业：飞行器动力工程**

**学 号：121143325**

**指导教师：张鸿燕**

**学 院：中欧航空工程师学院**

**2016年6月**

**创见性声明**

本人声明：所呈交的毕业论文是本人在指导教师的指导下进行的工作和取得的成果，论文中所引用的他人已经发表或撰写过的研究成果，均加以特别标注并在此表示致谢。与我一同工作的同志对本论文所做的任何贡献也已在论文中作了明确的说明并表示谢意。

毕业论文作者签名： 签字日期： 年 月 日

**本科毕业设计（论文）版权使用授权书**

本毕业设计（论文）作者完全了解中国民航大学有关保留、使用毕业设计（论文）的规定。特授权中国民航大学可以将毕业设计（论文）的全部或部分内容编入有关数据库进行检索，并采用影印、缩印或扫描等复制手段保存、汇编以供查阅和借阅。同意学校向国家有关部门或机构送交毕业设计（论文）的复印件和磁盘。

（保密的毕业论文在解密后适用本授权说明）

毕业论文作者签名： 指导教师签名：

签字日期： 年 月 日 签字日期： 年 月 日

摘 要

随着世界经济的全球化，国际金价波动的不稳定性与日俱增，基于线性或非线性模型的传统经济学理论已经不能很好地解释黄金价格的变化规律。计算机科学的飞速发展，为需要极其庞大的计算量的非线性算法提供了有力的支持。因此基于深度残差网络的国际金价预测的研究，具有广阔的研究前景。

首先在国际金价历史数据的处理过程中，应用R/S分析法（Rescaled Range Analysis，重标极差分析法）证明在14日内（10交易日内）国际金价的自相关性与可预测性；然后对XAU/USD（纽约现货金价/美元）的时间序列进行数据分析预处理及数据集封装，并建立基于GPU（Graphics Processing Unit，图形处理单元）计算的深度残差网络，用于数据集的训练和预测；最后经专家系统修正，得出具有一定准确率的预测结果。

在实际的研究过程中，经过训练的深度残差网络对于训练集达到的涨跌预测正确率，对于测试集达到的涨跌预测正确率，对于测试集涨跌幅的预测准确率达到。同时专家系统对互联网上金融专家或金融组织的国际金价预测数据进行数据挖掘并用于深度残差网络预测结果的修正。研究成果对于金融机构或普通投资者都有一定的参考价值。

**关键词：**国际黄金价格；深度残差网络；深度学习；Hurst指数；专家系统

# 

Abstract

International gold price volatility is daily multiplied with the globalization of the world economy. Therefore the traditional economic theory, with the linear or nonlinear model, could not explain the change of the gold price ideally. The computer science mushrooms provide a powerful support for the nonlinear algorithms which require a huge amount of computation. For this reason, the study of international gold price forecasting based on deep residual network has broad research prospects and great research value.

Firstly, the autocorrelation and predictability of international gold price have been proved within 14 days (10 trading days) by using Rescaled Range Analysis method in the processing procedure of the historical international gold price data. Secondly, the time series of XAU/USD has been analyzed and prepossessed into capsulations. Thirdly, the Deep Residual Network has been built for the training and testing of the data on the Graphics Processing Unit. Finally, the forecast result has a certain accuracy after the correction of expert system.

In the study, the trained Deep Residual Network has the accuracy of for the training data, the accuracy of for the testing data and the accuracy of for the amount of increase and decrease of the testing data. At the same time, the expert system has gotten the international gold price data which is used for result correction of Deep Residual Network under the aid of data-mining. There is a certain reference value of the study results for financing institutions or ordinary investors.

**Key Words:** international gold price; Deep Residual Network; Deep Learning; Hurst exponent; expert system

目 录

[第1章 绪论 1](#_Toc452398852)

[1.1 课题研究背景及意义 1](#_Toc452398853)

[1.1.1 课题研究背景 1](#_Toc452398854)

[1.1.2 课题研究意义 2](#_Toc452398855)

[1.2 国内外研究综述 2](#_Toc452398856)

[1.1.3 深度学习研究现状 2](#_Toc452398857)

[1.1.4 国际黄金价格预测的研究现状 3](#_Toc452398858)

[1.2 本文的研究内容与结构 3](#_Toc452398859)

[1.3 本文的创新之处 3](#_Toc452398860)

[第2章 R/S分析法与国际金价可预测性分析 5](#_Toc452398861)

[2.1 R/S分析法 5](#_Toc452398862)

[2.1.1 R/S分析法的起源 5](#_Toc452398863)

[2.1.2 Hurst指数的定义 5](#_Toc452398864)

[2.1.3 R/S分析法Hurst指数的估计 5](#_Toc452398865)

[2.1.4 Hurst指数值的解释 6](#_Toc452398866)

[2.2 国际金价的可预测性分析 6](#_Toc452398867)

[2.2.1 有效市场的定义与分类 6](#_Toc452398868)

[2.2.2 分形市场假说 6](#_Toc452398869)

[2.2.3 国际黄金市场的R/S分析 7](#_Toc452398870)

[2.3 本章小结 9](#_Toc452398871)

[第3章 深度残差网络 11](#_Toc452398872)

[3.1 深度残差网络概述 11](#_Toc452398873)

[3.1.1 深度学习简介 11](#_Toc452398874)

[3.1.2 深度学习的基本思想 11](#_Toc452398875)

[3.1.3 深度学习的原理 12](#_Toc452398876)

[3.1.4 深度残差网络的原理 13](#_Toc452398877)

[3.2 国际金价数据的获取与预处理 14](#_Toc452398878)

[3.2.1 国际金价走势历史数据的获取 14](#_Toc452398879)

[3.2.2 国际金价历史数据的预处理方法 14](#_Toc452398880)

[3.3 深度残差网络的构建 15](#_Toc452398881)

[3.3.1 系统环境设置 15](#_Toc452398882)

[3.3.2 构建适用于国际金价预测的深度残差网络 16](#_Toc452398883)

[3.3.3 利用预处理的数据集对深度残差网络进行训练 16](#_Toc452398884)

[3.3.4 深度残差网络的训练过程 16](#_Toc452398885)

[3.4 预测结果与分析 18](#_Toc452398886)

[第4章 专家系统 19](#_Toc452398887)

[4.1 专家系统 19](#_Toc452398888)

[4.1.1 专家系统简介 19](#_Toc452398889)

[4.1.2 专家系统的搭建环境 19](#_Toc452398890)

[4.1.3 专家系统在国际金价预测研究中的意义 19](#_Toc452398891)

[4.2 专家系统的搭建 19](#_Toc452398892)

[4.2.1 专家系统数据来源 19](#_Toc452398893)

[4.2.2 爬虫软件设计 20](#_Toc452398894)

[4.2.3 专家系统的数据提取实例 20](#_Toc452398895)

[4.3 预测结果与分析 21](#_Toc452398896)

[第5章 结论 24](#_Toc452398897)

[5.1 研究内容概述 24](#_Toc452398898)

[5.2 研究过程的不足之处 24](#_Toc452398899)

[5.3 继续研究的方向 24](#_Toc452398900)

[致 谢 27](#_Toc452398901)

[附录A：网址链接 28](#_Toc452398902)

[附录B：程序调用过程及专家系统知识库部分信息 29](#_Toc452398903)

[附录C：外文翻译资料 30](#_Toc452398904)

[附录D：外文翻译资料译文部分 43](#_Toc452398905)

1. 绪论
   1. 课题研究背景及意义
      1. 课题研究背景

黄金自古以来被人们视为永恒的金属，象征着至高无上的财富，在历史上曾被作为货币使用，直到现在也在一些国家和地区经济中保持一定的货币职能。国际金价与国际经济形势具有千丝万缕的联系，世界各国历来对国际金价的走势相当关注。同时各大公司、金融机构及个人也把黄金作为一种投资商品，希望通过对国际金价的预测谋取高额利润。因此，国际金价走势预测的研究不论是对国际金融研究领域还是具体至投资机构、个人，都具有极其重要的意义。

随着世界金融领域向着一体化发展，国际黄金市场的动态性和复杂性日益增加，这就使得国际黄金市场的预测分析更加复杂。现在的黄金市场已经表现出非常明显的非线性机制，若仅基于传统经济学理论，其中极其复杂的内在联系已经很难被经济学家完全掌握。

2006年后，深度学习算法是机器学习算法研究领域中，具有革命性的研究成果。其相对于神经网络具有更加“深”的层次，因此对抽象的数据具有更为突出的“深度”学习能力。尤其现在，深度学习算法在较为抽象的机器学习领域已经获得了巨大的成功，如语音识别和图像识别等[1]。对国际金价走势的预测而言，传统经济学暂时无法完全解释大规模金价数据的内在联系，依靠深度学习算法进行大规模数据的训练容易取得更好的效果。

2015年12月第六届ImageNet年度图像识别测试中，微软研究院的计算机图像识别的最新研究成果：深度残差网络（Deep Residual Network）[2]在几个类别的测试中获得第一名，其深度学习网络的深度甚至尝试过1000层。对于人工神经网络来说，其层数的深度对于数据的理解能力而言是一个关键因素，并且越深的网络其预测效果越好，对数据内在含义的理解也更为深刻。

2016年3月15日的“人机大战”中，AlphaGo以4:1大比分战胜“人类”。AlphaGo是Google针对围棋布阵最优化问题研发的一套设计周密，调用巨大的Google云计算资源和GPU通用计算能力的深度学习引擎。虽然现在AlphaGo的取胜并不完全意味着人工智能已经超越人类智能，但它足以证明深度学习算法研究方向无尽的潜力。

因此研究中搭建深度残差网络进行国际金价的预测，也有理由相信这个研究将获得优异的预测效果。

* + 1. 课题研究意义

本课题基于深度残差网络进行国际金价预测的研究具有的意义：

首先，对于个人来说，黄金投资是一种投资的手段。用黄金的投资可以抵御通货膨胀和经济动荡，从而达到保值、规避风险的目的，甚至还可以利用黄金波动过程中的差价谋取少量的利润。

然后，对于社会而言，目前各大公司和金融机构一直在尝试各种方法对国际汇率、国际金价进行预测，降低它们的波动所带来的风险、损失，或提高利用其波动所谋取的利润。金价的预测能为相关金融机构、组织的投资行为提供一定的支持，使得它们的资金分配与利用更加合理，减少损失或获取更高的利润。

最后，从国家的角度看来，2013年中国的黄金消费量已经突破1000吨，同比增长41%，中国也已经成为全球最大的黄金消费国。自古以来中国人就把黄金作为身份地位的象征，随着居民财富快速增加，正处于飞速发展阶段的中国的黄金消费量还将会继续增长。因此中国的黄金市场发展潜力巨大，国际金价的问题研究，能够增强国家在国际金融活动中的应变能力，对维持国家经济的稳定同样会有很大的作用。

* 1. 国内外研究综述
     1. 深度学习研究现状

Hochreiter与Schmidhuber 在1997研究出长时记忆神经网络（Long-Short Term Memory，LSTM）[3]。2009年，在没有任何先前知识储备的情况下，深度多维LSTM神经网络在手写识别方面，识别三种语言赢得2009年第三届ICDAR竞赛[4]。

在2010年，Dan Ciresan与他的同事[5]在瑞士人工智能实验室IDSIA表示，尽管人工神经网络在训练过程中会出现“梯度消失”的问题，但是，“GPU的超强处理能力使得具有多隐层的普通反向传播神经网络的训练变得可行，并且该方法在著名的MNIST手写数字识别训练库的训练中，表现优于以前的任何机器学习技术。”

同时，深度学习也进入语音识别的领域。微软研究院、多伦多大学相关领域的研究人员在2010年证实：利用深层神经网络的输出层接口，能够一定程度上减少大词汇量语音识别的工作量，如：语音搜索。这个方法最近已经被广泛应用于生活中，例如Google利用这一项技术极大提高了智能手机的语音识别率[6]。

直到2011年，正向传播卷积神经网络（Convolutional Neural Network，CNN）在全连接或稀疏连接的分类层具有很大的优势，因为它们的训练通常都不需要监督式的预训练。自2011年之后，在GPU的支持下，这个算法在众多识别竞赛中获胜，例如在2012年进行的分割神经元结构的ImageNet竞赛[1]，EM组挑战[7]等。

随着CNN网络的发展，大家发现人工神经网络的层数对于数据的理解能力而言是一个关键因素，并且越深的网络层数预测效果越好。但是随着网络层数的增加，问题也随之而来，例如梯度发散[8]和精度下降[9]的问题。前一个问题可以通过参数初始化的方法得到解决，对于后面这一个问题，性能下降并不是因为人们普遍认为的深度神经网络的过拟合的原因导致的。直到2015年，深度残差网络解决了深度神经网络训练中精度下降的问题，使用了152层甚至尝试了1000层神经网络。

* + 1. 国际黄金价格预测的研究现状

目前，各国政府、银行、经合组织以及私营部门等，都在从事金融方面的预测研究。对特定数据的预测研究中，历史数据对未来市场行情产生影响的假设被广泛应用于各种模型[10]。经济预测的过程类似于数据分析，又综合了未来的关键经济变量的估计值。现在的统计学模型在金融时间序列的预测当中，没法有效地解决其不确定性和不稳定性的问题。而深度学习神经网络，尤其是本研究中使用的深度残差网络能够“高深度”地学习金融时间序列当中的内在联系，有效地克服其不确定性和不稳定性的问题[11]。

因此深度残差网络具有以前的深度学习神经网络所无法比拟的巨大优势（一般可以达到超过100层的深度）。而在非常需要对数据内在联系进行理解学习的金融时间序列预测领域，暂时还没有人尝试利用深度残差网络进行这个领域预测相关的研究。

* 1. 本文的研究内容与结构

本文的研究思路是：基于经济学市场理论，利用R/S分析法研究国际金价的相关性和可预测性，首先证明本研究的可行性；接下来通过互联网免费获取可靠的国际金价15年历史数据；利用基于Python编程语言的程序，对得到的数据集进行预处理；然后将数据集用于深度残差神经网络（搭建在Linux系统上并利用GPU加速计算）的训练与测试；最后结合专家系统获取的信息修正得到预测结果。

在这里，第一章为引言部分，介绍课题研究的背景及意义，分析国内外在深度学习和国际金价预测方面的研究现状，说明本文的研究思路以及创新之处。

第二章介绍R/S分析法和金融市场的假说理论，综合研究国际黄金价格预测的可行性。最终得出：在一定时间尺度下，国际金价时间序列具有一定自相关性并是可预测的。

第三章给出深度学习和深度残差网络的定义及简单原理，详细介绍深度残差网络的输入数据来源、预处理方法，搭建系统的环境，训练过程和预测结果分析。

第四章详细描述专家系统知识库的数据获取、处理，推理库的修正权重计算方法和最终权重修正后的结果分析。

最后是论文的结论，对前面章节的研究过程和预测结果进行分析，对论文的研究过程中遇到的问题做出总结，并提出在未来基于本研究成果的研究方向。

* 1. 本文的创新之处

本文有两项创新：

一是首次将高深度的深度残差网络应用到国际金价预测的研究当中，用深度达到81层的深度残差神经网络来进行国际黄金价格走势的预测。与普通神经网络相比，通过训练得到的深度残差网络对于训练数据具有更深刻的学习与理解，同时解决了高深度的神经网络的精度下降问题。

二是将专家系统与深度残差网络相结合。专家系统整合互联网上的金融专家和金融组织的资源来提高深度残差网络预测的鲁棒性，巧妙地克服了深度学习神经网络由于信息获取不全而产生的弊端。

1. R/S分析法与国际金价可预测性分析
   1. R/S分析法
      1. R/S分析法的起源

“Hurst指数”或“Hurst系数”由研究员Harold Edwin Hurst在研究罗河旱涝更替的情况时，为研究水利的实际问题发明，以衡量时间序列的长期记忆能力[12]。Hurst指数又被称为“指数依赖性”或“指数长期依赖性”，它能够量化时间序列的相对趋势。

现在有很多Hurst指数估计值的算法，最有名的就是Mandelbrot和Wallis基于Hurst的水利研究结果使用的重标极差R/S方法[13]。

* + 1. Hurst指数的定义

设有一个序列 ，是观测到的时间跨度，是前*n*个值的取值范围，是它们的标准差，符号表示求期望值，是一个常数。则序列X的Hurst指数（后面以表示）的原始定义如式(2-1)：

|  |  |  |
| --- | --- | --- |
|  | 当*n* | (2-1) |

* + 1. R/S分析法Hurst指数的估计

利用R/S分析法，把一个长度总共为的时间序列分成长度分别为的短序列。对于每一个，便可以计算其重标极差。

而对于一个长度为的序列,便可以由以下步骤计算R/S：

1. 计算平均值：

|  |  |  |
| --- | --- | --- |
|  |  | (2-2) |

1. 计算均值调整序列：

|  |  |  |
| --- | --- | --- |
|  |  | (2-3) |

1. 计算累计偏离序列：

|  |  |  |
| --- | --- | --- |
|  |  | (2-4) |

1. 计算序列极差：

|  |  |  |
| --- | --- | --- |
|  |  | (2-5) |

1. 计算无偏标准差序列：

|  |  |  |
| --- | --- | --- |
|  |  | (2-6) |

1. 计算重标极差序列：

|  |  |  |
| --- | --- | --- |
|  |  | (2-7) |

1. 计算长度为的部分序列的平均值。

数据的Hurst指数满足幂定律，便可以画出关于的图形，通过拟合直线的斜率得到的值。

* + 1. Hurst指数值的解释

若一个时间序列的Hurst指数处在0.5至1之间，则表示它具有一定的自相关能力（一个较高值的序列之后更有可能会跟着另一个较高值的序列）。而Hurst指数处于0至0.5之间的时间序列，其相邻的值会不断在高值和低值之间切换（一个单一的高值更有可能会跟着一个低值，而后又跟着另一个高值）。因此我们可以把Hurst指数为0.5的时间序列当作前后无关的序列。

* 1. 国际金价的可预测性分析
     1. 有效市场的定义与分类

Eugene Fama在1965年首次提出有效市场的概念[14]：在有效的市场当中，充满了具有足够理性、追求利益最大化并且充分掌握当前市场中一切重要信息的投资者。他们当中的每一个人都积极地进行投资，试图通过自己对于市场足够理性的判断，预测出市场中商品的未来价格趋势。因此在这样定义的有效市场中，充分理智且对于市场具有充分认识的投资者们将会对市场产生这样的影响：在市场交易当中的任何时期，股票的实际市场价格都能够完全反映其中一切的市场信息（包括市场已有的价格以及即将发生改变的价格走势）。然后根据Eugene Fama有效市场的概念，可以将资本市场划分为[弱式有效市场](http://baike.baidu.com/view/1557884.htm)、半强式有效市场和强式有效市场三种形式[15]。

* + 1. 分形市场假说

在1994年，Edgar E. Peters的分形市场假说[16] 反驳了Eugene Fama的有效市场假说，他认为资本市场当中的投资者各自都具有不同的思考角度、投资起点和投资背景：

1. 对于市场信息的获取不完全对等；
2. 对于同样的市场信息，不同投资者的交易时间以及交易决策也都不完全相同；
3. 在某时间段内的交易也未必是完全均匀的；
4. 不同投资者的理性程度也是各不相同的，并不是所有投资者都会按照完全理性的方式进行投资；
5. 在对市场信息的反应速度上，有些投资者接受到市场信息之后马上就会作出相应的反应，然而大多数投资者会等待确认市场信息的可靠性，甚至等到趋势显现的已经十分明显时才会作出相应的反应。

Edgar E. Peters最后使用R/S分析法[13] ，证明资本市场中的价格收益符合有偏的随机游走规律。

* + 1. 国际黄金市场的R/S分析

根据第3章的方法，我们获取了纯文本格式的国际金价历史数据，并通过简单的预处理程序得到国际金价关于UTC（Universal Time Coordinated，国际标准时）变化的数组。我们选择时间段2001年1月2日至2014年12月31日总共3464交易日的国际金价走势，作为时间序列一（称为实际金价）。为得到对照组，我们把它看作一段时间序列，采用傅里叶变换滤去所选时间段内周期小于30交易日的信号，并在新曲线的基础上加上同样长度，最大振幅为当日金价3%的高斯白噪声，得到时间序列二（称为随机模拟金价）。

我们选用交易日及当日之前连续9个交易日的国际金价作为时间序列。根据算法2-1所示的流程计算此时间序列对应的作为当日的Hurst指数。

算法2‑ 国际金价序列Hurst指数计算流程

|  |
| --- |
| 算法：计算数列的Hurst指数 |
| 输入：  输出：  1. While do  2. 根据2-2式计算序列的平均值；  3. 根据2-3式构造均值调整序列；  4. 根据2-4式得到累计偏离序列；  5. 根据2-5式计算序列范围；  6. 根据2-6式得到此序列标准差；  7. 根据2-7式计算此序列的重标极差；  8.  9. ；  10. 为偶数：  11. ；  12. ；  13. 为奇数：  14. ；  15. ；  16. End while；  17. 拟合曲线 计算其斜率得到； |

我们对序列一和序列二分别计算结果如图2-1、图2-2所示。它们分别对应实际金价和随机模拟金价在3464交易日内的走势(The international gold price - The days)以及每交易日(The days)对应的Hurst指数走势(The Hurst Value)。

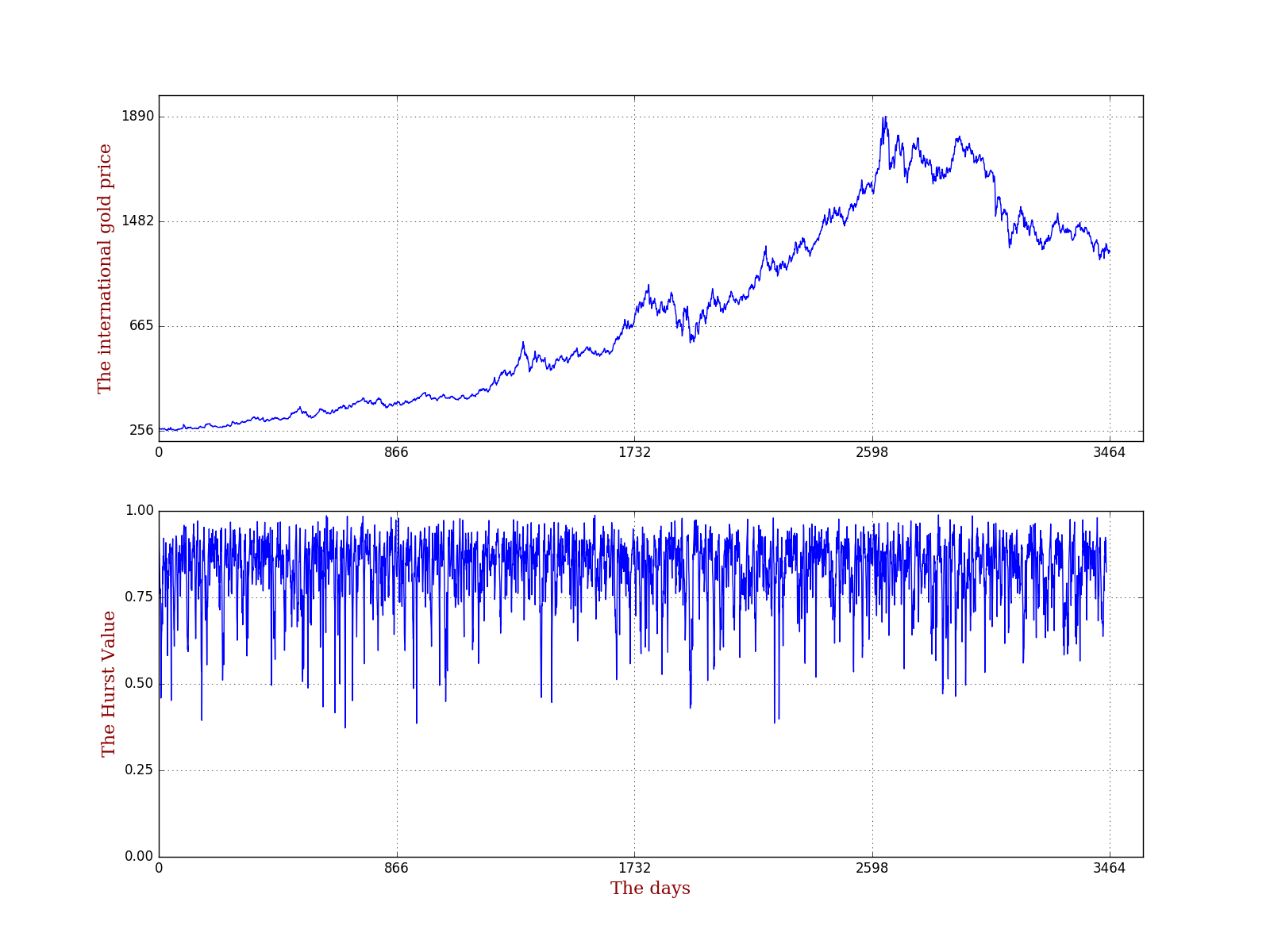


图2‑ 实际国际金价走势

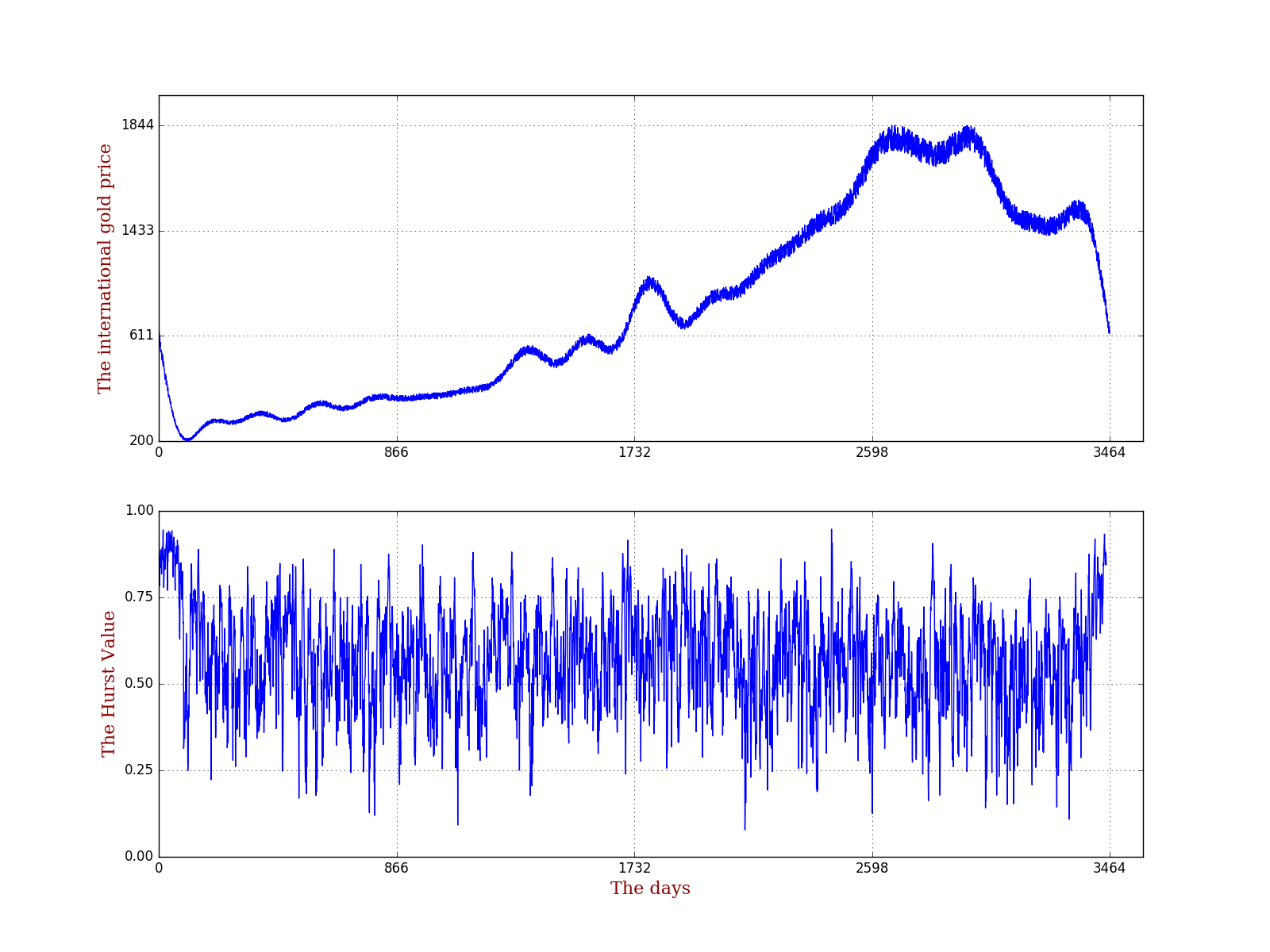


图2‑ 随机模拟金价走势（噪声比3%）

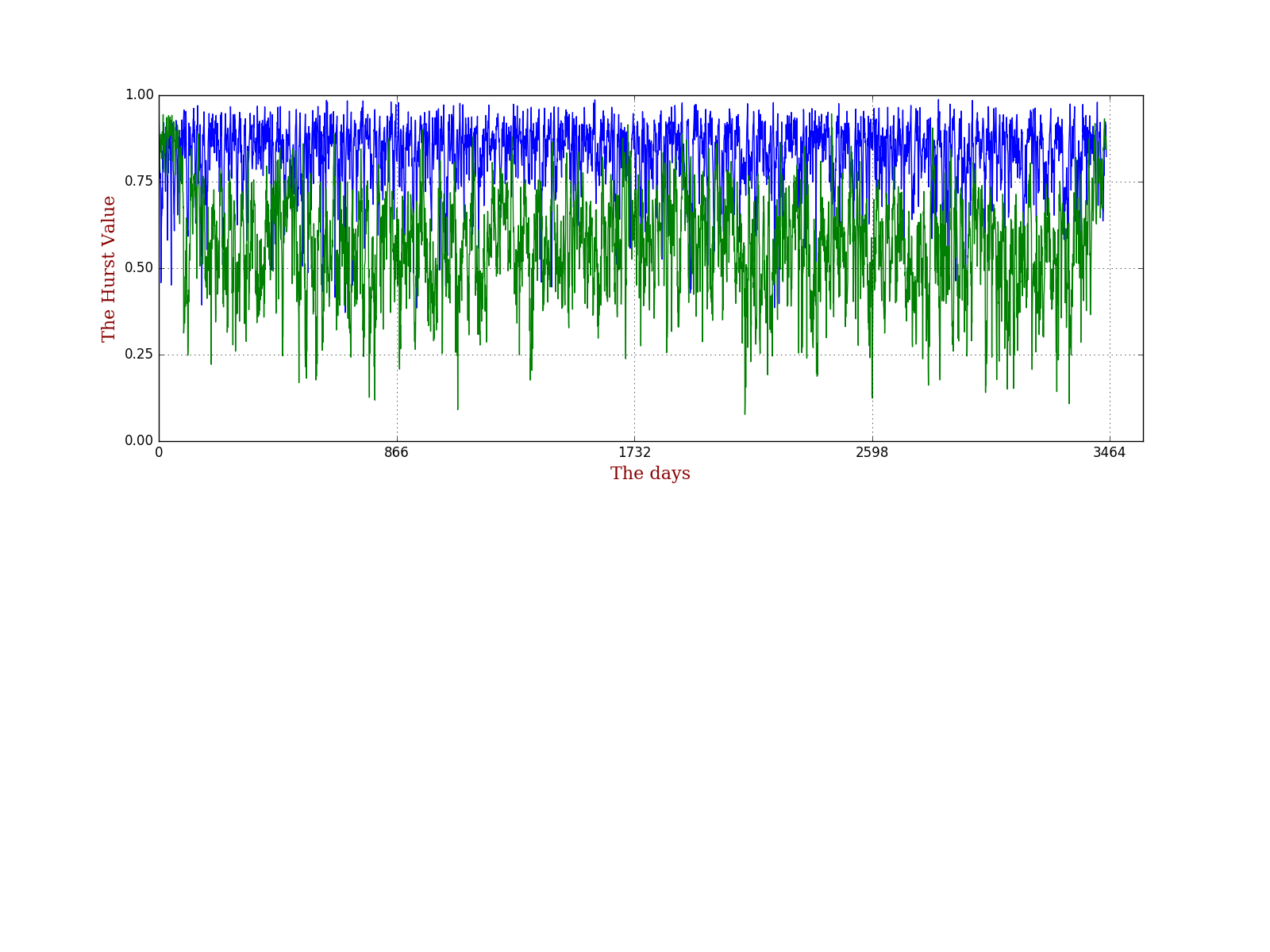


图2‑ 实际金价走势（蓝）与随机模拟金价走势（绿）Hurst指数对比

从图2-3可以看出，实际金价的Hurst指数（蓝色）与随机模拟金价的Hurst指数（绿色）之间有着非常显著的差异。实际金价的Hurst指数平均值为0.81，而随机游走金价的Hurst指数平均值为0.51。这个实验结果符合随机游走时间序列的Hurst指数接近于0.5的结论，同时也说明实际金价在14日（10交易日）内是有偏的随机游走时间序列（即具有一定自相关能力的时间序列）。

* 1. 本章小结

在本章中，分形市场假说表明每个人对于公开的市场信息的获取和理解程度都不相同，每个人对于市场信息的分析处理能力也不同。因此也就有可能利用超过普通投资者的市场分析能力的方法，做出更加理性的投资决策。我们在本章R/S分析结果也表明国际黄金价格是有一个有偏的时间序列，如果能够找到分析或预测这个“偏向”的方法，就能够在很大程度上提高该研究对于市场信息的处理能力。所以，可以认为国际黄金价格在一定程度上是可以被预测的，接下来本文将解释，为何选用深度残差网络与专家系统相结合的方法预测国际金价。

1. 深度残差网络
   1. 深度残差网络概述
      1. 深度学习简介

在20世纪80年代，深度学习被定义为一类利用多层由非线性单元组成的处理层，进行数据特征的提取与变换的算法。这些处理层都是连续的，并且相邻的每一层都会利用前一层的输出结果作为本层的输入。

相对于浅层学习算法，深度学习算法将输入数据传递了更多的处理层。在每一层当中，传入的信号被这一层当中的神经元进行非线性的变换，这些神经元的参数通过对深度神经网络的训练得出[17]。深度学习最终将利用大量训练得到的深度处理层给高维而抽象的数据（包含一系列复杂的非线性变换）建立模型[18]。

* + 1. 深度学习的基本思想

假设有一个系统，总共有个处理层。这个系统的输入输出分别为和，可以将这个过程表示为图3-1。如果输出的信息量等于输入的信息量，则表示输入经过系统的各个处理层后并无信息丢失（一般很难实现）。



图3‑ 深度学习数据处理流程

经过系统的处理过程，不仅不会增加信息量，反而会造成一些信息的丢失。假设系统能够保持信息不变，即输入信息与输出信息在它的任意处理层前后都没有差别，便可以说任何一层输出的信息都是输入信息的另外一种表达形式。

假设一个系统（总共有层，分别为）有输入，例如二维图像点集或者语音信号时间序列。通过对神经网络的训练（调整系统中参数），保持它的输出仍然是输入，那么通过训练获得输入的一系列层次特征，即输入数据另外的一类表达方式。

深度学习的基本思想就是构建出训练层的结构，并通过训练得出输入数据的表达层。通过这种方式实现对输入信息进行分级表达，进而利用这些分级表达层来进行数据的分类、处理和预测。

* + 1. 深度学习的原理

若*n*是输入数据的个数，假设这个输入为，则经过一个确定的映射之后得到这个输入数据的隐层表达式3-1：

|  |  |  |
| --- | --- | --- |
|  |  | (3-1) |

其中，s是一个非线性函数（例如Sigmoid函数），是的权重矩阵，是一个长度为的向量。

然后我们假设是的逆映射，则其表达式3-2为：

|  |  |  |
| --- | --- | --- |
|  |  | (3-2) |

其中，是此映射的逆权重矩阵，是一个长度为的向量。过程可以由图3-2表示，



图3‑ 自编码器值传递过程

在这个模型中，可以把看做为输入的重构向量，目标便是优化得出平均重构误差最小的映射。为达到这个目的，传统的计算方法为计算与的差的方差，如式3-3所示：

|  |  |  |
| --- | --- | --- |
|  |  | (3-3) |

如果输入非位向量的数据，便可以计算与的相互熵，如式3-4所示：

|  |  |  |
| --- | --- | --- |
|  |  | (3-4) |

通过对深度网络的各处理层参数的调整，优化出最小的，这样便可以把当作为的一个低损耗估计。由于对于不同的值，映射的损耗值都不一定相同。所以需要经过大量监督式的训练，优化深度网络的参数，最终得到自编码器。

* + 1. 深度残差网络的原理

为解决深度神经网络训练过程中精度下降的问题，本文引入一个深度残差网络的架构。在这个架构中，最终希望的是每一网络层的输出拟合输入数据的残差，而不是像其他深度神经网络一样拟合原来的输入数据[2]。

假设深度神经网络的基础映射为，然后构造一个映射如式3-5：

|  |  |  |
| --- | --- | --- |
|  |  | (3-5) |



图3‑ 深度残差网络值传递过程

因此，原来的映射就变成，深度残差网络的值传递过程如图3-3所示。尽管这两个映射的值都可以近似为理论真值，但是它们的训练难度是不一样的。研究证明深度残差网络训练过程中的优化问题变得非常简单，因为每一层的输出结果直接收敛到0[2]。

* 1. 国际金价数据的获取与预处理
     1. 国际金价走势历史数据的获取

Forextester是美国的一家外汇交易网站（网址见附录A），免费提供近15年的国际黄金价格XAU/USD（黄金/美元）的历史数据。通过此网站下载到的数据以纯文本格式储存，其数据的具体存储形式如表3-1所示。

表3‑ 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 |
| XAUUSD | 20010102 | 234100 | 268.8 | 268.8 | 268.8 | 268.8 | 4 |
| XAUUSD | 20010102 | 235000 | 268.8 | 268.8 | 268.8 | 268.8 | 4 |
| XAUUSD | 20010102 | 235200 | 268.6 | 268.6 | 268.6 | 268.6 | 4 |
| XAUUSD | 20010103 | 300 | 268.9 | 268.9 | 268.9 | 268.9 | 4 |
| XAUUSD | 20010103 | 2600 | 269.1 | 269.1 | 269.1 | 269.1 | 4 |

* + 1. 国际金价历史数据的预处理方法

对所有数据的读取可用简单的程序实现，并将时间转换为更实用的UTC格式，经过简单的处理后其数据具体的存储形式如表3-2所示。

表3‑ 时间格式转换后的数据（部分）

|  |  |  |
| --- | --- | --- |
| <DATE> | <UTC> | <CLOSE> |
| 20010102230900 | 978448140 | 268.8 |
| 20010102231200 | 978448320 | 268.8 |
| 20010102231600 | 978448560 | 268.9 |
| 20010102233800 | 978449880 | 268.8 |
| 20010102234100 | 978450060 | 268.8 |
| 20010102235000 | 978450600 | 268.8 |
| 20010102235200 | 978450720 | 268.6 |
| 20010103000300 | 978451380 | 268.9 |
| 20010103002600 | 978452760 | 269.1 |

在经过简单的数据格式调整之后，本研究希望构造用于深度残差网络训练以及测试的输入、输出数据集。在第二章已经表明，国际金价时间序列在14日（10交易日）内具有一定的自相关性。经过大量测试后，确定最合适的训练数据输入形式如下表3-3所示：

表3‑ 输入数据集格式

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 第一列 | 第二列 | 第三列 | 第四列 | 第五列 | 第六列 | 第七列 |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

其中,

代表第天对应的数据；

表示当天价格的平均值；

表示当天价格的最大值；

表示当天价格的最小值；

表示当天价格平均值的10日平均对数收益率（由下面公式4-1计算）

(4-1)

其中，表示第至第日的价格平均值；

表示当天价格最大值的10日平均对数收益率；

表示当天价格最小值的10日平均对数收益率。

最后，数据输出为第日的对数收益率。数据完成预处理之后，将其分为两部分。第一部分提供给深度残差网络用以训练，第二部分用于深度残差网络的预测误差测试。利用Python的pickle库将所有训练数据集和测试数据集封装成数据包。国际金价历史数据的预处理程序使用方法见附录A。

* 1. 深度残差网络的构建
     1. 系统环境设置

出于方便、效率的考虑，深度残差网络构建在Linux系统上，并同时调用NVDIA（英伟达公司）的CUDA库，利用GPU给深度残差网络的训练加速。具体的搭建环境如表3-4所示：

表3‑ 深度残差网络搭建环境表

|  |  |
| --- | --- |
| 系统版本 | Linux Ubuntu 14.04 LTS 64-bit |
| 内存 | 7.8 GiB |
| 硬盘 | 483.7 GB |
| CPU | Intel Core i5-2400 @ 3.10GHz \* 4 |
| GPU | NVDIA GeForce GTX 950/PCIe/SSE2 |
| 编程语言 | Python |

注：CPU（Central Processing Unit，中央处理器）。

* + 1. 构建适用于国际金价预测的深度残差网络

基于Linux系统构建深度残差网络时，利用Python编程语言主要调用“numpy”、“theano”、“lasagne”等开源库，建立81层总共包含2,245,005个参数的深度残差网络。其输入层形式为的矩阵，输出层形式为浮点数。经过本章的预处理方式得到用于其进行训练以及测试的数据包。

然后通过GitHub下载获得建立深度残差网络的基本Python源代码框架，并做如下修改：

1. 替换输入输出数据集的获取方式、输出数据集的形式由单个类别变量修改为5个浮点变量；
2. 修改输出层的节点个数为5个、训练以及测试误差的计算方式为预测值与实际值之差的方差、训练过程中更新参数的方式。
3. 增加自动记录测试准确率最高的参数功能、增加每100个训练周期自动备份参数功能、增加每个训练周期自动记录训练及测试准确率的功能、增加加载参数的功能。
   * 1. 利用预处理的数据集对深度残差网络进行训练

由本章得到的“Inputs.pkl”和“Outputs.pkl”两个以特定形式封装的包含4764日金价走势特征的数据集。利用程序“Training.py”将前1至4514日数据作为训练集，最后250日金价特征数据作为测试集。搭建的神经网络总共81层共2,245,005个参数，其训练的学习率按照“0.01”、“0.001”的方式降低，同时用于训练的数据集的也不断调整。具体的训练方式见附录A，最后得到训练成型的深度残差网络（总训练次数为162,000次）。

* + 1. 深度残差网络的训练过程

在深度残差网络的训练过程中定义：若某一交易日的国际金价大于前一日的国际金价，则称该日价格上涨，否则称为下跌。若某一交易日的国际金价相对于前一日的国际金价上涨或者下跌的幅度超过后一日金价的1%，则规定为大涨跌幅，否则将其规定为小涨跌幅。

同时规定：

* + 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。

根据以上规定，深度残差网络在实际训练过程当中的训练集的预测准确率、测试集的预测准确率以及测试集的涨跌幅预测准确率随着训练次数增加的变化曲线如图3-4所示，训练集的预测方差与测试集的预测方差如图3-5所示。

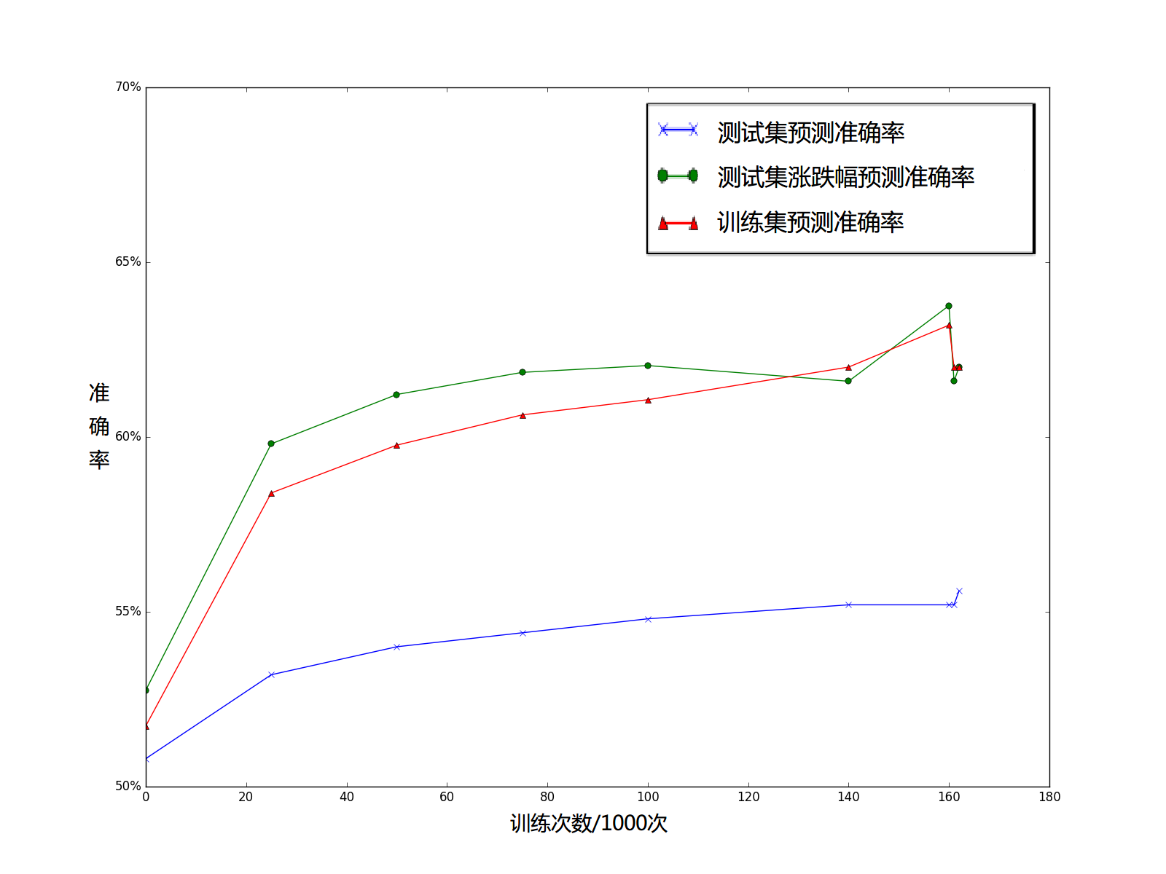


图3‑ 深度残差网络训练过程中的准确率变化曲线

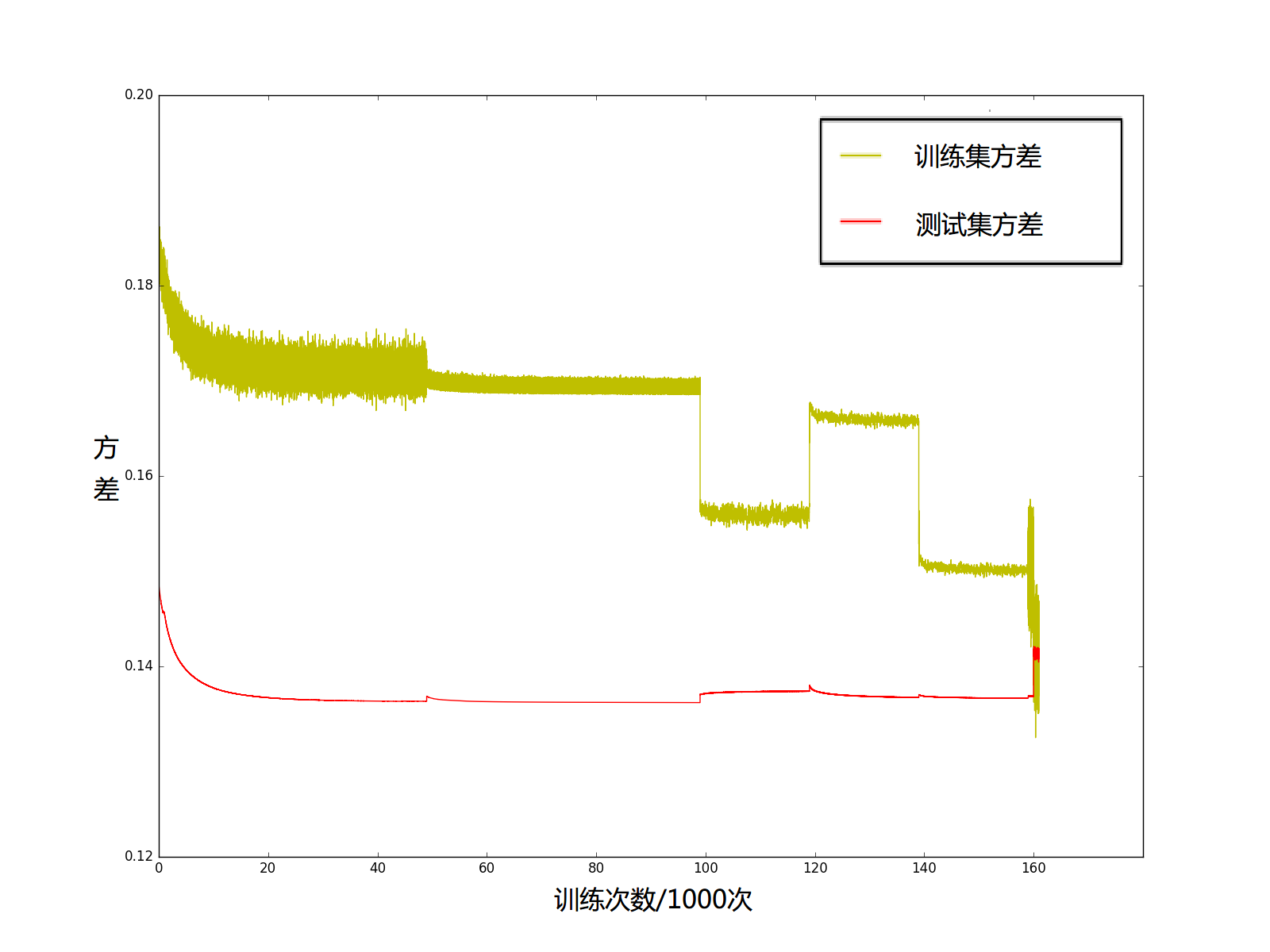


图3‑ 深度残差网络训练过程中的方差变化曲线

从图3-4中可以看出，随着对于深度残差网络的训练量的增加，深度残差网络对于国际金价的预测准确率逐渐增加。经过160,000次训练后，训练集的预测准确率在之间波动，测试集的预测准确率在之间波动，测试集涨跌幅的预测准确率在之间波动。

* 1. 预测结果与分析

利用深度残差网络和历史数据预测2016年4月25日至2016年5月27日共25交易日的金价走势，结果如图3-4所示，

图3‑ 深度残差网络训练过程中的方差变化曲线

从图中可以发现，预测值与实际值最大误差为6.1%（5月2日），但在第23交易日（5月25日）预测价格与实际价格几乎重合。分析这一时间段的财经新闻发现：美国联邦储备理事会在4月28日发表申明维持利率不变，导致在这个时间点前后XAU/USD的汇率大涨。由于深度残差网络无法完全获取外界的信息，仍然按照国际金价内在的联系进行预测，使得在本时间段第23交易日之前的预测结果与实际金价的相差甚远。在这一金融事件发生2周后，国际金融市场重新趋向于稳定，国际金价重新回归市场期望值，与深度残差网络的预测值重合。

从本章可以看出，虽然经过训练的深度残差网络能够学习到国际金价一定的内在联系，但是其预测准确率仅达到55%。尤其是深度残差网络无法获取足够的外界信息，几乎没有任何应对突发事件的能力。所以下一章就需要利用专家系统为此深度学习网络的预测结果提供一定的鲁棒性。

1. 专家系统
   1. 专家系统
      1. 专家系统简介

专家系统是一个仿人脑进行决策的计算机系统[19]，用于解决暂时无法通过简单的程序代码解决的复杂推理难题。第一个专家系统是在20世纪70年代创造的，而在20世纪80年代专家系统的数量激增，飞速发展以满足各个研究领域的需求。

一般来说，一个专家系统可以分为两个子系统：推理系统和知识库系统。知识库系统存储了由管理员提供或其自动收集到的信息，代表着已经收集存储的事实和规则。而推理系统将其设定好的或通过机器学习系统学习到的规则应用于知识库中存储的数据，从而推导出新的数据。其推理引擎同时还可以包括解释以及调试的功能。

* + 1. 专家系统的搭建环境

在本文的研究中，专家系统利用数据挖掘技术获取互联网当中与国际黄金市场相关的新闻信息，并从众多的信息中筛选出各金融专家或组织对于国际金价预测的信息，以修正深度残差网络的预测结果。此专家系统在Linux上使用Python编程语言实现，其过程自动执行，不需要人工干预。具体的硬件环境如第三章表3-4所示。

* + 1. 专家系统在国际金价预测研究中的意义

国际黄金价格是一个有偏的时间序列，因此能够利用深度残差网络学习到其内在的规律。可在国际经济全球化的今天，国际金价也是一个极易受到国际各大经济、政治事件影响的时间序列，各类突发事件都会严重改变国际金价短期或长期的走势。就目前来说，本文设计的深度残差网络适用的假设条件是：国际金价时间序列只在14日内产生相互影响，而在此期间国际金融市场环境以及政治背景没有发生变化。因此其预测结果在现实情况下将会产生非常大的误差。

专家系统能够实时获取国际金融专家和组织对于国际金价市场的市场预期，并通过推理库提取信息当中的关键词，最终利用数据化的结果修正深度残差网络的预测结果。这将大大提高深度残差网络的鲁棒性，弥补它在实时信息获取方面的不足。

* 1. 专家系统的搭建
     1. 专家系统数据来源

由于现在大多数门户网站使用了异步加载技术或异常请求检测程序，无法靠普通的爬虫软件获取数据或搜索次数过多会被目标网站禁止访问，所以暂时使用“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（网址见附录A）。

* + 1. 爬虫软件设计

爬虫软件的实现思路为，从选定的门户网站上搜索与国际黄金有关的最新信息，然后根据其中的关键词出现次数的统计与分析结果，筛选出包含金融专家或金融组织预测的相关信息并备份至系统中。最后将所有预测信息数据化，并综合分析计算得到一个用于修正深度残差网络预测结果的参数。

首先分析今日头条门户网站，让其对关键词“国际”、“黄金”进行搜索，按照时间顺序得到最近的50条新闻搜索结果（网址代码为“*count=50*”），并且同时发现搜索三个关键词“国际”、“黄金”、“预测”的结果与前面的搜索结果有很大的差别，因此可以得到两个搜索链接（见附录A）。

通过分析它们的网站源代码，找到了其满足Python中re模块的正则表达式用以得到各条新闻的链接：

item\_source\_url": "(.\*?)",

关键代码：

url = re.findall('item\_source\_url": "(.\*?)",', links.text, re.S)

其中*links*代表两个搜索链接，*url*返回值为各条新闻的链接。

对新闻内容的分析，可以发现涉及金融专家或投资机构对于国际金价预测信息的内容一般包含“预计”、“建议”、“迹象”、“阻力”、“支撑”、“做多”、“目标”等关键词，而涉及国际金价涨跌预测信息的内容一般包含“继续”、“保持”、“放缓”、“承压”、“走弱”、“走强”。通过对信息的筛选提取，将这些信息备份至专家系统知识库。

* + 1. 专家系统的数据提取实例

通过本章对专家系统数据库（信息存储格式见附录B）的分析，可以发现其中既有无用信息，也有包含金融机构的预测信息。通过大量的数据比对发现知识库数据中，预测信息的表现形式一般为“建议”、“做多”、“目标”、“阻力”、“支撑”。

再通过推理系统对于如表4-1中所示关键词的数据分析提取。其中关键词1表示提取的数据中可能包含有国际金价的预测价位的数值信息，关键词2和3分别金融机构或专家对于国际金价的走势判断为乐观（国际金价短期或长期内会有上涨）和悲观（国际金价短期或长期内会有下跌）态度。

汇总两个网站的备份知识库数据，便得到了如表4-2所示专家系统推理数据库。其中第一项为推理数据库的数据日期，后两项分别为预测黄金价格上涨和下跌关键词出现的次数，而表格中剩余的项为金融专家或机构预测的短期或长期的黄金目标。

表4‑ 专家系统关键词提取表

|  |  |  |
| --- | --- | --- |
| **关键词1** | **关键词2** | **关键词3** |
| 黄金…正试图…上破… | 预计黄金市场将继续走高 | 黄金的季节性放缓 |
| 黄金…有可能突破… | 黄金在近期内将走弱 | 黄金…市场…承压 |
| 黄金…目标位是… | 黄金…我们仍看多 | 黄金…继续走强 |
| 金价…会回调至… | 持有黄金是明智的 | 减…黄金仓位 |
| 金价…会跌至… | 黄金…继续走强 | 金价…下跌 |
| 今年有望涨至… | 黄金…保持看多 | 金价…走弱 |
| 黄金…会上探… | 建议做多黄金 | 金价…暴跌 |
| 黄金…预测价… | 增…黄金仓位 | 美元走强 |
| 建议…做多 | 美元将下跌 | 支撑美元 |

表4‑ 专家系统推理库5月18日信息表

|  |  |  |  |
| --- | --- | --- | --- |
| 信息获取日期：2015/5/18 | | | |
| 预测金价上涨关键词出现次数：69 | | 预测金价下跌关键词出现次数：100 | |
| 预测价格关键值 | | | |
| 1272 | 1270 | | 1270 |
| 1270 | 1270 | | 1270 |
| 1268 | 1272 | | 1270 |
| 1270 | 1270 | | 1270 |
| 1270 | 1270 | | 1265 |

* 1. 预测结果与分析

本章同样研究2016年4月25日至2016年5月27日共25交易日的金价走势，数据见表4-3。发现专家系统推理库在第5交易日（4月29日）预测到，美国联邦储备理事会在4月28日至4月29日发表申明维持利率不变的决定给国际金融市场包括黄金市场所带来的影响，如图4-1所示。

因此专家系统推理库的结果会给深度残差网络的结果给与一定的修正，根据最大后验概率原理，将修正的权重系数由式4-1定义：

|  |  |  |
| --- | --- | --- |
|  |  | (4-1) |

其中，分别表示专家系统知识库中，预测金价上涨和下跌的关键词的出现次数。当时，相信金价上涨和下跌的信息量一样，说明市场中并没有表现出特别强的趋势性，深度残差网络的预测值更为可信。当相差很大时表示市场趋势性强，专家系统推理库的预测值更为可信。于是得到修正后的金价，如式4-2所示：

|  |  |  |
| --- | --- | --- |
|  |  | (4-2) |

其中表示专家系统推理库的预测值，表示深度残差网络的预测值。通过专家系统推理库数据的加权修正，可以得到图4-2。

表4‑ 2016年4月25日至2016年5月27日总信息

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 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 |
| 20160425 | 1235.5 | 1201.8 | 1231 | 1 | 1 | 1201.8 |
| 20160426 | 1238.3 | 1205.2 | 1236 | 1 | 6 | 1229.3 |
| 20160427 | 1245.7 | 1202.8 | 1238 | 0 | 1 | 1220.4 |
| 20160428 | 1256.2 | 1204.7 | 1249 | 1 | 0 | 1226.8 |
| 20160429 | 1279.6 | 1209.1 | 1315 | 0 | 1 | 1262.1 |
| 20160502 | 1293.9 | 1214.9 | 1325 | 8 | 2 | 1301.1 |
| 20160503 | 1289.9 | 1219.2 | 1340 | 5 | 2 | 1287.1 |
| 20160504 | 1280.3 | 1222.4 | 1328 | 21 | 48 | 1318.8 |
| 20160505 | 1278.0 | 1225.1 | 1300 | 12 | 18 | 1265.9 |
| 20160506 | 1284.6 | 1231.1 | 1290 | 8 | 63 | 1288.6 |
| 20160509 | 1269.9 | 1240.9 | 1285 | 33 | 9 | 1282.0 |
| 20160510 | 1265.2 | 1241.1 | 1270 | 6 | 4 | 1249.3 |
| 20160511 | 1274.5 | 1241.9 | 1270 | 2 | 1 | 1248.9 |
| 20160512 | 1269.5 | 1242.0 | 1285 | 4 | 11 | 1274.9 |
| 20160513 | 1272.0 | 1239.2 | 1260 | 2 | 1 | 1244.4 |
| 20160516 | 1277.5 | 1238.7 | 1282 | 24 | 50 | 1277.7 |
| 20160517 | 1275.8 | 1236.0 | 1285 | 12 | 1 | 1280.2 |
| 20160518 | 1271.9 | 1236.2 | 1265 | 9 | 2 | 1259.7 |
| 20160519 | 1254.2 | 1234.9 | 1270 | 83 | 105 | 1260.1 |
| 20160520 | 1254.3 | 1232.3 | 1244 | 1 | 28 | 1243.5 |
| 20160523 | 1250.5 | 1230.1 | 1252 | 33 | 181 | 1251.7 |
| 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 |
| 20160527 | 1217.1 | 1215.8 | 1226 | 62 | 135 | 1226.0 |

图4‑ 专家推理库预测值与深度残差网络预测值对比

图4‑ 加权修正预测值与深度残差网络预测值对比

专家系统的推理库预测值虽然相对实际金价也有较大幅的波动，但能在一定程度上增加深度残差网络的鲁棒性，弥补它对于市场信息的获取的不足。

在本小节的25交易日时间段内，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。

1. 结论
   1. 研究内容概述

本文以预测国际金价走势为目的，首先分析国内外利用深度学习进行国际金价预测的研究现状，发现随着深度学习网络深度的提高，会出现精度下降的问题。深度残差网络是深度学习算法2015年的最新研究成果，更“深”的神经网络层数也意味着对国际金价数据更“深层次”的理解，在金融方面具有一定的研究价值。然后研究证明国际金价在10交易日内是有偏向的时间序列，即是可以被预测的。接下来研究搭建深度残差网络，完成国际金价历史数据的获取、预处理以及对深度残差网络的训练。对2016年4月25日至2016年5月27日共25交易日的金价走势预测，发现深度残差网络的一个弊端：无法及时对国际金融市场的变化做出预测调整。然后利用Python语言编写专家系统，实时获取金融专家和组织的预测信息，并经过反复测试得到专家系统修正权重系数的表达式，完成对深度残差网络预测结果的修正，大幅度降低其均方误差值。

* 1. 研究过程的不足之处

在研究深度残差网络的搭建过程中，曾尝试利用前十分钟的金价预测后十分钟的金价走势，却发现国际金价的训练集基本无法提高精度。后来发现在分钟级的时间段内，国际金价的Hurst指数均值接近0.5，即说明金价在分钟级的时间段内是不可预测的。由于完成搭建专家系统的时间比较晚，导致早期的专家系统数据获取不全（2016年5月23日之前），故在此之前的专家系统修正权重系数波动很大，也同时对修正的预测数据产生一定的影响。

* 1. 继续研究的方向

深度残差网络确实能够学习到一定的国际金价市场的内在规律，但是当今国际金融市场波动性与日俱增，同时也会受到众多国际金融事件的干扰，所以预测精度很难再得到提高。这方面的研究不建议继续进行。

从最后一章可以看出，专家系统的预测准确性极高，值得进一步的研究。今后可以尝试从增加专家系统搜索关键词的数量（增加其它与黄金相关的关键词，如白银、石油等）、增加门户网站的搜索数量（编写高级爬虫软件获取更多搜索引擎的资源）等方面，提高专家系统知识库的覆盖面。也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，对大数据具有更加充分的理解，从而提高专家系统推理库的预测能力。

参考文献

1. Krizhevsky A, Sutskever I, et al. ImageNet Classification with Deep Convolutional Neural Networks [J]. Advances in Neural Information Processing Systems, 2012, 25(2):2012.
2. He Kaiming, Zhang Xiangyu, et al. Deep Residual Learning for Image Recognition [R]. Computer Sciences, 2015, arXiv:1512.03385.
3. Hochreiter, Sepp; and Schmidhuber, et al. Long Short-Term Memory [J]. Neural Computation, 1997, 9(8):1735-1780.
4. Alex G, Marcus L, et al. A novel connectionist system for unconstrained handwriting recognition [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2009, 31(5):855-868.
5. D. C. Ciresan. Deep Big Simple Neural Nets for Handwritten Digit Recognition [J]. Neural Computation, 2010, 22, pp. 3207-3220.
6. Sak H., Senior A. et al. Fast and Accurate Recurrent Neural Network Acoustic Models for Speech Recognition [J]. Machine Learning, 2015, arXiv:1507.06947.
7. Giusti A, Ciresan D.C et al. Fast Image Scanning with Deep Max-Pooling Convolutional Neural Networks [J]. Computer Sciences, 2013:4034-4038.
8. Bengio.Y, Simard.P, et al. Learning long-term dependencies with gradient descent is difficult [J]. IEEE Transactions on Neural Networks, 1994, 5(2):157-66.
9. He.K, Sun J. Convolutional neural networks at constrained time cost [C]// Computer Vision and Pattern Recognition. IEEE, 2014.
10. Ramanathan R. Introductory Econometrics with Applications [M]// Introductory econometrics with applications. Dryden Press, Harcourt Brace College Publishers, 1998.
11. Philip A.A, Taofiki A.A, et al. Artificial Neural Network Model for Forecasting Foreign Exchange Rate [J]. World of Computer Science & Information Technology Journal, 2011, 1(3):2221-741110.
12. Hurst H.E. Long Term Storage Capacity of Reservoirs [J]. Transactions of the American Society of Civil Engineers, 1951, 116(12):776-808.
13. Mandelbrot B.B, Wallis J.R. Robustness of the rescaled range R/S in the measurement of noncyclic long run statistical dependence [J]. Water Resources Research, 1969, 5(5):967–988.
14. Fama E.F. The Behavior of StockMarket Prices [J]. Journal of Business, 1965, 38(1):34-105.
15. Fama E.F. Fama, E. Efficient Capital Markets: A Review of Theory and Empirical Work [J]. Journal of Finance, 1970, 25(2):383-417.
16. Peters E.E. Fractal market analysis: applying chaos theory to investment and economics [J]. Chaos Theory, 1994, 34(2):343-345.
17. Dechter R. Learning While Searching in Constraint Satisfaction Problems.[C]// National Conference on Artificial Intelligence. Philadelphia, Pa, August 11-15, 1986. Volume 1: Science. 1986:557-581.
18. Deng.L, Yu.D. Deep learning: methods and applications [J]. Foundations & Trends in Signal Processing, 2013, 7(3):197-387.
19. Myers.W. Introduction to Expert Systems [J]. IEEE Expert, 1986, 1(1):100-109.

致 谢

论文工作是在张鸿燕老师的悉心指导下完成的。张老师以他敏锐的洞察力、渊博的知识、严谨的治学态度、精益求精的工作作风和对科学的献身精神给我留下了刻骨铭心的印象，这些使我受益匪浅。

同时也感谢王子昊学长、罗家祯学长和周璐莎学姐、康鸣翠学姐在本课题研究当中给予的帮助。

附录A：网址链接

A.1 历史数据的获取地址

Forextester外汇交易网站网址：*http://www.forextester.com/data/datasources*

A.2 专家系统的资源地址

今日头条：*http://toutiao.com/*

今日头条两个搜索链接：

*http://toutiao.com/search\_content/?offset=0&format=json&keyword=%E5%9B%BD%E9%99%85+%E9%BB%84%E9%87%91&autoload=true&count=50&\_=1460862899094*

*http://toutiao.com/search\_content/?offset=0&format=json&keyword=%E5%9B%BD%E9%99%85+%E9%BB%84%E9%87%91+%E9%A2%84%E6%B5%8B&autoload=true&count=50&\_=1460864327503*

附录B：程序调用过程及专家系统知识库部分信息

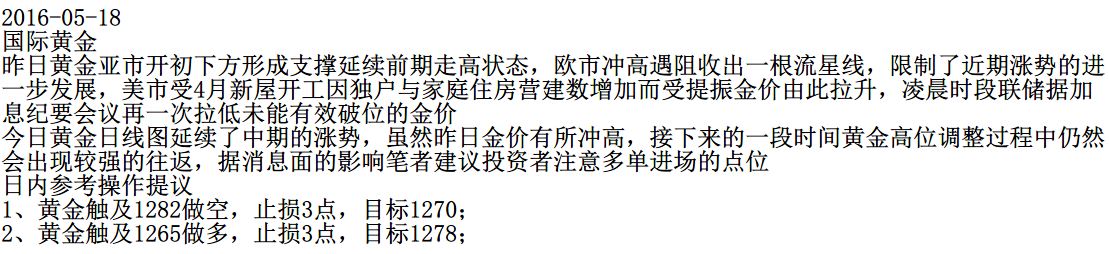
B.1 国际金价历史数据的预处理程序

研究通过3.2.1 提供的途径下载从2001年1月2日至2016年4月29日总共4764交易日的国际黄金价格/美元(XAU/USD)的数据，其中总共包括3,583,882个时刻对应的国际黄金价格，以“XAUUSD.txt”的文件名保存。

然后通过Python语言程序“Predata.py”中的“file2data(filename)”函数将逐行读取金价历史数据并返回两个长度为3,583,882的日期与金价列表。并调用“MakePriceEachDay(Date, Price)”函数便可将每日对应的金价以“PriceEachDay.pkl”形式保存至同一目录。接下来使用“Save\_HMEachDay()”函数便可读取刚刚存储的每日金价数据并生成0中所需的“Mean”、“Max”、“Min”以及当日金价的对数、十日平均金价的对数数据并将所有数据保存到同一目录。最后调用“Pre\_datas()”函数读取上述的数据并生成可直接提供给深度残差网络进行训练和测试的数据集，以“Inputs.pkl”和“Outputs.pkl”的形式保存至同一目录。

B.2 利用预处理的数据集对深度残差网络进行训练

搭建的神经网络总共81层共2,245,005个参数，其训练的学习率根据“0.01”、“0.001”，每训练50,000次调整一次。

由于相隔时间越长，国际金价时间序列的相关性越小，因此在以上训练的基础上分别选用第1000日至4514日，第2000日至4514日以及第3000日至4514日数据集，以“0.001”的学习率分别对此神经网络再次训练20,000次；第4000日至4514日的数据集，以“0.001”的学习率对此神经网络再次训练1,000次；第4314日至4514日数据集，以“0.001”的学习率分别对上一个训练的结果再次训练1000次，最后得到一个已经训练成型的深度残差网络（总训练次数为162,000次）。

图B.1 专家系统知识库5月18日部分数据

附录C：外文翻译资料

(Article Source：IASTED conference on Financial Engineering and Applications (FEA 2004). 2004: 203-209.)

HURST EXPONENT AND FINANCIAL MARKET PREDICTABILITY

Bo Qian Khaled Rasheed

Department of Computer Science University of Georgia

Athens, GA 30601 USA

[qian, khaled]@cs.uga.edu

**ABSTRACT**

The Hurst exponent (H) is a statistical measure used to classify time series. H=0.5 indicates a random series while H>0.5 indicates a trend reinforcing series. The larger the H value is, the stronger trend. In this paper we investigate the use of the Hurst exponent to classify series of financial data representing different periods of time. Experiments with back propagation Neural Networks show that series with large Hurst exponent can be predicted more accurately than those series with H value close to 0.50. Thus Hurst exponent provides a measure for predictability.

**KEY WORDS**

Hurst exponent, time series analysis, neural networks, Monte Carlo simulation, forecasting

**Introduction**

The Hurst exponent, proposed by H. E. Hurst [1] for use in fractal analysis [2],[3], has been applied to many research fields. It has recently become popular in the finance community [4],[5],[6] largely due to Peters’ work [7],[8]. The Hurst exponent provides a measure for long- term memory and fractality of a time series. Since it is robust with few assumptions about underlying system, it has broad applicability for time series analysis. The values of the Hurst exponent range between 0 and 1. Based on the Hurst exponent value H, a time series can be classified into three categories. (1) H=0.5 indicates a random series. (2) 0<H<0.5 indicates an anti-persistent series. (3) 0.5<H<1 indicates a persistent series. An anti- persistent series has a characteristic of “mean-reverting”, which means an up value is more likely followed by a down value, and vice versa. The strength of “mean- reverting” increases as H approaches 0.0. A persistent series is trend reinforcing, which means the direction (up or down compared to the last value) of the next value is more likely the same as current value. The strength of trend increases as H approaches 1.0. Most economic and financial time series are persistent with H>0.5.

In time series forecasting, the first question we want to answer is whether the time series under study is predictable. If the time series is random, all methods are expected to fail. We want to identify and study those time series having at least some degree of predictability. We know that a time series with a large Hurst exponent has strong trend, thus it’s natural to believe that such time series are more predictable than those having a Hurst exponent close to 0.5. In this paper we use neural networks to test this hypothesis.

Neural networks are nonparametric universal function approximators [9] that can learn from data without assumptions. Neural network forecasting models have been widely used in financial time series analysis during the last decade [10],[11],[12]. As universal function approximators, neural networks can be used for surrogate predictability. Under the same conditions, a time series with a smaller forecasting error than another is said to be more predictable. We study the Dow-Jones index daily return from Jan. 2, 1930 to May. 14, 2004 and calculate the Hurst exponent of each period of 1024 trading days. We select 30 periods with large Hurst exponents and 30 periods with Hurst exponents close to random series, and then we use these data to train neural networks. We compare forecasting errors for these two groups and find that the errors are significantly different. This research is done using Matlab. All Matlab programs generating result for this paper can be downloaded from www.arches.uga.edu/~qianbo/research.

The remainder of the paper is organised as follows: Section 2 describes the Hurst exponent in detail. Section 3 then describes the monte carlo simulation process we used to generate data with similar structure to the financial series of interest to us. Section 4 describes a scramble test that we conducted to help verify that there is structure in the series due to the order of samples. Section 5 describes neural networks and their use to verify that sequences with larger values of the Hurst exponent can be more accurately learned and predicted than those with lower Hurst exponent values. Finally, the paper is concluded in section 6.

1. **Hurst exponent and R/S analysis**

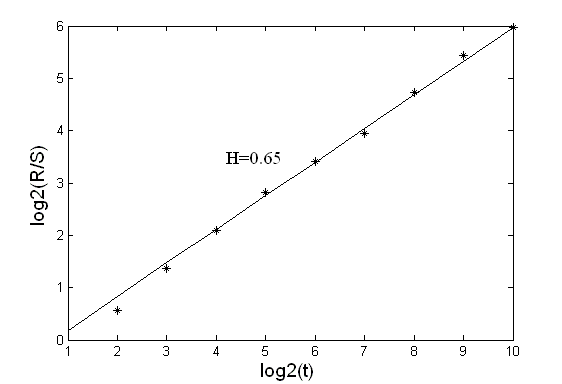
The Hurst exponent can be calculated by rescaled range analysis (R/S analysis). For a time series, X = X1, X2, … Xn, R/S analysis method is as follows:

* 1. Calculate mean value m.
  2. Calculate mean adjusted series Y
  3. Calculate cumulative deviate series Z
  4. Calculate range series R
  5. Calculate standard deviation series S
  6. Calculate rescaled range series (R/S)

Note is averaged over the regions until where m=floor(n/t). In practice, to use all data for calculation, a value of t is chosen that is divisible by n.

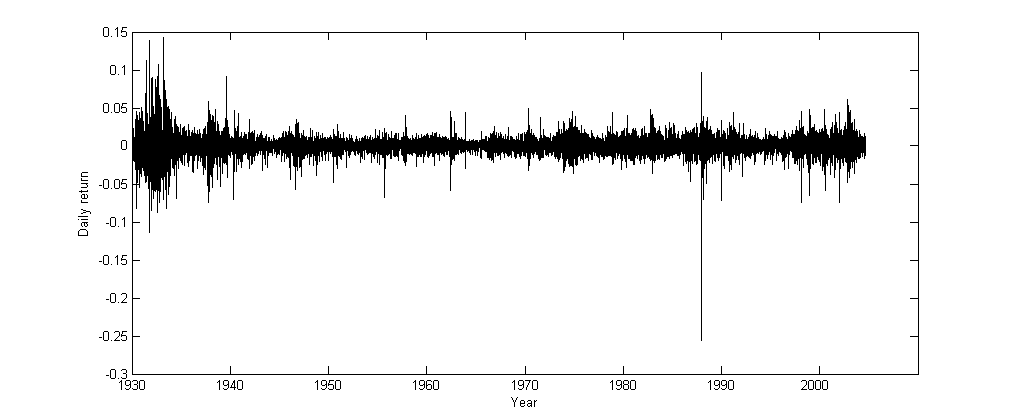
Hurst found that (R/S) scales by power-law as time increases, which indicates

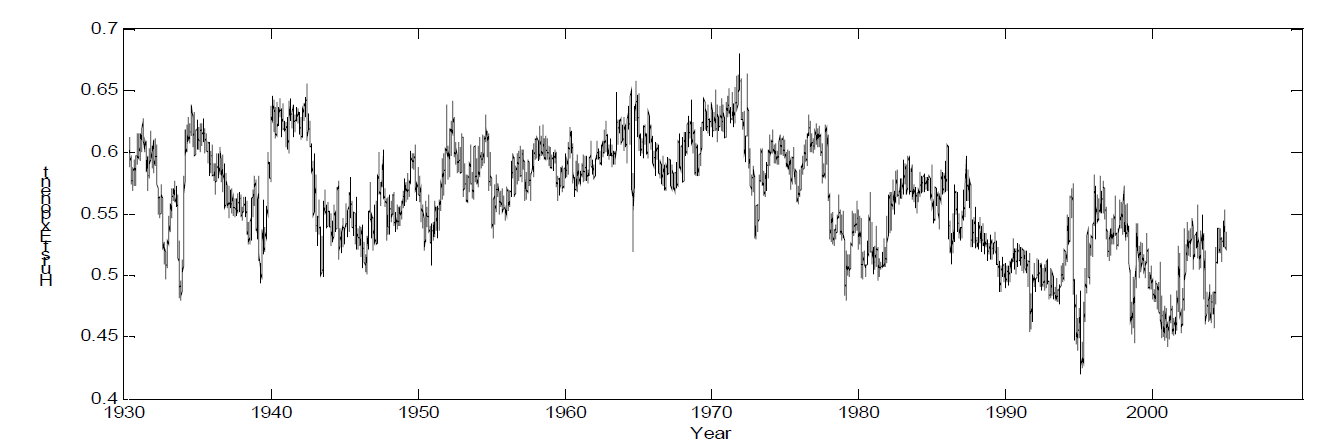
Here c is a constant and H is called the Hurst exponent. To estimate the Hurst exponent, we plot (R/S) versus t in log-log axes. The slope of the regression line approximates the Hurst exponent. For t<10, (R/S)t is not accurate, thus we shall use a region of at least 10 values to calculate rescaled range. Figure 2.1 shows an example of R/S analysis.



**Figure 2.1. R/S analysis for Dow-Jones daily return from 11/18/1969 to 12/6/1973**

In our experiments, we calculated the Hurst exponent for each period of 1024 trading days (about 4 years). We use to do regression. In the financial domain, it is common to use log difference as daily return. This is especially meaningful in R/S analysis since cumulative deviation corresponds to cumulative return. Figure 2.2 shows the Dow-Jones daily return from Jan. 2, 1930 to May 14, 2004. Figure 2.3 shows the corresponding Hurst exponent for this period. In this period, Hurst exponent ranges from 0.4200 to 0.6804. We also want to know what the Hurst exponent would be for a random series in our condition.



**Figure 2.2. Dow-Jones daily return from 1/2/1930 to 5/14/2004**

**Figure 2.3. Hurst exponent for Dow-Jones daily return from 1/2/1930 to 5/14/2004**

1. **Monte Carlo simulation**

For a random series, Feller [13] gave expected (R/S)t formula as 3.1.

|  |  |  |
| --- | --- | --- |
|  |  | (3.1) |

However, this is an asymptotic relationship and is only valid for large t. Anis and Lloyd [14] provided the following formula to overcome the bias calculated from (3.1) for small t:

|  |  |  |
| --- | --- | --- |
|  |  | (3.2) |

For t>300, it is difficult to calculate the gamma function by most computers. Using Sterling’s function, formula (3.2) can be approximated by:

|  |  |  |
| --- | --- | --- |
|  |  | (3.3) |

Peters [8] gave equation (3.4) as a correction for (3.2).

|  |  |  |
| --- | --- | --- |
|  |  | (3.4) |

We calculate the expected (R/S) values for t=24, 25,…,210 and do least squares regression

at significance level α=0.05. Results are shown in table 3.1.

|  |  |  |  |
| --- | --- | --- | --- |
| *log2(t)* | *log2(E(R/S))* | | |
| *Feller* | *Anis* | *Peters* |
| 4 | 0.7001 | 0.6059 | 0.5709 |
| 5 | 0.8506 | 0.7829 | 0.7656 |
| 6 | 1.0011 | 0.9526 | 0.9440 |
| 7 | 1.1517 | 1.1170 | 1.1127 |
| 8 | 1.3022 | 1.2775 | 1.2753 |
| 9 | 1.4527 | 1.4345 | 1.4340 |
| 10 | 1.6032 | 1.5904 | 1.5902 |
| **Regression** | **0.5000** | **0.5436** | **0.5607** |
| **Slope (H)** | 5.5511e-016 | 0.0141 | 0.0246 |

**Table 3.1. Hurst exponent calculation from Feller, Anis and Peters formula**

From table 3.1, we can see that there are some differences between Feller’s, Anis’ and Peters’ formulae. Moreover, their formulae are based on large numbers of data points. In our case, the data is fixed at 1024 points. So what is the Hurst exponent for random series in our case?

Fortunately, we can use Monte Carlo simulation to derive the result. We generate 10,000 Gaussian random series. Each series has 1024 values. We calculate the Hurst exponent for each series and then average them. We expect the average number to approximate the true value. We repeated this process 10 times. Table 3.2 below gives the simulation results.

|  |  |  |
| --- | --- | --- |
|  | *Simulated Hurst Exponent* | *Standard deviation (Std.)* |
| 1 | 0.5456 | 0.0486 |
| 2 | 0.5452 | 0.0487 |
| 3 | 0.5449 | 0.0488 |
| 4 | 0.5454 | 0.0484 |
| 5 | 0.5456 | 0.0488 |
| 6 | 0.5454 | 0.0481 |
| 7 | 0.5454 | 0.0487 |
| 8 | 0.5457 | 0.0483 |
| 9 | 0.5452 | 0.0484 |
| 10 | 0.5459 | 0.0486 |
| **Mean** | **0.5454** | **0.0485** |
| **Std.** | **2.8917e-004** |  |

**Table 3.2. Monte Carlo simulations for Hurst exponent of random series**

From table 3.2, we can see that in our situation, the Hurst exponent calculated from Monte Carlo simulations is 0.5454 with standard deviation 0.0485. Our result is very close to Anis’ formula. Based on the above simulations, with 95% confidence, the Hurst exponent is in the interval 0.5454±1.96\*0.0485, which is between 0.4503 and 0.6405. We choose those periods with Hurst exponent greater than 0.65 and expect those periods to be bearing some structure different from random series. However, since these periods are chosen from a large sample (total 17651 periods), we want to know if there exists true structure in these periods, or just by chance. We run a scramble test for this purpose.

1. **Scramble** **test**

To test if there exists true structure in the periods with Hurst exponent greater than 0.65, we randomly choose 10 samples from those periods. For each sample, we scramble the series and then calculate the Hurst exponent for this scrambled series. The scrambled series has the same distribution as the original sample except that the sequence is random. If there exists some structure in the sequence, after scrambling the structure will be destroyed and the calculated Hurst exponent should be close to that of a random series. In our experiment, we scramble each sample 500 times and then the average Hurst exponent is calculated. The results are shown in table 4.1 below.

|  |  |  |
| --- | --- | --- |
|  | *Hurst exponent after scrambling* | *Standard deviation* |
| 1 | 0.5492 | 0.046 |
| 2 | 0.5450 | 0.047 |
| 3 | 0.5472 | 0.049 |
| 4 | 0.5454 | 0.048 |
| 5 | 0.5470 | 0.048 |
| 6 | 0.5426 | 0.048 |
| 7 | 0.5442 | 0.051 |
| 8 | 0.5487 | 0.048 |
| 9 | 0.5462 | 0.048 |
| 10 | 0.5465 | 0.052 |
| **Mean** | **0.5462** | **0.048** |

**Table 4.1. The average Hurst exponent on 500 scrambling runs**

From table 4.1, we can see that the Hurst exponents after the scrambling of samples are all very close to 0.5454 which is the number from our simulated random series. Given this result, we can conclude that there must exist some structure in those periods making them different from random series and that scrambling destroys the structure. We hope this structure can be exploited for prediction. Neural networks, as universal function approximators, provide a powerful tool to learn the underlying structure. They are especially useful when the underlying rules are unknown. We expect neural networks to discover the structure and thus benefit from it. We use neural network prediction error as a measure of predictability. Below we compare prediction errors of the periods with Hurst exponents greater than 0.65 with those between 0.54 and 0.55.

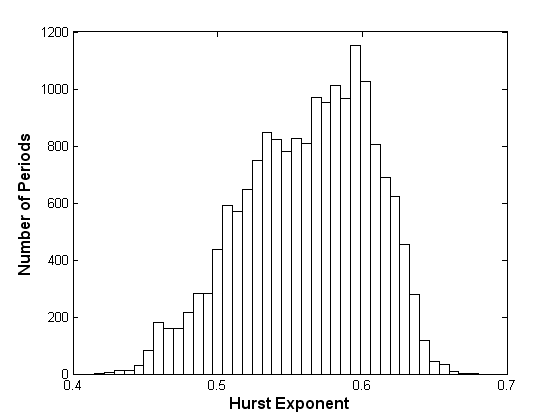
1. **Neural networks**

In 1943, McClloch and Pitts [15] proposed a computational model simulating neurons. This work is generally thought as the beginning of artificial neural networks research. Rosenblatt [16],[17] popularized the concept of perceptrons and created several perceptron learning rules. However, in 1969, Minsky and Papert [18] pointed that perceptrons cannot solve any non “linearly separable” problems. People knew that multi-layer perceptrons (MLP) can simulate non “linearly separable” functions, but no one knew how to train them. Neural network research was then nearly stopped until 1986. In 1986, Rumelhart [19] used the backpropagation algorithm to train MLP and thus resolved this long obsessed problem among connectionists. Since then neural networks have regained considerable interest from many research fields. Neural networks have become popular in the finance society and the research fund for neural network applications from financial institutions is the second largest [20].

A neural network is an interconnected assembly of simple processing nodes. Each node calculates a summation of weighted inputs and then outputs its transfer function value to other nodes. The Feedforward backpropagation network is the most widely used network paradigm. Using the backpropagation training algorithm, neural network adjusts the weights so that it will minimize the square difference (error) between its observed outputs and their target values. The backpropagation algorithm uses a gradient descent method to find a local minimum on the error surface. It calculates the partial derivative of the square error for each weight. The opposite of these partial derivatives (gradient) gives the direction in which the error decreases most. This direction is called the steepest descent direction. The standard backpropagation algorithm adjusts the weights along the steepest descent direction. Although the error in the steepest descent direction is decreased most rapidly, it usually converges slowly and tends to be stranded due to oscillation. Therefore many backpropagation variants were invented to improve performance by optimizing direction and step size. To name a few, we have backpropagation with momentum, conjugate gradient, Quasi-Newton and Levenberg-Marquardt [21]. After training, we can use the network to do prediction given unseen inputs. In neural network forecasting, the first step is data preparation and pre-processing. After training, we can use the network to do prediction given unseen inputs. In neural network forecasting, the first step is data preparation and pre- processing.

**5.1. Data preparation and pre-processing**

For Dow-Jones daily return data, we calculated the Hurst exponent for each period of 1024 trading days from 1/2/1930 to 5/14/2004. Among the total of 17651 periods, there are 65 periods with Hurst exponents greater than 0.65 and 1152 periods with Hurst exponents between 0.54 and 0.55. Figure 5.3 below shows the histogram of Hurst exponents for all periods.



**Figure 5.3. Histogram of all calculated Hurst exponents**

We randomly chose 30 periods from those with Hurst exponent greater than 0.65 and 30 periods from those with Hurst exponent between 0.54 and 0.55. These two groups of 60 samples constituted our initial data set.

Given a time series , how do we construct a vector from to predict ? Taken’s theorem [22] tells us that we can reconstruct the underlying dynamical system by time-delay embedding vectors if we have appropriate and . Here d is called the embedding dimension and τ is called the separation. Using the automutual information and false nearest neighbour methods [23], we can estimate and . We used the TSTOOL [24] package to run the auto-mutual information and false nearest neighbour methods for our 60 data sets. Separations of all data sets are suggested to be 1 by automutual information method. This is consistent with our intuition since we have no reason to use separated values. As for embedding dimension, our data sets are suggested to be from 3 to 5. We shall examine this later.

After building the time-delay vector and target value , we normalized the inputs and output to mean 0 and standard deviation 1. We had no need to normalize the output to a limited range, say –0.85 to 0.85, to avoid saturation because we used a linear transfer function in the output layer.

We used a commonly used approach to deal with the over-fitting problem in neural networks. We split the data set to three parts for training, validation and testing. The training data are used to adjust the weights through error back propagation. The validation data are used to stop training when the mean square error in the validation data increases. The network’s prediction performance is judged by the testing data. We used the first 60% of the data for training, the following 20% for validation and the last 20% for testing. In this way, we can have more confidence using the final network model for prediction since the forecasting data follow the testing data directly.

**5.2 Neural network construction**

Although neural networks are universal function approximators, we still need to pay much attention to their structure. How many layers should we use? How many nodes should we include in each layer? Which learning algorithm should we choose? In practice, most applications use one hidden layer since there is no apparent advantages for multi-hidden-layer networks over single-hidden-layer networks. Thus we used a single hidden layer network structure in our study. For learning algorithm, we tested the Levenberg-Marquardt, conjugate gradient method, and the back propagation with momentum algorithms. We find that Levenberg-Marquardt beats the other algorithms consistently in our test samples. We chose the Levenberg-Marquardt learning algorithm with the sigmoid transfer function in the hidden layer and a linear transfer function in the output layer. Now we need to determine the embedding dimension and the number of hidden nodes. A heuristic rule to determine the number of hidden nodes is that the total degrees of freedom of a network should equal one and half of the square root of the total data numbers. Based on this rule, we have the following equation:

|  |  |  |
| --- | --- | --- |
|  | (#input nodes+1)\*(#hidden nodes) + (#hidden nodes+1)\*(#output nodes) = 1.5sqrt(#data) | (5.2.1) |

The solution for (#hidden nodes) is 10. Similarly, we find that the number of hidden nodes for dimensions 4 and 5 are 8 and 7 respectively. For each dimension, we examine 5 network structures with the number of hidden nodes adjacent to the number suggested. For example 8, 9, 10, 11, 12 hidden nodes for dimension 3. We randomly choose 5 periods from each group (the group with Hurst exponents greater than 0.65 and the group with Hurst exponents between 0.54 and 0.55) to train each network. Each network is trained 100 times, and then the minimum NRMSE (Normalized Root Mean Square Error) is recorded. NRMSE is defined as:

|  |  |
| --- | --- |
|  | (5.2.2) |

In (5.2.2), O is output value and T is target value. NRMSE gives a performance comparison to the mean prediction method. If we always use the mean value to do prediction, NRMSE will be 1. NRMSE is 0 when all predictions are correct.

Table 5.1-5.3 gives the training results for various network structures.

|  |  |  |
| --- | --- | --- |
|  | *Dimension 3, Hidden nodes* | *MIN Std.* |
| *8 9 10 11 12* |
| 1  2  3  4  5  6  7  8  9  10 | 0.9572 0.9591 0.9632 0.9585 **0.9524**  0.9513 0.9531 0.9560 **0.9500** 0.9523  0.9350 0.9352 0.9332 **0.9328** 0.9359  0.9359 0.9426 0.9383 0.9351 **0.9313**  0.9726 0.9686 0.9652 0.9733 **0.9647**  0.9920 0.9835 0.9892 **0.9793** 0.9931  0.9831 0.9813 **0.9725** 0.9845 0.9825  0.9931 0.9852 **0.9832** 0.9877 0.9907  **0.9684** 0.9790 0.9773 0.9815 0.9862  **0.9926** 1.0044 1.0047 1.0014 1.0039 | 0.9524 0.0039  0.9500 0.0023  0.9328 0.0013  0.9313 0.0042  0.9647 0.0040  0.9793 0.0059  0.9725 0.0048  0.9832 0.0040  0.9684 0.0066  0.9926 0.0051 |
| Mean  Std. | **0.9681** 0.9692 0.9683 0.9684 0.9693  0.0225 0.0215 0.0221 0.0232 0.0255 | 0.9627 0.0042  0.0207 0.0015 |

**Table 5.1. NRMSE for dimension 3 with hidden nodes 8, 9, 10, 11 and 12**

|  |  |  |
| --- | --- | --- |
|  | *Dimension 4, Hidden nodes* | *MIN Std.* |
| *6 7 8 9 10* |
| 1  2  3  4  5  6  7  8  9  10 | 0.9572 0.9572 **0.9557** 0.9558 0.9633  0.9534 0.9554 0.9523 **0.9518** 0.9574  **0.9373** 0.9406 0.9414 0.9437 0.9404  0.9419 0.9471 0.9392 **0.9332** 0.9376  0.9691 0.9678 **0.9597** 0.9669 0.9662  0.9907 0.9939 **0.9836** 0.9948 0.9876  0.9902 0.9816 0.9872 **0.9766** 0.9855  0.9852 0.9842 **0.9802** 0.9865 0.9878  0.9809 0.9729 0.9722 **0.9669** 0.9741  **0.9916** 0.9957 0.9991 1.0025 0.9959 | 0.9557 0.0031  0.9518 0.0023  0.9373 0.0023  0.9332 0.0052  0.9597 0.0037  0.9836 0.0046  0.9766 0.0053  0.9802 0.0029  0.9669 0.0050  0.9916 0.0041 |
| Mean  Std. | 0.9698 0.9696 **0.9671** 0.9679 0.9696  0.0210 0.0193 0.0204 0.0225 0.0203 | 0.9637 0.0038  0.0196 0.0012 |

**Table 5.2. NRMSE for dimension 4 with hidden nodes 6, 7, 8, 9 and 10**

|  |  |  |
| --- | --- | --- |
|  | *Dimension 5, Hidden nodes* | *MIN Std.* |
| *5 6 7 8 9* |
| 1  2  3  4  5  6  7  8  9  10 | 0.9578 **0.9560** 0.9617 0.9589 0.9622  0.9441 0.9466 0.9456 0.9456 **0.9427**  0.9410 **0.9395** 0.9449 0.9428 0.9396  0.9546 0.9453 0.9414 **0.9409** 0.9479  0.9659 **0.9501** 0.9671 0.9653 0.9653  0.9906 0.9919 0.9898 0.9901 **0.9891**  **0.9803** 0.9819 0.9805 0.9840 0.9837  **0.9912** 0.9980 1.0009 0.9991 1.0049  0.9770 0.9747 0.9742 0.9761 **0.9689**  **0.9909** 0.9984 0.9975 0.9977 0.9951 | 0.9560 0.0026  0.9427 0.0015  0.9395 0.0023  0.9409 0.0056  0.9501 0.0071  0.9891 0.0010  0.9803 0.0017  0.9912 0.0050  0.9689 0.0032  0.9909 0.0031 |
| Mean  Std. | 0.9693 **0.9682** 0.9704 0.9701 0.9699  0.0194 0.0233 0.0220 0.0226 0.0227 | 0.9650 0.0033  0.0217 0.0020 |

**Table 5.3. NRMSE for dimension 5 with hidden nodes 5, 6, 7, 8, and 9**

From table 5.1 to 5.3, we can see that the differences of NRMSE for nodes within each dimension are very small. The number of hidden nodes with minimum average NRMSE for dimension 3, 4, 5 are 8, 8 and 6 respectively. Thus we use 8, 8, and 6 hidden nodes network for dimension 3, 4, and 5 to do prediction. Each network is trained 100 times and the minimum NRMSE is recorded. Final NRMSE is the minimum of 3 dimensions. Table 5.4 gives NRMSE of our initial 60 samples for two groups.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | *H>0.65* | *0.55>H>*  *0.54* |  | *H>0.65 0.55>H>*  *0.54* | |
| 1 | 0.9534 | 0.9863 | 16 | 0.93 | 0.9747 |
| 2 | 0.9729 | 0.9784 | 17 | 0.9218 | 0.9738 |
| 3 | 0.9948 | 0.9902 | 18 | 0.9256 | 0.9635 |
| 4 | 0.9543 | 0.9754 | 19 | 0.9326 | 0.957 |
| 5 | 0.9528 | 0.9773 | 20 | 0.937 | 0.9518 |
| 6 | 0.9518 | 0.9477 | 21 | 0.9376 | 0.9542 |
| 7 | 0.9466 | 0.9265 | 22 | 0.9402 | 0.9766 |
| 8 | 0.9339 | 0.9598 | 23 | 0.9445 | 0.9901 |
| 9 | 0.9299 | 0.9658 | 24 | 0.9498 | 0.9777 |
| 10 | 0.9435 | 0.9705 | 25 | 0.948 | 0.9814 |
| 11 | 0.9372 | 0.9641 | 26 | 0.9486 | 0.9778 |
| 12 | 0.9432 | 0.9824 | 27 | 0.9451 | 0.9968 |
| 13 | 0.9343 | 0.9557 | 28 | 0.9468 | 0.9966 |
| 14 | 0.9338 | 0.9767 | 29 | 0.9467 | 0.9977 |
| 15 | 0.9265 | 0.9793 | 30 | 0.9542 | 0.9883 |
| **Mean** | **0.9439** | **0.9731** | | | |
| **Std.** | **0.0145** | **0.0162** | | | |

**Table 5.4. NRMSE for two groups**

We ran the unpaired student’s t test for the null hypothesis that the mean values of the two groups are equal. After calculation, the t statistic is 7.369 and p-value is 7.0290e-010. This indicates the two means are significantly different and the chance of equality is essentially 0. This result confirms that the time series with larger Hurst exponent can be predicted more accurately.

1. **Conclusion**

In this paper, we analyze the Hurst exponent for all 1024-trading-day periods of the Dow-Jones index from Jan.2, 1930 to May 14, 2004. We find that the periods with large Hurst exponents can be predicted more accurately than those with H values close to random series. This suggests that stock markets are not totally random in all periods. Some periods have strong trend structure and this structure can be learnt by neural networks to benefit forecasting.

Since the Hurst exponent provides a measure for predictability, we can use this value to guide data selection before forecasting. We can identify time series with large Hurst exponents before we try to build a model for prediction. Furthermore, we can focus on the periods with large Hurst exponents. This can save time and effort and lead to better forecasting.

**References:**

[1] H.E. Hurst, Long-term storage of reservoirs: an experimental study, Transactions of the American society of civil engineers, 116, 1951, 770-799.

[2] B.B. Mandelbrot & J. Van Ness, Fractional Brownian motions, fractional noises and applications, SIAM Review, 10, 1968, 422-437

[3] B. Mandelbrot, The fractal geometry of nature (New York: W. H. Freeman, 1982)

[4] C.T. May, Nonlinear pricing : theory & applications (New York : Wiley, 1999)

[5] M. Corazza & A.G. Malliaris, Multi-Fractality in Foreign Currency Markets, Multinational Finance Journal, 6(2), 2002, 65-98.

[6] D. Grech & Z. Mazur, Can one make any crash prediction in finance using the local Hurst exponent idea? Physica A: Statistical Mechanics and its Applications, 336, 2004, 133-145

[7] E.E. Peters, Chaos and order in the capital markets: a new view of cycles, prices, and market volatility (New York: Wiley, 1991).

[8] E.E. Peters, Fractal market analysis: applying chaos theory to investment and economics (New York: Wiley, 1994).

[9] K. Hornik, M. Stinchcombe & H. White, Multilayer feedforward networks are universal approximators, Neural networks, 2(5), 1989, 259-366

[10] S. Walczak, An empirical analysis of data requirements for financial forecasting with neural networks, Journal of management information systems, 17(4), 2001, 203-222

[11] E. Gately, Neural networks for financial forecasting (New York: Wiley, 1996)

[12] A. Refenes, Neural networks in the capital markets (New York: Wiley, 1995)

[13] W. Feller, The asymptotic distribution of the range of sums of independent random variables, The annals.of mathematical statistics, 22, 1951, 427-432

[14] A.A. Anis & E.H. Lloyd, The expected value of the adjusted rescaled hurst range of independent normal summands, Biometrika, 63, 1976, 111-116

[15] W. McCulloch and W. Pitts, A logical calculus of the ideas immanent in nervous activity, Bulletin of Mathematical Biophysics, 7, 1943,:115 - 133.

[16] F. Rosenblatt, The Perceptron: a probabilistic model for Information storage and organization in the brain, Psychological Review, 65(6), 1958, 386-408.

[17] F. Rosenblatt,, Principles of neurodynamics, (Washington D.C.: Spartan Press, 1961).

[18] M. Minsky & S. Papert, Perceptrons (Cambridge, MA: MIT Press, 1969)

[19] D.E. Rumelhart, G.E. Hinton & R.J. Williams, Learning internal representations by error propagation, in Parallel distributed processing, 1, (Cambridge, MA: MIT Press, 1986)

[20] J. Yao, C.L. Tan & H. Poh, Neural networks for technical analysis: a study on LKCI, International journal of theoretical and applied finance, 2(3), 1999, 221-241

[21] T. Masters, Advanced algorithms for neural networks: a C++ sourcebook (New York: Wiley, 1995)

[22] F. Takens, Dynamical system and turbulence, lecture notes in mathematics, 898(Warwick 1980), edited by A. Rand & L.S. Young (Berlin: Springer, 1981)

[23] A.S. Soofi & L. Cao, Modelling and forecasting financial data: techniques of nonlinear dynamics (Norwell, Massachusetts: Kluwer academic publishers, 2002)

[24] C. Merkwirth, U. Parlitz, I. Wedekind & W. Lauterborn, TSTOOL user manual, http://www.physik3.gwdg.de/tstool/manual.pdf, 2002

附录D：外文翻译资料译文部分

（文章出处：第二届国际金融工程与应用会议记录2004：203-209。）

**Hurst指数和金融市场可预测性**

Bo Qian Khaled Rasheed

计算机科学系佐治亚大学，Athens，GA 30601，USA，[qian,khaled]@cs.uga.edu

**摘要**

Hurst指数（H）是一个统计学测量用来分类时间序列。当H=0.5时，表示一个完全随机的序列。而当 H>0.5 时，表示了一个具有保持趋势倾向能力的序列。H 的值越大，这个序列的倾向也越强。我们接下来将要研究如何利用Hurst指数来将不同时期的金融序列数据进行分类。BP神经网络的实验表明，具有高Hurst指数的序列比那些Hurst 指数接近于 0.50 的序列能够被更加精确的预测。因此，Hurst 指数提供了一种预测方法。

关键词：Hurst指数；时间序列分析；神经网络；蒙特卡诺模拟；预测

1. 引 言

Hurst指数是H. E. Hurst[1]提出用来作分形分析的[2][3]，现在已经被用在许多研究领域。最近，由于Peter的相关工作[7][8]，它在金融领域也变的十分热门[4][5][6]。Hurst指数为长期记忆和时间序列的分形提供了一种方法。由于它是高鲁棒性的基本系统的几个假设，现在已经被广泛用于时间序列分析。Hurst指数的值在 0 和 1 之间。基于 Hurst 指数 H，一个时间序列能够被分为三种类型：（1）H=0.5表明了序列可以用随机游走来描述。（2）0<H<0.5表明了序列具有反持续性。（3）0.5<H<1表明序列具有持续性。一个反持续性序列具有均值回复的特性，即意味着一个上升的值更有可能紧接着一个下降的值，反之亦然。H的值越接近于0.0，序列均值回复的能力也越强。而一个持续性序列具有保持倾向的能力，即下一时刻的值相对于现在值的变化，更有可能与这一时刻相对于上一时刻值的变化一致。H的值越接近于1.0，序列保持倾向的能力也越强。大多数的经济和金融时间序列具有持续性，即H>0.5。

在时间序列的预测当中，我们首先需要解决的问题是我们想要研究的这个时间序列是否可以被预测。如果这个时间序列是随机的，一切的方法都是无效的。我们想要确定这些序列具有一定的可预测等级。我们知道一个具有很高H值的时间序列是具有很强的倾向性的，所以我们自然地认为这样的时间序列要比那些H值接近于0.5的时间序列更可能被预测。接下来，我们将要使用神经网络来测试这个假设。

神经网络是无参数的通用函数逼近[9]，可以无假设地从数据中进行学习。在过去的十年里，神经网络预测模型已经被广泛应用于金融时间序列分析[10][11][12]。神经网络可以被用来代替通用函数逼近，进行预测。在同样的条件下，一个时间序列如果比另外一个时间序列具有更小的预测误差，我们便说它更容易被预测。从 1930 年1月2日到2004年5月14日，我们研究每日的道琼斯指数，计算每1024交易日的Hurst指数。从当中选出30个具有最大的Hurst指数与30个Hurst指数接近于随机序列的周期，然后我们用这些数据来训练我们的神经网络。我们对比这两组数据的预测误差，发现他们的预测误差完全不同。这个研究是通过Matlab来实现的，这篇 文 章 所 有 的 Matlab 程 序 生 成 的 结 果 都 可 以 从 [www.arches.uga.edu/~qianbo/research下](http://www.arches.uga.edu/~qianbo/research)载。

在这篇论文接下来的部分：第二部分将会详细描述Hurst指数，第三部分我们将利用蒙特卡洛模拟过程来构造一个类似让我们感兴趣的金融序列，第四部分描述了一个我们模拟生成的混乱序列用来验证根据样本顺序构造的模型，第五部分描述了神经网络和他们用来验证高 Hurst 指数的序列能够比低 Hurst 值的序列更加准确地被学习和预测。最后，这篇论文将在第六部分作出结论。

1. Hurst指数与R/S分析

Hurst指数能够通过重标极差分析(R/S)分析。对于一个时间序列，，R/S分析方法如下：

1. 计算平均值：

|  |  |  |
| --- | --- | --- |
|  |  | (2-1) |

1. 计算均值调整序列：

|  |  |  |
| --- | --- | --- |
|  |  | (2-2) |

1. 计算累计偏离序列：

|  |  |  |
| --- | --- | --- |
|  |  | (2-3) |

1. 计算序列范围：

|  |  |  |
| --- | --- | --- |
|  |  | (2-4) |

1. 计算无偏标准差序列：

|  |  |  |
| --- | --- | --- |
|  |  | (2-5) |

1. 计算重标极差序列：

|  |  |  |
| --- | --- | --- |
|  |  | (2-6) |

我们把记为区间平均值直到其中m=floor(n/t)。事实上，为了计算所有的数据，t的值是可以整除n的。

Hurst发现(R/S)随着时间的增加具有指数增长的规律，研究表明

|  |  |  |
| --- | --- | --- |
|  |  | (2-7) |

在这里c是一个常数，H被称为Hurst指数。为了分析Hurst指数，我们画出(R/S)随着t变化的log图。这条回归直线的斜率便可以用来估计Hurst指数。当t<10时，是不准确的。所以我们将利用至少10个值来计算重标极差。图2.1便是一个R/S分析的例子。

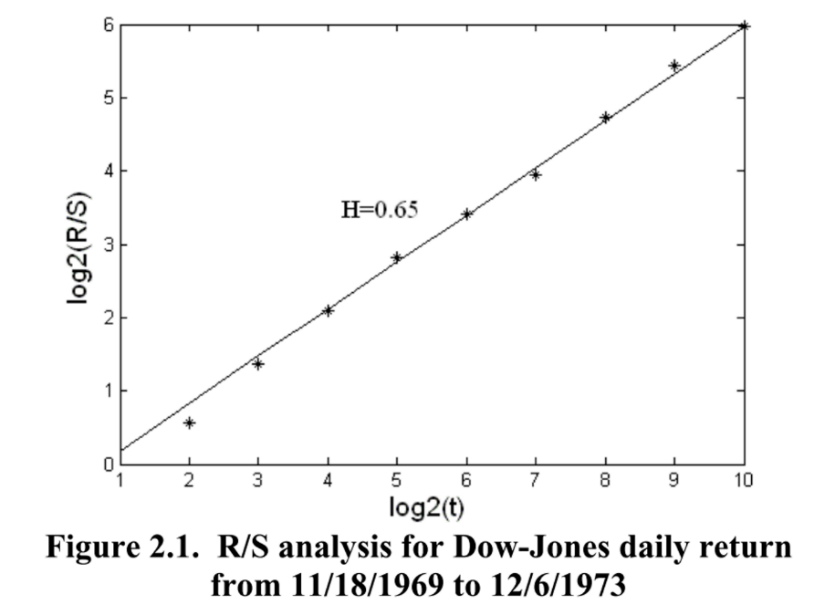


图2.1 从1969年11月18日至1973年12月6日道琼斯指数日回报率的R/S分析

在我们的实验中，我们计算了一个1024交易日周期的Hurst指数。我们用来进行回归计算。在金融领域，采用对数变化率的计算每日收益的方法十分普遍。对于累积变化对应的累积收益率，这在R/S分析中是非常有意义的。图2.2展示的是道琼斯指数从1930年1月2日至2004年5月14日的日收益率。图2.3展示的是在这个期间内对应的Hurst指数。在这个期间内，Huest指数从0.4200至0.6804波动。我们同样也想知道什么样的Hurst指数能够满足我们的条件。

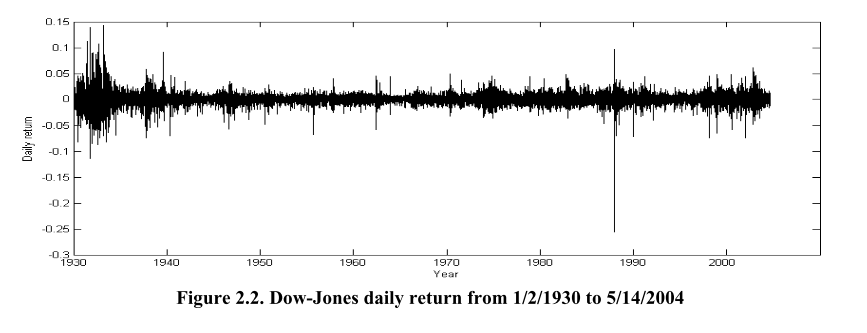


图2.2 从1930年1月1日至2004年5月14日的道琼斯日收益率

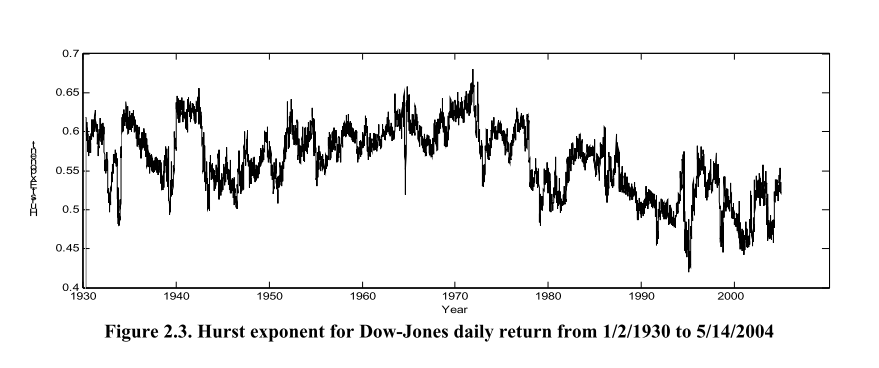


图2.3 从1930年1月2日至2004年5月14日道琼斯日收益率的Hurst指数

1. 蒙特卡洛模拟

对于一个随机序列，Feller[13]给出了一个预计的(R/S)t公式:

|  |  |  |
| --- | --- | --- |
|  |  | (3-1) |

然而，这是一个近似的关系并且只在t很大时才有效。Anis和Lloyd [14] 为了在t很小时克服这个误差提供了下面这个公式：

|  |  |  |
| --- | --- | --- |
|  |  | (3-2) |

当t>300时，大多数的计算机都很那来计算gamma函数了。利用Sterling的方程，此公式还能够被近似为：

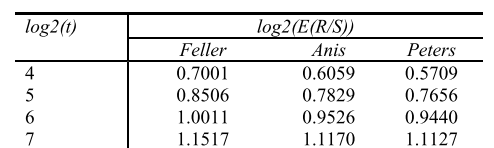
|  |  |  |
| --- | --- | --- |
|  |  | (3-3) |

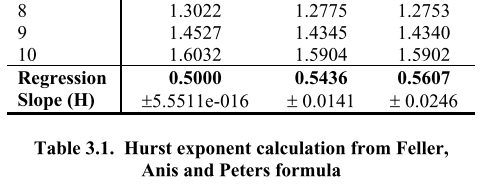
Peter又给出了公式3.2的一个修正(公式3-4)

|  |  |  |
| --- | --- | --- |
|  |  | (3-4) |

我们计算了对应期望的(R/S)值并且在的置信水平下作了平方回归。结果如表格3.1所示。

表格3.1 通过Feller，Anis，Peters公式的Hurst指数计算

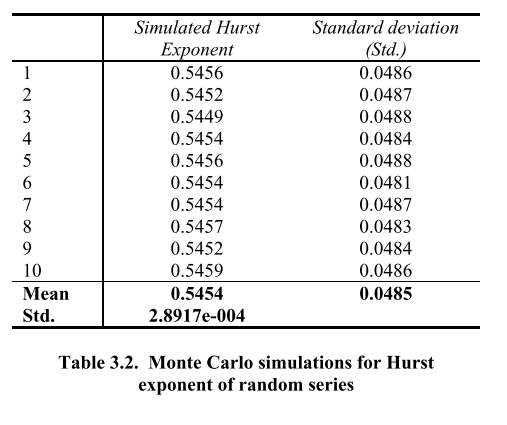




从表格3.1可以看出，对于Feller, Anis和Peter之间的公式有不少的差别。更进一步，他们的公式都是基于大量的数据点的计算。我们现在的数据量固定在1024个点。所以在我们这种情况下，随机序列的Hurst指数是怎样的？

幸运地，我们可以利用蒙特卡洛估计法来得到结果。我们生成了10000个高斯随机序列。每个序列都有1024个值。我们计算了每个序列的Hurst指数和平均值。我们希望这个平均值尽量地接近实际的值。然后重复这个过程10次。下面的表格3.2给出了这个模拟结果。

表格3.2 对于随机序列的Hurst指数Monte Carlo模拟

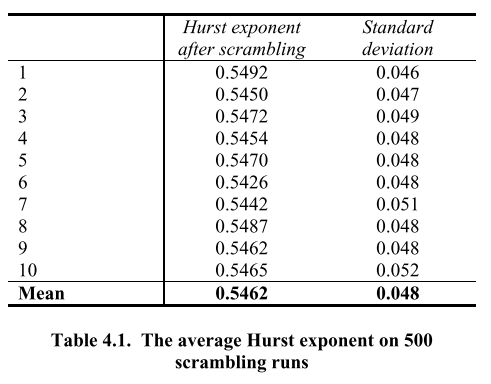


从表格3.2，我们可以看出在我们实验的情况下，可以计算出蒙特卡洛模拟的Hurst指数为0.5454，标准差为0.0485。这个结果是非常接近于Anis的公式的。通过以上的模拟，在95% 置信度的情况下，Husrt指数处在的区间，即0.4503到0.6405之间。我们选择一个Hurst指数大于0.65的区间，希望找到这些区间一些不同于随机序列的规律。然而，从这些大样本(总共17651个周期)中选出的样本中，我们想知道是否存在或碰巧存在这些时期的真实模型。为了达到这个目的，我们进行了一个混乱测试。

1. 混乱测试

为了测试是否存在一个真实的模型能满足Hurst指数大于0.65时期的样本，我们从中随机选择了10个样本。对于每一个样本，我们打乱这个序列，然后计算这个混乱序列的Hurst值。被打乱的序列与非随机序列的原始样本具有相同的分布。如果对于这些序列存在某些模型，在打乱顺序之后，这些模型将会被破坏并且计算的Hurst值也会接近于随机序列的Hurst值。在我们的实验当中，我们将每个样本打乱了500次，然后计算了平均的Hurst值。结果如表格4.1所示。

表格4.1 500步随机游走的平均Hurst值



从表格4.1我们可以看出，在样本被打乱顺序之后，Hurst指数都非常接近于0.5454，这与我们的模拟随机序列一致。通过这个结果，我们可以得出在这些时期内必然存在某些模型使得时间序列不同于随机序列，并且会被颠倒顺序被破坏。我们希望这个模型能够被用来进行预测。神经网络作为一个通用方程的近似器，提供了一个强大的工具用来学习这个潜在的模型。当潜在的规则未知时，他们变得尤为有用。我们希望用神经网络来发现这个模型并从中受益。我们用神经网络来控制错误率在可控制的范围内。接下来，我们对比Hurst指数大于0.65的时期与处于0.54和0.55之间的时期的预测误差。

1. 神经网络

在1943年，McClloch和Pitts[15]提出了一个模拟神经估算模型。这项工作被普遍认为是人工神经网络研究的开端。Rosenblatt[16][17]普及了感知器的概念并创造了很多感知器的学习规则。然而，在1969年，Minsky和Papert[18]发现感知器无法解决一些非线性可分的问题。人们认识到多层感知器(MLP)能够模拟非线性可分方程，但是没人知道如何训练他们。神经网络的研究在1986年几乎停止了。在1986年，Rumelhart[19]使用了反向传播算法来训练MLP，终于解决了这个长时间困扰人们的问题。从那以后，神经网络在很多领域重新得到了重视。神经网络开始在金融领域变得热门，并且在金融方面的神经网络研究投入资金排名第二[20]。

一个神经网络是一系列简单的相互联系的处理节点。每个节点计算加权输入，然后输出其传递函数对其他节点的值。前馈反向传播网络是最广泛使用的网络范式。利用反向传播算法训练算法，神经网络将调整权重，使它减少所观察到的输出之间的平方差（误差）和他们的目标值。反向传播算法使用梯度下降法寻找局部极小值误差曲面。它对每一个权重计算平方差的偏导数。偏导数（梯度）的相反数给出了使误差减小的方向。这个方向被称为最速下降方向。标准反向传播算法调整权重沿最速下降方向。虽然沿最速下降方向的偏差减少的最快，它通常收敛缓慢，并且可能收敛于局部最小值并振荡。因此，许多反向传播变种算法被发明，他们通过优化方向和步长来提高性能。比如说几个有名的，我们有反向传播动量，共轭梯度，准牛顿以及LM[21]。经过训练，我们可以使用这个网络来预测给定的不可见的输入。在神经网络预测中，第一步是数据的准备和预处理。经过训练后，我们可以用神经网络来预测给定的不可见的输入。

* 1. 数据的准备和预处理

对于道琼斯日收益率数据，从1930年1月2日至2004年5月14日，我们计算每1024交易日的Hurst指数。在总共17651的时期中，有65个时期的Hurst指数大于0.65,1152个时期的Hurst指数处于0.54至0.55之间。图5.3展示了所有周期的条形图。

我们从Hurst指数大于0.65的样本中随机选择了30个周期，并从Hurst指数处于0.54至0.55之间的也选择了30个周期。两组总共60个样本作为我们的原始数据集。

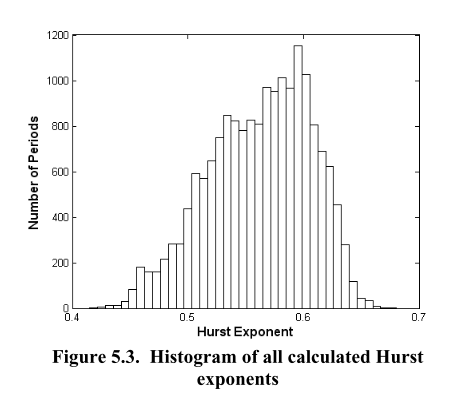


图5.3 各种情况的Hurst指数

给定一个时间序列，基于我们应该如何构造一个向量用来预测呢？Taken的理论[22]告诉我们，如果我们有合适的d和，我们可以通过延时向量来重构一个潜在的系统。这里d被称为嵌入维度，叫做分离。使用自动同步信息和假最近邻方法[23]，我们可以估算d和。我们用TSTOOL[24]包对我们的60个数据集来运行自动同步信息和假最近邻方法。建议使用自动同步信息将所有数据集的分离信息的方法。这符合我们的直觉，因为我们没有理由使用分离的值。至于嵌入维数，我们的数据集建议从3到5来取。我们将在后面进行检验。

在构造好了时间延迟向量和目标值，我们将输入和输出规范化到平均值为0，标准差为1的分布。我们没有必要将输出规范化至一个有限的区间里，比如说-0.85到0.85，以避免出现饱和的现象。因为我们在输出层使用了一个线性的转换方程。

我们使用了一个通常用来解决过拟合问题的神经网络。我们将数据集分成了三部分以用来进行训练，生效和测试。训练的数据集通过误差的反向传播来调整权重。生效集是通过判定此集合中的均方差开始增加时，停止训练。神经网络的预测性能是通过测试集来评定的。我们一般使用前60%的数据来进行训练，接下来的20%用来生效，最后的20%作为测试数据。在这种方式下，我们能够对最后通过神经网络模型进行预测的结果更加有信心。

* 1. 神经网络的构造

尽管神经网络是通用函数的近似器，我们仍然需要注意一下他们的构型。我们到底需要多少层？我们应该使用哪一种算法？实际上，到目前为止还没有直接的证据表明多隐含层的神经网络比单隐含层的神经网络更有优势，大多数在构建过程中都只使用了一层隐含层。因此，在我们的研究中将使用单隐含层的结构。对于学习算法，我们测试了Levenberg-Marquardt， 共轭梯度法，以及包含动量算法的反向传播算法。我们发现Levenberg-Marquardt 比其它的算法具有非常明显的优势。我们因此选择了Lvenberg-Marquardt算法，隐含层选用了Sigmoid传递函数，输出层则选用了一个线性函数。现在，我们需要确定隐藏的维数和隐含层的节点数。有一个探索式的规则用来确定隐含层的节点数，即神经网络总的自由度需要等于1.5倍的总数据量的平方根。根据这一条规则，我们得到以下的方程：

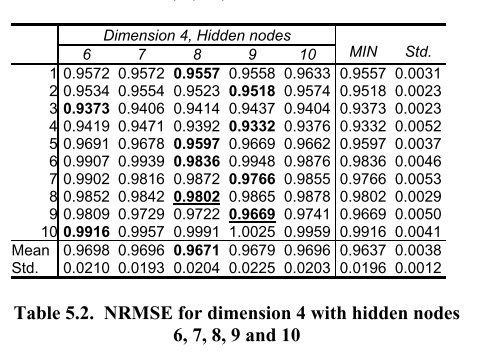
|  |  |  |
| --- | --- | --- |
|  | (#输入节点数+1)\*(#隐含层节点数)+(#隐含层节点数+1)\*(#输出节点数)=1.5\*sqrt(#数据总数) | (5-2-1) |

方程得到隐含层节点数的解是10.类似的，我们发现维度为4和5的网络隐含层的节点数为8和7。对于每一个维度，根据建议的隐含层节点数，我们测试了5种神经网络构型。例如，对于3维度的8,9,10,11,12的隐含层节点数的神经网络。我们从每一组中(Hurst指数大于0.65的组和Hurst指数处于0.54和0.55之间的组)随机选择了5个时期用来训练神经网络。每一个神经网络被训练100次，然后最小的规范误差均方根NRMSE被记录下来。NRMSE被定义为：

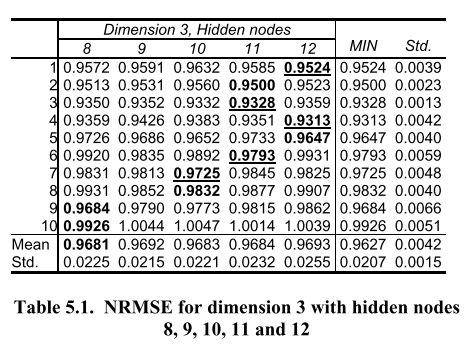
|  |  |  |
| --- | --- | --- |
|  |  | (5-2-2) |

在(5-2-2)中，O是输出的值，T是目标值。NRMSE给出了一个对于平均预测的性能比较方法。如果我们经常使用这个平均值去预测，NRMSE将会变成1.当NRMSE的值为0时表明所有的预测都是正确的。

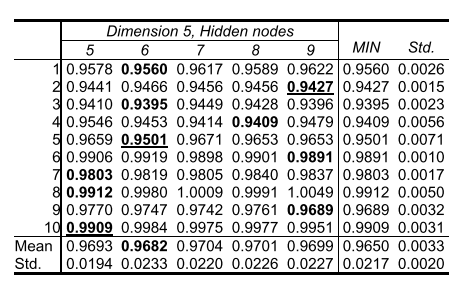
表格5.1-5.3给出了对于不同的神经网络的训练结果。

表格5.1 3维8,9,10,11,12个隐藏节点的NRMSE值

表格5.2 4维6,7,8,9,10个隐藏节点的NRMSE值



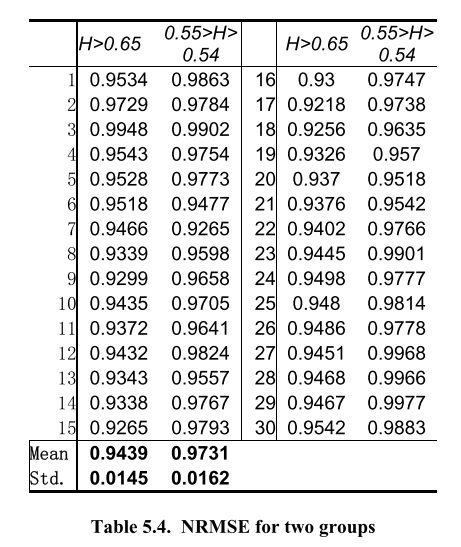
表格5.3 5维5,6,7,8,9个隐藏节点的NRMSE值



从表格5.1到5.3我们可以看出，对于每一个维度，不同的隐含层的节点数，NRMSE的差异非常小。对于3,4,5维度的网络，拥有最小平均NRMSE值的隐含层节点数是8,8,6.然后我们利用8,8,6的隐含层节点数的神经网络来对3,4,5维度的数据进行预测。每一个神经网络将会被训练100次，最小的NRMSE将会被记录下来。最后NRMSE将会是维度3的最小值。表格5.4给出了我们最初60个样本分成两组后的NRMSE值。

我们将两组不同的学生测试作为零假设，即两个组的平均值相同。再计算之后发现，t的统计量是7.369，p的值是7.0290e-010.这表示两个平均值是完全不同的，或者说他们相同的概率为0.这个结论证明了具有高Hurst值的时间序列更容易被预测精准。

表格5.4 两个不同组的NRMSE值



1. 结论

在这篇论文当中，我们分析了从1930年1月2日至2004年5月14日道琼斯股票所有1024交易日周期的Hurst指数。我们发现具有高Hurst值的数据比那些Hurst值接近于随机序列的数据更加容易被预测精确。这表明了股票市场并不是在所有时刻都是随机的。有一些时期会有很强的走势倾向模型，这些模型能够被神经网络学习并用于预测。

自从Hurst指数提供了一个评价可预测性的方法，在预测之前，我们能够使用这个值来指导我们进行数据的筛选。我们可以选择具有高Hurst值的模型来进行预测。更进一步，我们可以只关心具有高Hurst值的时期。这样能够大大地节约物力财力，指导我们更有效率低进行预测。

参考文献

1. H.E. Hurst, Long-term storage of reservoirs: an experimental study, Transactions of the American society of civil engineers, 116, 1951, 770799.
2. B.B. Mandelbrot & J. Van Ness, Fractional Brownian motions, fractional noises and applications, SIAM Review, 10, 1968, 422-437
3. B. Mandelbrot, The fractal geometry of nature (New York: W. H. Freeman, 1982).
4. C.T. May, Nonlinear pricing : theory & applications (New York : Wiley, 1999)
5. M. Corazza & A.G. Malliaris, Multi-Fractality in Foreign Currency Markets, Multinational Journal, 6(2), 2002, 65-98.
6. D. Grech & Z. Mazur, Can one make any crash prediction in finance using the local Hurst exponent idea? Physica A: Statistical Mechanics and its Applications, 336, 2004, 133-145
7. E.E. Peters, Chaos and order in the capital markets: a new view of cycles, prices, and market volatility (New York: Wiley, 1991).
8. E.E. Peters, Fractal market analysis: applying chaos theory to investment and economics (New York: Wiley, 1994).
9. K. Hornik, M. Stinchcombe & H. White, Multilayer feedforward networks are requirements for universal. Neural networks, 2(5), 1989, 259-366 [10]
10. S.Walczak, An empirical analysis of financial forecasting with approximators, data neural networks, Journal of management information systems, 17(4), 2001, 203-222
11. E. Gately, Neural networks for financial forecasting (New York: Wiley, 1996)
12. A. Refenes, Neural networks in the capital markets (New York: Wiley, 1995)
13. W. Feller, The asymptotic distribution of the range of sums of independent mathematical statistics, 22, 1951, 427-432
14. A.A. Anis & E.H. Lloyd, The expected value of the adjusted rescaled hurst summands, Biometrika, 63, 1976, 111-116
15. W. McCulloch and W. Pitts, A logical calculus of the ideas immanent in Mathematical Biophysics, 7, 1943,:115 - 133.
16. F. Rosenblatt, The Perceptron: a probabilistic model for Information storage and organization in the brain, Psychological Review, 65(6), 1958, 386-408.
17. F. Rosenblatt,, Principles of neurodynamics, (Washington D.C.: Spartan Press, 1961).
18. M. Minsky & S. Papert, Perceptrons (Cambridge, MA: MIT Press, 1969) [19] D.E. Rumelhart, G.E. Hinton & R.J. Williams, Learning internal representations by error propagation, in Parallel distributed processing,
19. 1, (Cambridge, MA: MIT Press, 1986)
20. J. Yao, C.L. Tan & H. Poh, Neural networks for technical analysis: a study on LKCI, International journal of theoretical and applied finance, 2(3), 1999, 221-241
21. T. Masters, Advanced algorithms for neural networks : a C++ sourcebook (New York: Wiley, 1995)
22. F. Takens, Dynamical system and turbulence, lecture notes in mathematics, 898(Warwick 1980), edited by A. Rand & L.S. Young (Berlin: Springer, 1981)
23. A.S. Soofi & L. Cao, Modelling and forecasting financial data: techniques of nonlinear dynamics(Norwell, Massachusetts: Kluwer academic publishers, 2002)
24. C. Merkwirth, U. Parlitz, I. Wedekind & W. Lauterborn, TSTOOL user manual, http://www.physik3.gwdg.de/tstool/manual.pdf, 2002

附件一

**中国民航大学**

**毕 业 设 计（论 文）任 务 书**

中欧航空工程师 学院 飞行器动力工程 专业 学号: 121143325

课 题：基于深度残差网络的国际金价预测研究

完成期限：自 2016 年 1 月 11 日至 2016 年 5 月 25 日

学生姓名： 汤 吉

学 号： 121143325

学 院：中欧航空工程师学院

专 业： 飞行器动力工程

指导教师： 张 鸿 燕

主管院长： （签章）

批准日期： 2015年 11月30 日

|  |
| --- |
| 一、课题的目的与意义  本课题针对国际黄金价格预测问题，采用国际黄金价格机器学习系统(IGMLS, International Gold price Machine Learning System)和私人专家系统(PES, Private Expert System)相结合的方法，开发快速、精确、可行、高鲁棒性的国际黄金价格分析系统。 本项目不论是对国际宏观经济和微观经济方面，还是对国家间经济活动方面都有十分重要的意义。 |
| 二、课题在专业技术上的要求和具体的量化指标（包括外文资料翻译、论文的字 数、完成期限等）  1、论文主要研究内容是在研究机器学习算法在经济学中的应用。  2、翻译（原文及译文），其中中文部不少于 2000 字，应按指定的外文翻译资料翻译，并且应在毕业设计（论文）开始后两周内完成；  3、论文正文部分字数应不少于 10000 字；  4、论文完成期限：自 2016 年 1 月 11 日至 2016 年 5 月 25 日。 |
| 三、课题完成后应提交的文本、计算、图表、实验报告等清单  1、毕业设计论文文本及中文和英文摘要（纸质和电子版）；  2、测试程序 ；  3、外文资料翻译（原文及译文）。 |
| 四、主要参考文献与资料  [1] Hansen A ,, Schmittbuhl J ,, Batrouni G G. Distinguishing fractional and white noise in one and two dimensions [J]. Physical Review E Statistical Nonlinear & Soft Matter Physics, 2000, 63(6):062102.  [2] Lloyd E H, Hurst H E, Black R P, et al. Long-Term Storage: An Experimental Study [J]. Journal of the Royal Statistical Society, 1966, 129(4).  [3] Kantelhardt J W, Koscielny-Bunde E, Rego H H A, et al. Detecting long-range correlations with detrended fluctuation analysis [J]. Physica A Statistical Mechanics & Its Applications, 2001, 295(s 3–4):441-454  [4] Abbad M, Oesterlé B. Lift Force Effects on the Behavior of Bubbles in Homogeneous Isotropic Turbulence [C]// ASME 2005 Fluids Engineering Division Summer Meeting. American Society of Mechanical Engineers, 2005:659-666.  [5] Mccauley J L, Bassler K E, Gunaratne G H. Martingales, detrending data, and the efficient market hypothesis [J]. Physica A Statistical Mechanics & Its Applications, 2008, 387(1):202-216.  [6] Yu S, Piao X, Hong J, et al. Bloch-like waves in random-walk potentials based on supersymmetry [J]. Nature Communications, 2015, 6.  [7] Provost F, Kohavi R. Guest Editors' Introduction: On Applied Research in Machine Learning [J]. Machine Learning, 1998, 30(2-3):127-132.  [8] Stoyanov J. Probability for Statistics and Machine Learning: Fundamentals and Advanced Topics A. DasGupta, 2011 New York, Springer xx + 782 pp., £81.00 ISBN 978-1-441-99633-6 [J]. Journal of the Royal Statistical Society, 2014, 177(2):566-566.  [9] Novotný F, Raková M. Assessment of Consensus Forecasts Accuracy: The Czech National Bank Perspective [J]. Working Papers, 2010, 61(4):348-366.  [10] Shafiee S, Topal E. An overview of global gold market and gold price forecasting [J]. Resources Policy, 2010, 35(3):178–189. |

附件二

中国民航大学毕业设计（论文）开题报告

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 姓 名 | 汤吉 | 班 级 | 126041A | 学 号 | 121143325 |
| 课题名称 | 基于深度残差网络的国际金价预测研究 | | | 指导教师 | 张鸿燕 |
| 一、课题的目的和意义 | | | | | |
| 本项目针对国际黄金价格预测问题，采用国际黄金价格机器学习系统(IGMLS, International Gold price Machine Learning System)和私人专家系统(PES, Private Expert System)相结合的方法，开发快速、精确、可行、高鲁棒性的国际黄金价格分析系统。本项目不论是对国际宏观经济和微观经济方面，还是对国家间经济活动方面都有十分重要的意义。 | | | | | |
| 二、课题的任务和要求 | | | | | |
| 1）毕业设计论文文本及中文和英文摘要（纸质和电子版）；  2）结合历史数据的计算分析结果和数据挖掘的结果，设计并测试程序系统；  3）外文资料翻译（原文及译文）；  4）利用系统对国际黄金价格走势进行一定精度的预测。 | | | | | |
| 三、主要设计思路 | | | | | |
| 本项目旨在利用最近十年的国际黄金价格、大宗商品价格以及国际经济政治事件历史数据，建立并优化出最适合的经济模型，开发对应的IGMLS。并利用各国际著名财经分析师的实时预测分析结果，开发出对应的PES。最后将两系统的计算数据加权结合，建立一套完整的国际黄金价格预测系统。  该项目的主要研究内容有:  1）构建国际黄金价格预测深度学习神经网络；  2）利用最近十年的国际黄金价格、大宗商品价格以及国际经济政治事件历史数据对神经网络进行训练；  3）不断训练优化此神经网络，建立IGMLS；  4）对互联网上各著名财经分析师的实时分析结果进行数据提取，分析，优化，结合，建立PES；  5）将两系统的计算数据加权结合，建立一套完整的国际黄金价格预测系统。 | | | | | |
| 四、毕业设计结束后须提交的材料清单 | | | | | |
| 1）论文主要研究内容是在研究机器学习算法在经济学中的应用。  2）翻译（原文及译文），其中中文部不少于2000字，应按指定的外文翻译资料翻译，并且应在毕业设计（论文）开始后两周内完成；  3）论文正文部分字数应不少于10000字；  4）论文完成期限：自2016年1月11日至2016年5月25日。 | | | | | |

毕 业 设 计（论 文）进 度 安 排 表

|  |  |
| --- | --- |
| 时间 | 计 划 完 成 内 容 |
| 第1周 | 查阅国内外相关文献 |
| 第2周 | 撰写开题报告 |
| 第3周 | 查询外文文献 |
| 第4周 | 翻译外文文献 |
| 第5周 | 了解金融学基础知识 |
| 第6周 | 学习机器学习基础算法 |
| 第7周 | 学习机器学习高级算法 |
| 第8周 | 了解神经网络基本原理，并进行简单的计算机算法实现 |
| 第9周 | 查阅神经网络应用方面的相关文献，优化神经网络算法代码 |
| 第10周 | 了解学习深度学习算法，查找相关开源代码 |
| 第11周 | 查阅深度学习应用方面的相关文献 |
| 第12周 | 用R/S分析法从国际黄金价格历史数据中筛选出符合用于神经网络训练条件的数据集 |
| 第13周 | 尝试构建用于预测国际黄金价格的深度学习神经网络，并利用数据集进行训练 |
| 第14周 | 测试机器学习系统的预测结果，查找BUG并改良系统 |
| 第15周 | 利用互联网资源，搭建专家系统搜寻金融领域信息并对预测结果进行修正 |
| 第16周 | 利用机器学习系统和私人专家系统，对实时国际黄金价格进行预测并分析结果 |
| 第17周 | 利用结果分析各系统中的不足，并调试分析 |
| 第18周 | 撰写毕业论文 |
| 第19周 | 准备答辩 |
| 指  导  教  师  意  见 | 指导教师签字： 年 月 日 |

附件三

中国民航大学毕业设计（论文）指导教师评阅表

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 课题名称 | 基于深度残差网络的国际金价研究 | | 指导教师 | 张鸿燕 |
| 学生姓名 | 汤吉 | 专 业 | 飞行器动力工程 | |
| 毕业设计（论文） | 评 价 内 容 | | | 评分 |
| 毕业设计（论文）文稿规范性，主要包括：摘要、正文、参考文献等（满分20） | | |  |
| 毕业设计（论文）质量，主要包括：问题综述、基本概念及问题分析、解决问题的基本方法步骤和设计方案及实现、总结等（满分50） | | |  |
| 外文翻译（满分10分） | | |  |
| 工作态度、网上周志填写及任务完成情况（满分20） | | |  |
| 总 分 | | |  |
| 评 语：  指导教师签字：  年 月 日 | | | | |

附件四

中国民航大学毕业设计（论文）评阅教师评阅表

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 课题名称 | 基于深度残差网络的国际  金价预测研究 | | 评阅教师 |  |
| 学生姓名 | 汤吉 | 专 业 | 飞行器动力工程 | |
| 毕业设计（论文） | 评 价 内 容 | | | 评 分 |
| 毕业设计（论文）文稿规范性，主要包括：摘要、正文、参考文献等（满分20） | | |  |
| 毕业设计（论文）质量，主要包括：问题综述、基本概念及问题分析、解决问题的基本方法步骤和设计方案及实现、总结等（满分50） | | |  |
| 符合专业培养目标，工作量饱满，具有一定的难度（满分20） | | |  |
| 对前人工作有改进或突破，或有独特见解  （满分10分） | | |  |
| 总 分 | | |  |
| 评阅教师签字：  年 月 日 | | | | |

附件五

中国民航大学毕业设计（论文）答辩评语及成绩评定表

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 课题名称 | 基于深度残差网络的国际金价预测研究 | | | | | | | | |
| 学生姓名 | 汤吉 | 专业 | 飞行器动力工程 | | | 学号 | | 121143325 | |
| 答辩小组成员 |  | | | | | | | | |
| 评价内容 | 具体要求 | | | | 满分 | | | | 评分 |
| 论文内容 | 思路清晰；概念清楚，语言表达准确。论点正确；方法科学，分析归纳合理；结构严谨；有应用价值。 | | | | 50 | | | |  |
| 创 新 | 对前人工作有改进或突破，或有独特见解。 | | | | 10 | | | |  |
| 答 辩 | 回答问题有理论依据，基本概念清楚，主要问题回答准确。 | | | | 40 | | | |  |
| 总 分 | | | | | | | | |  |
| 组长签字：  答辩成员签字：  年 月 日 | | | | | | | | | |
| 指导教师评分（30%） | 评阅教师评分（20%） | | | 答辩小组评分（50%） | | | 毕业设计（论文）  成绩 | | |
|  |  | | |  | | |  | | |

注：1）毕业设计（论文）的成绩应由答辩小组根据指导教师评定成绩（30%）、论文评阅成

绩（20%）和答辩成绩（50%）综合评定。

2）五分制与百分制对照：优（100-90）；良（89-80）；中 （79-70）；及格（69-60）；

不及格（59以下）。