

深度学习能更好地预测 人民币汇率吗?

© 周颖刚 陈海鹏^①

内容摘要: 本文试图回答深度学习的新技术是否能更好地预测人民币汇率波动。为此,我们运用深度学习方法改善并提出多层卷积神经网络(CNN),分别构建了预测汇率波动的长期与短期模型,分析发现网络搜索次数等新兴媒体指标能够提高短期模型的预测结果,同时合并长、短期模型的结果优于直接利用所有影响因素进行预测的结果,也优于采用传统的神经网络以及时间序列分析方法(如广义自回归条件异方差、贝叶斯平均分类回归等模型)进行预测的结果。

关键词: 汇率预测 CNN 神经网络 小波分析 主成分分析 新兴媒体指标

中图分类号: TP18; F832.6 **文献标识码:** A

一、引言

1994年1月,中国实现了官方汇率与外汇调剂市场汇率并轨,并于2005年7月21日起建立了以市场供求为基础、单一有管理的浮动汇率制。2007年5月,中国人民银行决定,银行间即期外汇市场人民币兑美元交易价浮动幅度,由0.3%扩大到了0.5%。而在2010年6月19日,为了增加人民币汇率弹性,中国人民银行宣布进一步推进人民币汇率形成机制改革,到了2012年4月,即期外汇市场人民币兑美元交易价浮动幅度进一步扩大到1%,2014

① 作者简介:周颖刚,厦门大学经济学院与王亚南经济研究院,教授,研究方向:国际金融与贸易等。

陈海鹏,厦门大学信息科学与技术学院,硕士研究生。

基金项目:本文获厦门大学中央高校基本科研业务费项目(20720161012)和闽都中小银行教育发展基金会的资助。

年3月,又扩大至2%。2015年8月11日,央行宣布调整人民币对美元汇率中间价报价机制,做市商参考上日银行间外汇市场收盘汇率进行报价。2016年10月1日,人民币正式纳入国际货币基金组织特别提款权(SDR)货币篮子。2017年10月,中共十九大在北京召开,要求“深化利率和汇率市场化改革”。随着人民币国际化和汇率市场化改革的不断深化,人民币汇率的变化受到国内外的高度关注,如何预测人民币汇率成为一个非常重要的研究课题。

传统的汇率波动预测均是以历史汇率作为基本数据的技术分析法或者是基于多重影响指标的马尔科夫链等线性预测模型,不仅无法很好地描述市场变化的随机性与波动性,而且没有考虑诸如网络舆论与媒体导向等新兴网络因素对于汇率波动的影响。而如今应用广泛的深度学习对于时间序列数据的学习与预测有着较好的效果,其对于波动性的判断与预测有着良好的模拟性。

2006年Hinton等人提出了深度学习的概念,其主要目的是为了解决神经网络内部原先存在的隐含层数过多而难以训练的问题。现代神经网络最早可追溯到20世纪40年代,由Warren McCulloch与Walter Pitts联合提出理论上神经网络可以计算任何算术与逻辑函数。而在20世纪50年代后期,由Frank Rosenblatt第一次构造出神经网络的实际应用模型。由此神经网络步入第一个研究热潮,但早期由其搭建的模型都有着固有的局限性,网络内部结构简单运算量小,仅能初步解决一些简单问题,许多人认为神经网络已经步入了死胡同,直到20世纪80年代,随着计算机计算能力的提高和James Anderson等人提出了能够完成记忆的新型网络,神经网络又焕发青春,其中诺贝尔奖得主Edelman提出的Darwinism将协同学引入神经网络之中,拓宽了神经网络的应用范围。但随着大数据时代的到来,原先神经网络的隐含层数少,对于高维数据或者大数据量难以训练的局限使其再次陷入瓶颈。直到由Hinton等人提出深度学习的概念,使得内部具有复杂结构、隐含层多的神经网络可以通过空间参数修正来完成训练目标,并提高训练性能,达到了良好的训练效果,基于深度学习的神经网络算法才又掀起一股热潮。

而经过深度学习改善过的神经网络模型如今正处于一个飞速发展与应用的时期,其以生物学的神经元传播原理为基础,通过一定数量的数据学习来完成网络的构造,最后输入自变量便可以完成因变量的预测。拥有自我学习、随机模拟、联想识别等特点,是一个非线性的动态系统。与传统的线性预测模型相比,神经网络的精度更高,也能更好地描述自变量的随机变化,并能通过训练误差与泛化误差很好地确定自变量对于网络的影响。

本文将运用经过深度学习改善过的神经网络进行人民币汇率波动预测,由于汇率的影响因素中包括了长期影响部分与短期影响部分,因此我们对于汇率的长短期影响部分分别进行相关的预测研究,有利于研究汇率波动的长期走势与短期突变,进行精确的汇

率预测。下文的结构安排如下：第二部分是文献综述与本文贡献；第三部分选择人民币汇率预测的因素；第四部分介绍多层卷积神经网络（CNN）及小波分析方法；第五部分是基于深度学习的人民币汇率预测实证研究；第六部分是结论。

二、文献综述与本文贡献

90年代初，Haken把协同学与神经网络相结合，他认为协同学可以在网络检验与梯度验算中减小网络的损失度，并据此发表了著作 *Synergetic and Cognition: Atop-Down Approach to Neural Nets*。由此神经网络的运用到达了新的巅峰，也渐渐地开始有经济学家利用神经网络对于汇率波动进行预测。Kuan和Liu（1995）使用同质神经网络对5种不同货币汇率进行预测，实证表明英镑、日元的预测结果均方差（MSE）很低，而马克、瑞士法郎、加拿大元的效果一般。Hu和Zhang（1998）则是在利用多层后向神经网络对汇率进行预测时，得出了神经网络的预测效果比线性模型好的结论，并提出，预测期越短，神经网络相较于其他网络的预测性能越好。黄志刚（2013）完成了利用GABP神经网络对于汇率波动弹性的测度，并得出神经网络可以较好地拟合政策与市场变化的事实。TD Chaudhuri（2016）在通过神经网络对汇率的预测结果进行分析时提出，神经网络在时间序列的预测中具有良好的优越性。

高精度的神经网络预测模型的前提是输入变量的选取程度，影响汇率波动的学说有购买力平价理论、利率平价理论、货币主义的汇率决定理论、汇率资产组合理论等，每个学说都提出了相应的影响因素，同时随着互联网的普及与发展，对于汇率市场中的消费者而言，网络舆论以及网络新闻媒体的报道往往能对他们的行为产生较大的影响，这种新兴媒体的舆论指标同样能影响汇率波动。

同时还需注意的是，汇率的影响因素可以分为长期与短期部分，而长期与短期的汇率影响部分对于汇率波动的影响形式是不尽相同的。短期影响部分可能只对即期汇率产生影响，作用时间快，持续时间短，而长期影响部分对即期汇率影响甚微，其影响区间往往是在未来的一段时间内。郭莹莹（2015）就利用马尔科夫模型分析界定了影响汇率波动的长短期因素对于汇率波动的影响周期与作用大小。因此对于影响因素，应该加以区分，而后选取不同的模型进行预测。本文选取的影响因素，有许多经济学家已经对其进行了界定。Clark和Macdonald（2000）完善了行为汇率均衡模型（BEER），从短期暂时影响、中长期影响两个角度分析，认为长期均衡汇率的影响因素主要有巴拉萨-萨缪尔逊效应、净国外资产与贸易条件。高铁梅等（2008）从产品市场与资本市场出发，构建

人民币汇率的行为均衡模型,分析表明,生产力水平、贸易条件、开放度、政府支出等因素在长期和短期内对汇率产生冲击。Ricci等(2008)研究政府消费在实际有效汇率的长期决定因素中的显著性,他们发现政府支出/支出每增加1%都会使实际有效汇率上升3%,政府支出在长期的汇率波动中具有显著性。与此同时,Williamson(1983)提出改进的基本要素均衡汇率理论(FEER),其中指出汇率波动取决于长期影响因素,如经常项目余额、国民生产总值、政府支出等。Karfakis(2006)采用协整分析和格兰特检验分析得出,货币供给量(M2)与投资率等长期影响因素与汇率有着长期稳定的关系。杨帆(1995)通过对历史人民币汇率的研究指出,人民币汇率波动的长期影响因素是外汇储备和货币供给,短期影响因素为利率、通货膨胀率和经常项目余额。

对于各影响因素而言,需要对其数据中的长期部分与短期部分进行分离。在金融数据的分离中,因为其波动大、频谱宽,是时序不平稳序列,传统的滤波方法如傅里叶变换、卡尔曼滤波理论和维纳滤波(1989)均不能适用,而小波分析以其紧支性、多分辨性、衰减性和快速逼近的特点,更适合来做金融数据长短期部分的分离(2004)。钱舒、袁修贵(2002)等就曾使用小波分析对股票数据进行了分离,并取得了良好的效果。

进行输入数据的分离确定之后就可以将分离之后的长期影响因素部分与短期影响因素加上长期因素的短期部分分别使用卷积神经网络(CNN)进行预测拟合,将预测结果与传统的神经网络预测结果进行对比。YSUN(2012)与徐缘圆(2013)分别使用了传统的BP神经网络对人民币兑美元汇率进行了预测,分别得到了97.3%与98.2%的预测精度。之后,L Falat(2016)利用改造后的径向基神经网络,将预测精度提高到了98.72%。

从已有的相关文献来看,利用神经网络对汇率波动进行预测会随着时间的延续而有越来越多人研究。实践表明,从基本面分析的预测结果在长期走势上拥有较高的准确度,但是在中短期的预测方面,技术面分析的精度更高,同时,大多数文献均是从预测方法层面对汇率预测提出改进,但是将长短期影响因素分开进行研究,并对于长短期影响因素进行内部分离,同时加入新兴媒体指标来解释汇率突变现象,还鲜有人提出讨论。

而本文的主要贡献在于首先利用小波分析,对汇率影响因素的长短期因素进行分离,并对汇率的长短期部分进行相应的预测,这种分长短期进行预测的方法是较为新颖的;其次利用经深度学习方法改善而提出的卷积神经网络对人民币汇率进行预测,将其与传统的神经网络和时间序列的预测方法自回归平均异方差、贝叶斯分类回归等算法预测结果进行对比,表明本文采用方法的优越性;最后初步尝试利用新兴媒体指标来解释汇率波动,证明了其对于汇率的波动性预测有一定的帮助。

三、选择预测汇率的因素

在对汇率进行预测之前，需要了解决定、影响汇率的影响因素，从而选取合适的解释变量对其进行研究。我们首先根据购买力平价理论、利率平价理论、货币主义汇率假说、汇率资产组合理论和汇率新闻理论选择以下影响汇率波动的因素，如表 1 所示。

表 1 汇率影响因素汇总表

购买力评价理论	物价水平
	通货膨胀率
利率平价理论	利率
货币主义的汇率决定理论	货币供应量（M2）
	物价水平
	国民收入
	利率
汇率的资产组合理论	经常项目
汇率的新闻模型	百度指数——人民币汇率上升搜索次数
	中国国家外汇管理局网站点击率
	国际原油价格
	恐慌指数（VIX）

应该指出的是，汇率的新闻模型认为，不仅仅是诸如利率、货币供应量、经常项目等因素影响着汇率波动，与汇率毫不相关的信息诸如国际石油价格、政府政策变化等新闻，也会作为随机误差项干扰着汇率波动。该学说提倡将之前理论未曾关注到的场外因素加入到影响汇率波动的因素中来。这也符合如今经济全球化与互联网普及的潮流。现如今每天人们都可以从社交媒体、报纸电视广播上摄取到各类信息。而主流媒体与舆论导向的影响力也越来越大，有能力引导改变投资者的预期从而影响汇率波动。因此，主流媒体论调与舆论导向也是汇率波动的影响因素之一。在这里，我们选取搜集了百度网站上人民币汇率上升的搜索次数作为影响因素来进行研究。

同时，恐慌指数（VIX）是芝加哥交易所（CBOE）提出的由指数期权隐含波动率加权平均后所得的指数，其数值越高，代表市场参与者预期后市波动程度会更加激烈，同时也反映了市场参与者的心理状态，在此我们认为其在一定程度上反映了部分外汇市场参与者的心理状态，也纳入影响因素之中。

此外，Clark 和 MacDonald（2000）完善了行为汇率均衡模型 BEER 模型，从短期暂

时影响、中长期影响两个角度分析,认为长期均衡汇率的影响因素主要有巴拉萨-萨缪尔逊效应、净国外资产与贸易条件。高铁梅等(2008)从产品市场与资本市场出发,构建人民币汇率的行为均衡模型,分析表明,生产力水平、贸易条件、开放度、政府支出等因素在长期和短期内对汇率产生冲击。Ricci等(2008)研究政府消费在实际有效汇率的长期决定因素中的显著性,他们发现政府支出/支出每增加1%都会使实际有效汇率上升3%,政府支出在长期的汇率波动中具有显著性。

我们在表1的基础上再次加入外汇储备(FER)、政府支出(GB)、贸易条件(NBTT)、投资率(INN)、经济开放度(EOR),同时把国民生产总值转化为经济增长率(RGDP)、经常项目转化为净资本流入(CIN)。

在选取好影响因素之后,我们还需要对这些影响因素进行分类,因为不同的影响因素对于汇率波动的时间周期是不同的,如果同时一起研究这些影响,那么得到的结果会是片面与不精确的。Williamson(1983)提出改进的基本要素汇率均衡理论(FEER)模型,其中指出汇率波动取决于长期影响因素,如经常项目余额、国民生产总值、政府支出等。Karfakis(2006)采用协整分析和格兰特检验分析得出,货币供给量(M2)与投资率等长期影响因素与汇率有着长期稳定的关系。杨帆(1995)通过对历史人民币汇率的研究指出,人民币汇率波动的长期影响因素是外汇储备和货币供给,中期影响因素是经济开放度和贸易条件,短期影响因素为利率、通货膨胀率和经常项目余额。依据上述的文献分析,把上述选取的影响因素分类为短期影响因素与中长期影响因素,分类结果如表2所示。

表2 长/短期影响因素汇总表

中长期影响因素	货币供给(M2)
	经济增长率(RGDP)
	投资率(INN)
	贸易条件(NBTT)
	经济开放度(EOR)
短期影响因素	国内利率(DIR)
	国内通货膨胀率(DINF)
	国外通货膨胀率(FINF)
	国外利率(FIR)
	净资本流入(CIN)
	国际原油价格(OIL)
	百度指数——人民币汇率上升搜索次数
	恐慌指数(VIX)

四、多层卷积神经网络及小波分析

人工神经网络（Artificial Neural Network, ANN）又简称神经网络（NN），是现代人工智能的基础，神经网络是基于生物学的人脑神经元传导，构造可计算的单元模拟人脑的神经元细胞，通过可计算单元的连接来构造网络。

20 世纪 90 年代初，Haken 把协同学与神经网络相结合，他认为协同学可以在网络检验与梯度验算中减小网络的损失度，并据此发表了著作 *Synergetic and Cognition: Atop-Down Approach to Neural Nets*。此后 Edelman 也提出了著名的 Darwinism 模型，正是这两个理论的完善，把神经网络推向了新的高度。而自 2000 以来，神经网络开始不再局限于模式识别的领域，而是广泛运用于图像处理、最优化分析、预测、经济调研、深度学习方面。

神经网络是一个非线性的分布式信息网络，由于是基于生物学基础创建而来，所以内部的每一个计算单元都称为神经元，每个神经元可以有一个或者多个输入端，但是只能有一个输出端，每个神经元内部具有相应的数学公式或者逻辑公式，同时，每一个输入输出端所构成的连接通道都附有相应的权重值。

（一）CNN 神经网络

CNN 是一个多层的卷积神经网络，每一层包含多个二维平面，每一个平面中又包括了多个独立的神经元。其内部包含了对于输入数据的特征提取过程，特征提取过程一般有 1 到 3 个阶段，每一个阶段紧随着一个全连接神经网络的分类器。

数据输入之后即被放入输入层，此后每一层的输入均为前一层的一组小的局部临近单元，首先对输入层分别在六个训练滤波器进行卷积运算，加上偏置之后就通过激励函数得到初步的特征矩阵，之后再行下采样，再经过一个激励函数即得到特征图。之后将特征图再重复一次如上操作，并将所得的特征图连接，即为提取的特征向量，将特征向量采用 CNN 算法进行训练，反向传播以误差最小化为目标。由于每一层只取前一层的局部领域作为输入，因此称为局部感受野，而每个特征映射生成的平面上的所有神经元权值相等，这就是权值共享原则。

CNN 神经网络的工作原理正是如前所述，分为正向传播与反向传播两个阶段，在正向传播中由输出端开始，通过连接通道进行逐层的传递，每一层的输出仅能传播到下一层神经元的输入端，一旦输出端的数值与真实值出现偏差，便产生一个误差信号，开始进行反向传播回溯，逐层地进行梯度搜索，利用最小二乘算法（Least Mean Squares），逐层地修改每一层的阈值函数与权值矩阵，不断地逼近真实值，最后达到训练值的误差平

方和最小, 图 1 给出 CNN 的网络结构图。

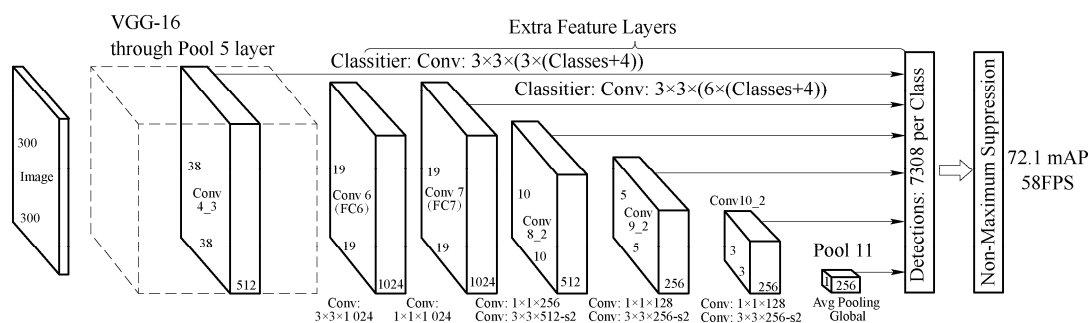


图 1 CNN 神经网络结构图

CNN 神经网络是一个良好的非线性优化系统, 它的优点为结构简单且工作状态稳定, 对于输入的数据, CNN 神经网络能够高精度地自学习, 而且其训练的网络泛化性能好。但是它同样有其缺陷之处, 首先就是 CNN 神经网络因为采用了梯度下降的优化算法, 所以其训练时间大大加长, 不利于大数据的输入学习, 同时, 随着 CNN 神经网络的训练次数增加, 其泛化精度会随之下降。这导致我们如果使用大数据训练, 可能会出现“过度训练”的问题, 而我们的短期影响因素刚好具有数据量少、需要高精度与高泛化程度的需求, 完全符合利用 CNN 神经网络的优势, 而长期影响由于已经被切分成了数个滞后期的重复叠加运算, 恰好可以利用 CNN 神经网络中的卷积计算能够描述前后数据关联性的优点。

(二) 小波分析

小波是一种长度有限、均值为 0 的波形, 其特点为在时域上具有紧支性而直流分量为 0。小波分析是将特定波形的小波进行伸缩平移变换从而多尺度地细化分析, 由于金融数据通常不是平稳序列, 所以对于高频与低频数据部分的分离无法通过经典维纳来进行, 而小波变换具有时频局部化、多分辨率、解相关、选机灵活的特性, 恰好适合用来对长期影响因素进行高低频的分离。

小波分析中, 高低频部分的小波系数幅值随尺度变化呈现不同的变换趋势, 高频部分的幅值会随着尺度的增加而衰减为零, 而低频部分的幅值则大致相当不会有太大变化。所以小波分解的大致步骤可分为: 小波分解, 选择特定小波对数据进行正交分解, 得到分解的每一层的频率系数; 而后依据前一步的分解结果进行阈值的设定, 对每一个分解层的系数按照阈值进行处理。最后按照最低一层的分解层的系数和第二步经过阈值处理过的系数进行数据的重构, 即可得到低频部分。

五、基于深度学习的人民币汇率预测实证研究

(一) 数据

基于上述第二部分的论述,我们选取了货币供给(M2)、经济增长率(RGDP)、投资率(INN)、贸易条件(NBTT)、经济开放度(EOR)作为中长期影响因素;而选择国内利率(DIR)、国外利率(FIR)国内通货膨胀率(DINF)、国外通货膨胀率(FINF)、净资本流入(CIN)、国际原油价格(OIL)、百度词条搜索次数、恐慌指数(VIX)为短期影响因素。选取的研究区间为2005年至2015年,采用月度数据研究。其中经济增长率为了消除统计局季度调剂我们采用了同比增长率并采用三样条插值法进行插值,而投资率则通过全国固定资产完成额除以GDP得到,贸易条件为出口价格指数与进口价格指数的比值,经济开放度等于进出口额除以GDP,净资本流入为外汇储备增加额减去经常账户后的余额。数据来源为国家统计局、万德数据库、中国银行官网以及Google学术,百度词条搜索次数则由百度官网提供。

虽然影响汇率波动的因素分为长期与短期,但是长期因素对于汇率波动的影响是一个持续的过程,其内部的短期部分依旧对于即期汇率波动产生了影响,而短期影响因素中除了含有对于即期汇率的影响因素之外,也包含了影响远期汇率的长期影响部分,因此采用小波分析分离长期和短期因素中的长期影响部分与短期影响部分,并用两类影响因素的长期部分预测人民币汇率的长期部分,短期亦是如此。

db4小波具有不对称性、没有线性相位、对于时频局部化能力较强、支集宽度较大的优点,可使频域分辨率增加,因此选其来实现小波分析。

其公式为: $P(y) = \sum_{k=0}^{N-1} C_k^{N-1} y^k$, 其中, C_k^{N-1} 为二项式系数, 则:

$$|m_0(\omega)|^2 = \left(\cos^2 \frac{\omega}{2} \right)^N P \left(\sin^2 \frac{\omega}{2} \right) \quad (1)$$

其中: $m_0(\omega) = \frac{1}{\sqrt{2}} \sum_{k=0}^{2N-1} h_k e^{-ik\omega}$ 。

采用db4小波对长/短期影响因素以及人民币汇率进行小波分析,并绘制其分析前后的波形图对比,我们选取部分长期因素中的货币供应量(M2)、投资率(INN),短期因素中的国内利率、国际原油价格以及人民币兑美元汇率的分解结果进行展示,如图2~图6所示。

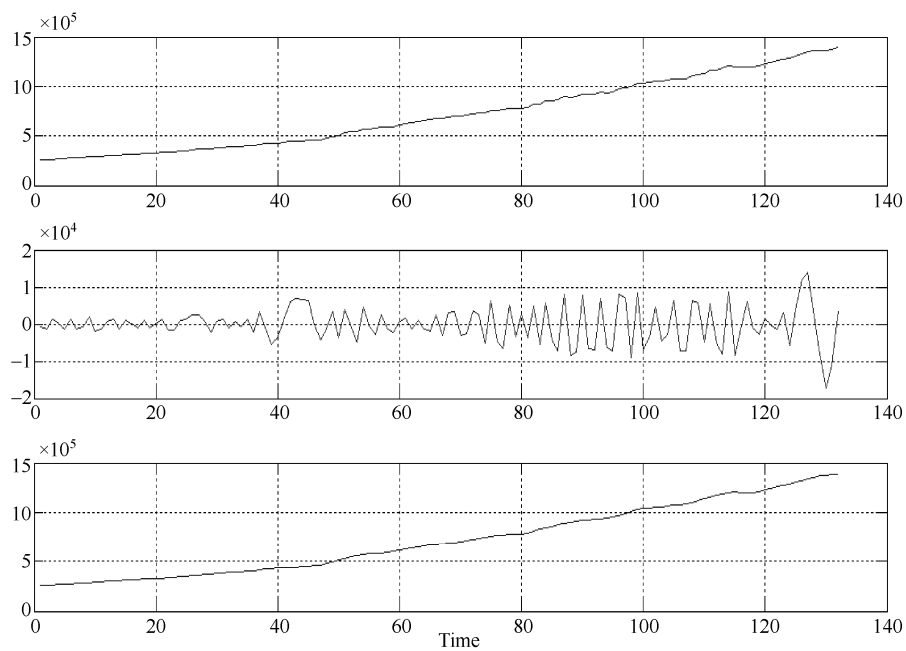


图2 货币供应量原波形与高低频分量对比

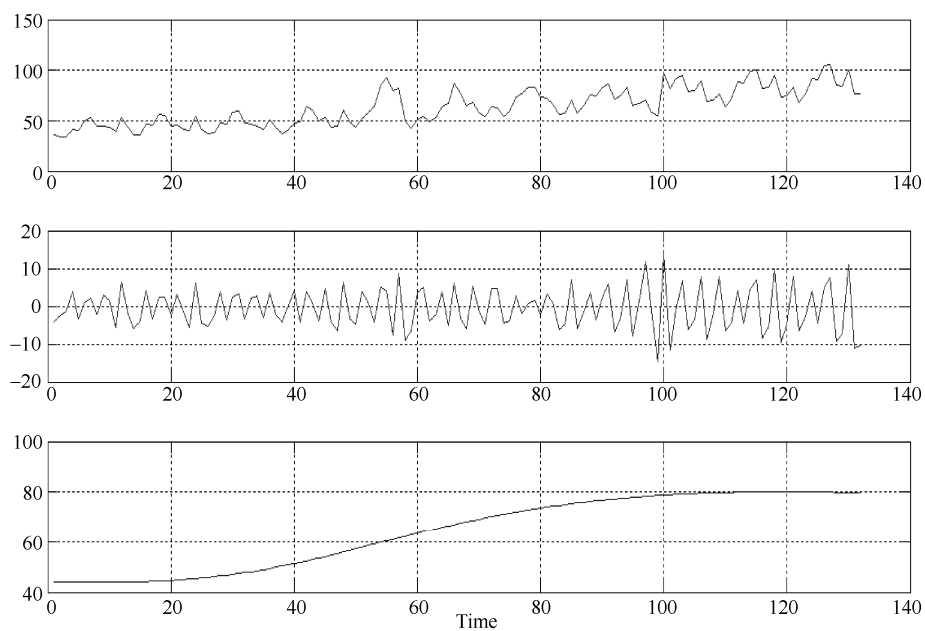


图3 投资率原波形与高低频分量对比

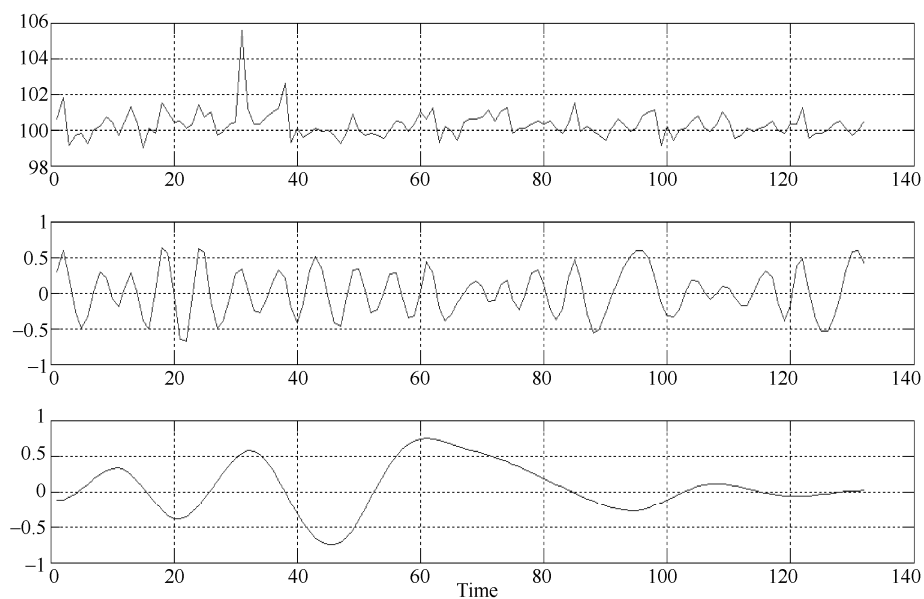


图4 国内利率原波形与高低频分量对比

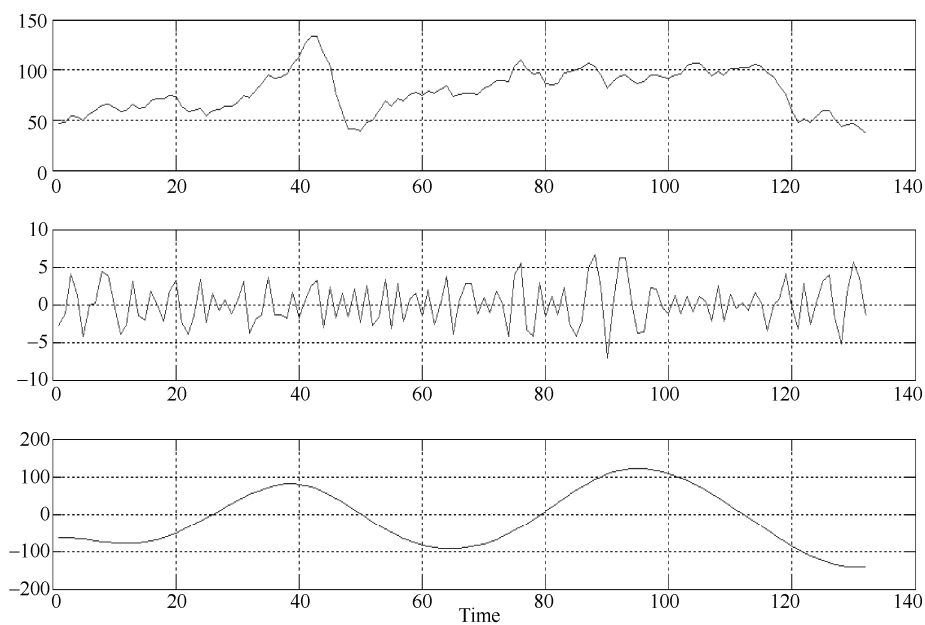


图5 国际原油价格原波形与高低频分量对比

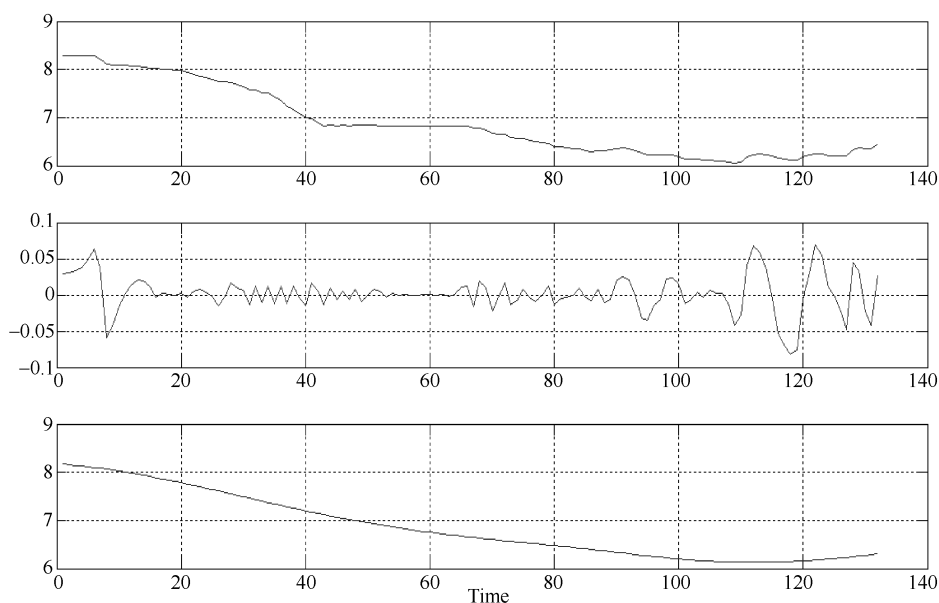


图 6 人民币汇率小波分析前后对比图

可以观察到各个原信号在原有的总体趋势之下，还有不同的波动幅度，而进行分解之后，短期影响部分的波动频率较大，变化较为频繁，而长期影响部分则变化得较为平缓，体现着一种总体趋势。接下来分别选取各个影响因素的短期部分来对汇率的短期部分进行预测，选取各个影响因素的长期部分对汇率的长期部分进行预测，最后结合所有影响因素对于汇率进行预测。

（二）短期神经网络模型

上述所有影响因素的短期影响部分共有 13 个指标，这对于神经网络的输入端影响太过冗杂，因此首先采用主成分分析对短期影响指标进行降维处理，主成分分析的碎石图如图 7 所示。

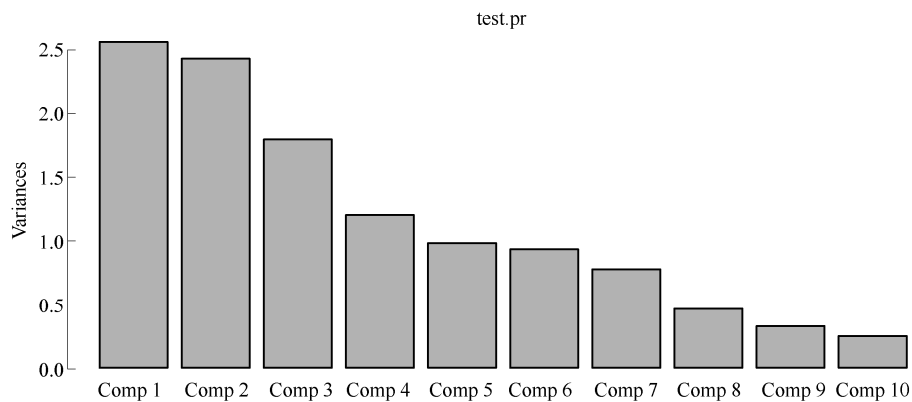


图 7 主成分分析因子方差贡献率

依据方差贡献率达 85% 的判断选取前 7 个主成分因子作为神经网络的输入端。在输入神经网络前, 需要对输入数据进行归一化处理, 公式如下:

$$y = (x - x_{\min}) / (x_{\max} - x_{\min}) \quad (2)$$

由于输入神经元个数为 7 个, 因此采用 3*3 的卷积核, 使用 6 层卷积层与池化层的嵌套, 加最后的一层全连接层的网络结构, 设置训练迭代次数 epochs 为 500, 均方差最小值 goal 为 0, 初始学习速率 mu 为 0.001, 学习速率减少系数 mu_dec 为 0.1, 学习速率增加系数 mu_inc 为 10, 最小梯度 min_grad 为 1e-07。选取 2004 年 12 月至 2014 年 12 月的数据为训练集, 选取 2015 年 1 月至 2015 年 12 月的数据作为测试集进行网络训练, 网络结构如图 8 所示。

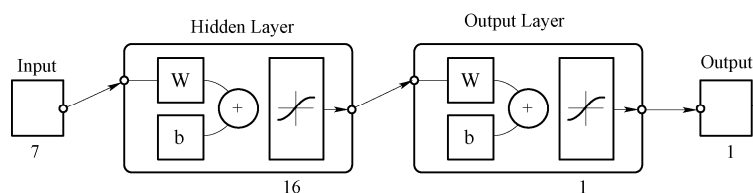


图 8 短期模型神经网络示意图

经训练之后的泛化误差对比图如图 9 所示。

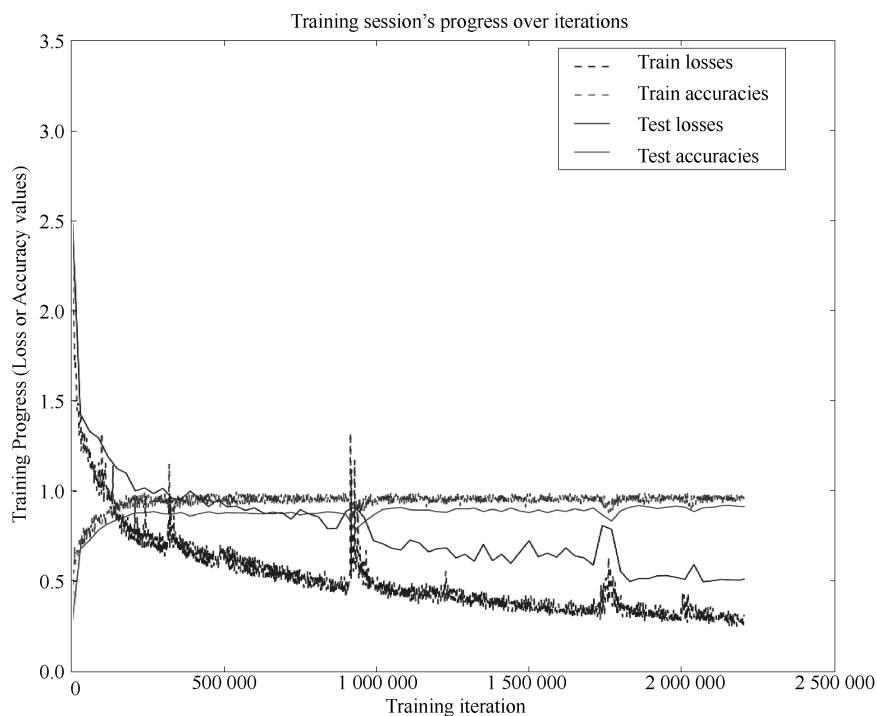


图 9 短期模型训练误差走势图

测试集的数据与真实值比较情况如表 3 所示。

表 3 短期模型测试集预测结果表

	反归一化输出值	真实数值	精确率
2015.01	1.060 6	1.061 5	99.91%
2015.02	1.062 5	1.064 4	99.82%
2015.03	1.063 6	1.066 3	99.75%
2015.04	1.063 8	1.067 1	99.69%
2015.05	1.064 2	1.067 1	99.73%
2015.06	1.051 3	1.066 2	98.60%
2015.07	1.050 4	1.064 1	98.71%
2015.08	1.045 1	1.060 6	98.54%
2015.09	1.031 4	1.055 7	97.69%
2015.10	1.024 5	1.49	97.64%
2015.11	1.010 7	1.040 4	97.15%
2015.12	1.001 4	1.029 7	97.25%
Overall	***	***	98.70%

表 3 我们选取所有影响因素的短期部分对于汇率的短期部分进行预测,可以发现前五个月的预测准确率均在 99%以上,而后 7 个月的预测精度则下降到了 98%~99%区间,这符合神经网络对于近期数据预测精度较高的设计结构,总体精度为 98.70%。可见 CNN 在一定程度上较好地拟合了汇率短期部分的波动,而根据我们之前的分析,指标百度搜索词条次数是一个短期影响因素(如图 10 所示),其指标内在波动反映市场参与者的投资预期与心态波动,所以其影响作用更多是对于因变量的短期部分,下面我们去除百度搜索词条次数的短期部分数据,仅用其余 12 个指标的短期部分对于汇率的短期部分进行预测。

经主成分分析之后选取 6 个主成分因子进行归一化处理输入神经网络,此网络同样采用 3*3 的卷积核,使用 8 层卷积层与池化层的嵌套,加最后的一层全连接层的网络结构,设置训练迭代次数 epochs 为 500,均方差最小值 goal 为 0,初始学习速率 mu 为 0.001,学习速率减少系数 mu_dec 为 0.1,学习速率增加系数 mu_inc 为 10,最小梯度 min_grad 为 1e-07。选取 2004 年 12 月至 2014 年 12 月的数据为训练集,选取 2015 年 1 月至 2015 年 12 月的数据作为测试集进行网络训练,得到的结果如表 4 所示。

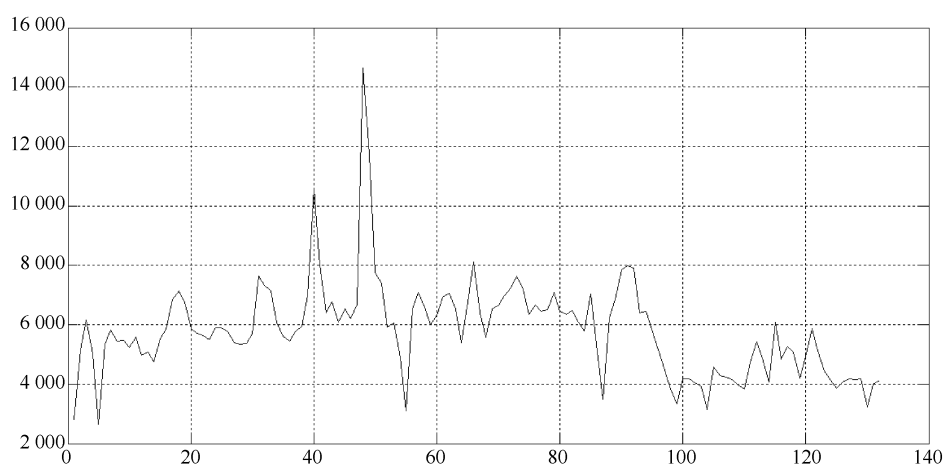


图 10 网络词条搜索量示意图

表 4 短期模型测试集预测结果表（不加入网络词条搜索量）

	反归一化输出值	真实值	精确率
2015.01	1.057 4	1.061 5	99.61%
2015.02	1.058 9	1.064 4	99.48%
2015.03	1.057 9	1.066 3	99.21%
2015.04	1.060 7	1.067 1	99.40%
2015.05	1.060 8	1.067 1	99.41%
2015.09	1.020 2	1.055 7	96.64%
2015.06	1.051 6	1.066 2	98.63%
2015.07	1.044 9	1.064 1	98.19%
2015.08	1.038 4	1.060 6	97.91%
2015.10	1.014 1	1.49	96.67%
2015.11	0.992 1	1.040 4	95.36%
2015.12	0.981 4	1.029 7	95.31%
Overall	***	***	97.99%

对比两次的拟合结果，没加入网络搜索次数时的泛化精度为 97.99%，而加入网络搜索次数的泛化精度为 98.70%。仔细观察每个月的预测精度可得，前五个月两个预测模型的预测精度均较高，当缺少网络搜索量指标时，波动较强的后六个月模型的预测精度为 96.67%，而当加入了网络搜索量之后，波动性较强的后六个月预测精度为 97.84%，可见网络搜索量指标对于解释因变量的短期成分具有一定的影响力，加入了网络词条搜索量之后对汇率波动的预测有所提高。随着新兴媒体的发展，来自媒体的信息指标对于经济指标的预测具有一定的作用，但是由于媒体的信息指标（诸如网络搜索次数）还有着一

些缺陷,例如网络搜索次数可能是因为一个人在短时间内反复搜索而增加的,或者由于经济课程的讲解需要导致学生前去搜索(但是并不会对汇率市场产生影响)所以并不能完全地解释指标波动。

而从另一方面观察,在短期的神经网络建模过程中,拟合精度已经达到 98.70%,虽然训练精度只有 98.91%,但是相当于舍弃了训练精度来达到泛化误差的减小,一定程度上减少了过拟合的现象。同时无论是否加入网络搜索量进行训练,前 5 个月的拟合精度均高于后期,这是因为该金融指标呈现非线性,而 CNN 神经网络对于非线性的数据的短期拟合效果较高,预测周期越长,其精度就越低。

(三) 长期影响部分的神经网络模型

长期影响因素不同于短期影响因素,其对于汇率波动的影响具有滞后期,由于无法确定滞后期的时长,于是使用逐期试验的方法,每次进行网络调试确定最优的隐含层节点与训练步数,并记录每次的训练结果,对比每次试验的训练误差,从而确定一个较优的滞后期来进行汇率预测。

选取 2004 年 12 月至 2014 年 12 月的数据为训练集,选取 2015 年 1 月至 2015 年 12 月的数据作为测试集进行网络训练,第 i 次的神经网络的输入端选取 $T-1$ 期至 $T-i$ 期的长期因素, $i=1, 2, \dots, 12$ 。先进行主成分分析降维,之后将主成分因子依次输入神经网络并调试出最优结果,如表 5 所示。

表 5 长期模型预测结果表

滞后期	输入训练集区间	主成分分析 输入变量	训练节点	训练步数	泛化误差
1	2004.12-2014.11	5	6	500	2.14%
2	2004.11-2014.11	10	12	597	1.98%
3	2004.10-2014.11	15	15	683	1.57%
4	2004.09-2014.11	20	21	798	1.16%
5	2004.08-2014.11	25	25	840	0.86%
6	2004.07-2014.11	30	34	900	1.71%
7	2004.06-2014.11	35	38	913	1.91%
8	2004.05-2014.11	40	44	930	2.16%
9	2004.04-2014.11	45	49	954	2.30%
10	2004.03-2014.11	50	53	985	2.41%
11	2004.02-2014.11	55	58	1 003	2.45%
12	2004.01-2014.11	60	62	1 023	2.44%

从表 5 观察可得，泛化误差呈现先降后升的趋势，在滞后期为 1–5 时，因为即期汇率受到前几个月的长期影响因素的影响，所以随着滞后期输入变量的增加，汇率波动得到了更好的解释，所以泛化误差随之下降，在滞后期为 5 时达到了最优。而在滞后期为 6–12 时，因为 5 个月前的长期影响因素没有足够的影响期影响到即期汇率，因此相对于神经网络而言，之后滞后期的输入变量是不符合汇率波动的干扰项。此外，随着滞后期的增加，主成分因子个数也随之增加，而训练集个数有限，不可避免出现了过拟合的现象，这也导致了泛化误差的上升。当滞后期为 5 时，泛化误差为 0.86%，优于短期模型中加入网络搜索量的误差 1.30%。所以可以认为长期滞后期为 5 的模型是优于短期模型的，其预测的结果如表 6 所示。

表 6 滞后期为 5 的长期模型预测结果表

	反归一化数值	真实值	精确率
2015.01	5.122 3	5.167 4	99.13%
2015.02	5.135 6	5.176 4	99.21%
2015.03	5.140 1	5.186 2	99.11%
2015.04	5.144 5	5.196 9	98.99%
2015.05	5.143 4	5.208 2	98.76%
2015.06	5.161 3	5.220 3	98.87%
2015.07	5.160 9	5.233 0	98.62%
2015.08	5.262 2	5.246 3	99.70%
2015.09	5.305 8	5.260 1	99.13%
2015.10	5.301 2	5.274 3	99.47%
2015.11	5.310 3	5.289 0	99.60%
2015.12	5.252 6	5.304 1	99.03%
Overall	***	***	99.14%

（四）两种汇率预测模型总结

经过上文的论述比较，由于长期影响部分的波动幅度较小，具有平缓的趋势，所以其模型的预测精度为 99.14%，而短期部分的波动大，预测难度大，因此模型的预测精度较低，为 98.7%。总体来看，CNN 神经网络对于长期影响部分的预测效果较好，而我们合并两个模型的预测结果，即预测所得的汇率短期部分与长期部分，将其相加得到总的预测汇率值如表 7 所示。

表 7 合并长短期预测结果表

	短期部分	长期部分	预测汇率	真实汇率	精确率
2015.01	1.057 4	5.122 3	6.182 9	6.218 1	99.43%
2015.02	1.058 9	5.135 6	6.198 1	6.251 8	99.14%
2015.03	1.057 9	5.140 1	6.203 7	6.238 6	99.44%
2015.04	1.060 7	5.144 5	6.208 3	6.201 0	99.88%
2015.05	1.060 8	5.143 4	6.207 6	6.203 5	99.93%
2015.06	1.051 6	5.161 3	6.212 6	6.205 2	99.88%
2015.07	1.044 9	5.160 9	6.211 3	6.208 5	99.95%
2015.08	1.038 4	5.262 2	6.307 3	6.338 3	99.51%
2015.09	1.020 2	5.305 8	6.337 2	6.367 6	99.52%
2015.10	1.014 1	5.301 2	6.326 6	6.350 5	99.62%
2015.11	0.992 1	5.310 3	6.321 0	6.364 0	99.32%
2015.12	0.981 4	5.252 6	6.254 0	6.449 1	96.97%
Overall	***	***	***	***	99.39%

由短期与长期影响部分预测的结果合并之后的预测精度为 99.39%，介于短期模型的预测精度与长期模型的预测精度之间。接下来我们选取所有影响因素直接对汇率进行预测，而后对比先进行短长期部分的分离、分别进行预测之后合并结果，与直接利用影响因素来进行预测的模型的精度高低，结果如表 8 所示。由表 8 结果观察可得，直接利用影响因素进行预测的结果为 99.07%，而由长短期分开进行预测的结果精度为 99.39%，这表明首先对汇率影响部分进行长短期分离，而后各自分别利用相应部分的因素进行预测，最后合并的预测效果更好。

表 8 直接分析预测结果表

	反归一化数值	真实值	精确率
2015.01	6.183 3	6.218 1	99.44%
2015.02	6.193 5	6.251 8	99.06%
2015.03	6.199 2	6.238 6	99.37%
2015.04	6.208 8	6.201 0	99.87%
2015.05	6.240 1	6.203 5	99.41%
2015.06	6.223 4	6.205 2	99.71%
2015.07	6.230 1	6.208 5	99.65%
2015.08	6.275 8	6.338 3	99.01%
2015.09	6.301 2	6.367 6	98.96%

续表

	反归一化数值	真实值	精确率
2015.10	6.294 5	6.350 5	99.11%
2015.11	6.302 1	6.364 0	99.03%
2015.12	6.203 0	6.449 1	96.18%
Overall	***	***	99.07%

(五) 与其他预测方法的结果对比

将本文的最终预测精度 99.39%与之前的参考文献中 BP 神经网络所得到的 97.3%和 98.2%的预测精度、径向基神经网络的精度 98.72%以及马尔科夫模型得到的预测精度 98.32%进行对比,表明基于深度学习的神经网络相较于之前传统的神经网络有着较优的预测性能。

由于汇率数据本身就是一个时间序列,因此采用时间序列的预测算法进行汇率预测也大有人在,其中毕玉江(2016)就采用了贝叶斯平均分类回归模型进行了汇率预测,李艳丽等(2016)以损失函数与 DM 检验为基准,证明了指数平滑模型具有优秀的波动预测能力。本文选取贝叶斯平均分类回归模型(BACRN)、指数平滑模型与广义自回归条件异方差模型(GARCH)与 CNN 模型的结果按照预测精度、均方根误差、平均绝对误差以及方向预测精度四个统计指标进行结果对比。

均方根误差(RMSE):

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (3)$$

平均绝对误差(MAE):

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (4)$$

方向预测精度(DA):

$$DA = \frac{1}{N} \sum_{i=1}^N a_i, a_i = \begin{cases} 1, & (y_{i+1} - y_i)(\hat{y}_{i+1} - y_i) > 0 \\ 0 & \end{cases} \quad (5)$$

由表 9 对比可得,四种方法中预测精度最高的是我们所选用的神经网络预测模型,同时,其均方根误差与平均绝对误差也是四种方法中最小的,但是在预测方向精度指标上,神经网络的表现较差,原因是神经网络在进行反向传播时判别的标准是均方误差最小,而没有去考虑目标增减的方向。就总体而言,CNN 模型的预测效果占优,且误差较小,优于传统的时间序列的预测方法。

表 9 四种方法对比

	精确率	RMSE	MAE	DA
BACRN	97.13%	0.076	0.042	0.818 2
指数平滑模型	97.89%	0.075	0.045	0.636 4
GARCH	99.03%	0.072	0.047	0.727 3
CNN	99.39%	0.063	0.039	0.636 4

六、结论

运用基于深度学习方法改善而提出的卷积神经网络模型,可以更好地预测人民币汇率变动。首先,我们利用小波分析把影响汇率波动的主要因素分为长期部分与短期部分,由此分别利用长期与短期两个预测模型进行分析,发现网络搜索次数等新兴媒体指标能够提高短期模型的预测结果,同时,合并长、短期模型的结果得到的预测精度为 99.39%,此结果优于直接利用所有影响因素进行预测的结果,也优于采用传统的神经网络以及贝叶斯平均分类回归模型(BACRN)、指数平滑模型与广义自回归条件异方差模型(GARCH)预测的结果。

本文的研究说明深度学习的新技术能更好地预测汇率波动,随着人民币国际化和汇率市场化,这些新技术将在汇率及其他资产价格预测中发挥不可忽视的作用。同时,加入新兴媒体指标来解释突变现象,可以更好地拟合汇率波动。对突发跳变的预测,是未来的一个研究方向。

参考文献

- [1] 毕玉江,王双成. 汇率预测方法及比较:基于贝叶斯平均分类模型的检验[J]. 统计与决策, 2016 (3).
- [2] 谷宇,高铁梅. 国际资本流动背景下人民币汇率的均衡水平及短期波动[J]. 金融研究, 2008, 5: 1-12.
- [3] 郭莹莹. 人民币汇率的长短期影响因素对比分析和汇率波动主导因素界定[J]. 科学决策, 2015 (3).
- [4] 黄志刚,郑国忠. 基于 GABP 神经网络汇率波动弹性空间测度[J]. 当代财经,2013(9):49- 60.
- [5] 兰秋军,马超群,文凤华. 金融时间序列去噪的小波变换方法[J]. 科技管理研究, 2004, 24 (6): 117-120.

- [6] 李艳丽, 邓贵川, 李辰阳. 人民币汇率波动的预测—基于损失函数与 DM 检验的比较分析[J]. 国际金融研究, 2016 (2).
- [7] 钱舒. 小波在股市数据分析中的应用[J]. 经济数学, 1999, 19 (4): 80–84.
- [8] 徐缘圆. BP 神经网络在汇率预测中的应用[J]. 时代金融旬刊, 2013 (1).
- [9] 杨帆. 人民币实际汇率研究[J]. 管理世界, 1999, 5:44–58.
- [10] 袁修贵, 侯木舟. 小波分析在证券分析中的应用[J]. 中南工业大学学报, 2002, 33 (1): 103–106.
- [11] CHAUDHURI T D, GHOSH I. Artificial neural network and time series modeling based approach to forecasting the exchange rate in a multivariate framework[R]. Working Paper,2016.
- [12] CLARK P B, MACDONALD R. Filtering the BEER: a permanent and transitory decomposition[R]. IMF Working Paper, 2010,144.
- [13] GIOQINANG ZH, MICHAEL HU. Neural network forecasting of the british Pound/US Dollar exchange rate[J]. Omega,1998,26(4):495–506.
- [14] HARVEY A C. Forecasting, structural time series model and the Kalman filter[M]. Cambridge: Cambridge University Press, 1999, 102–158.
- [15] KARFAKIS C. Is there an empirical link between the dollar price of the euro and the monetary fundamentals [J]. Applied Financial Economics, 2006 (16): 973–980.
- [16] KUAN C M, LIU T. Forecasting exchange rates using feed forward and recurrent neural networks[J]. Journal of Applied Econometrics, 1995,10(4): 347–64.
- [17] LEE J, FERRETTI M, RICCI L. Exchange rate assessments: CGER methodologies [M]. Washington, DC: International Monetary Fund,2008.
- [18] SUN Y. RMB exchange rate forecast approach based on BP neural network[J]. Physics Procedia, 2012, 33:287–293.
- [19] WILLIAMSON. The exchange rate system [M]. Institute for International Economics, MIT Press,1983.

Can Deep Learning Better Forecast RMB Exchange Rate?

ZhouYinggang¹ ChenHaipeng²

(1. *School of Economics, Xiamen University, Xiamen 361005, China;*

2. *School of Information Science and Engineering, Xiamen University, Xiamen 361005, China)*

Abstract: This paper aims to answer whether deep learning can better forecast RMB exchange rate fluctuations. We use the CNN neural networks, which are improved by deep learning, to construct long-term and short-term models of forecasting RMB exchange rate changes respectively. The result shows that the emerging media indicators of internetsearch can improve the short-term model's forecast. Meanwhile, combined forecasts of the long and short term models are better than those of the traditional neural network and time series analysis methods, such as GARCH and Bayesian mean classification regression models.

Keywords: Exchangerate Forecast, CNN Neural Network, Wavelet Analysis, Principal Component Analysis, Media Indicators