



Y3221222

单位代码: 10010

学 号: 2014200829



北京化工大学

硕士研究生学位论文

题 目 基于深度学习的外汇市场

预测与风险测度研究

专 业 工商管理

研 究 生 代文慧

指导教师 贺凯健

日 期: 2017 年 5 月 27 日

北京化工大学学位论文原创性声明

本人郑重声明： 所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不含任何其他个人或集体已经发表或撰写过的作品成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到本声明的法律结果由本人承担。

作者签名： 代文慧 日期： 2017.5.25

关于论文使用授权的说明

学位论文作者完全了解北京化工大学有关保留和使用学位论文的规定，即：研究生在校攻读学位期间论文工作的知识产权单位属北京化工大学。学校有权保留并向国家有关部门或机构送交论文的复印件和磁盘，允许学位论文被查阅和借阅；学校可以公布学位论文的全部或部分内容，可以允许采用影印、缩印或其它复制手段保存、汇编学位论文。

☐ 论文暂不公开（或保密）注释：本学位论文属于暂不公开（或保密）范围，在_____年解密后适用本授权书。

☒ 非暂不公开（或保密）论文注释：本学位论文不属于暂不公开（或保密）范围，适用本授权书。

作者签名： 代文慧 日期： 2017.5.25

导师签名： 贺庆 日期： 2017.5.27

学位论文数据集



中图分类号	F830.92	学科分类号	63040	
论文编号	1001020170829	密 级	非保密	
学位授予单位代码	10010	学位授予单位名称	北京化工大学	
作者姓名	代文慧	学 号	2014200829	
获学位专业名称	工商管理	获学位专业代码	1202	
课题来源	国家自然科学基金项目	研究方向	财务管理	
论文题目	基于深度学习的外汇市场预测与风险测度研究			
关 键 词	汇率预测, 汇率风险度量, Deep Belief Network (DBN) 模型, Long Short Term Mneory (LSTM) 模型, Value at Risk (VaR) 模型			
论文答辩日期	2017.05.18	* 论 文 类 型	基础研究	
学位论文评阅及答辩委员会情况				
	姓名	职称	工作单位	学科专长
指导教师	贺凯健	教授	湖南科技大学	金融工程、财务管理
评阅人1	张 文	教授	北京化工大学	管理科学与工程
评阅人2	周 波	副教授	对外经贸大学	财政与税务管理
评阅人3				
评阅人4				
评阅人5				
答辩委员会主席	吴 军	教授	北京化工大学	物流管理
答辩委员1	王 璇	副教授	北京化工大学	物流管理
答辩委员2	任继勤	副教授	北京化工大学	运营管理
答辩委员3	吴卫红	副教授	北京化工大学	项目管理
答辩委员4	李小燕	副教授	北京化工大学	财务管理
答辩委员5				

注: 一. 论文类型: 1. 基础研究 2. 应用研究 3. 开发研究 4. 其它

二. 中图分类号在《中国图书资料分类法》查询。

三. 学科分类号在中华人民共和国国家标准 (GB/T 13745-9) 《学科分类与代码》中查询。

四. 论文编号由单位代码和年份及学号的后四位组成。

基于深度学习的外汇市场预测与风险测度研究

摘 要

随着全球一体化的加速发展与国际贸易范围的不断扩大,国际金融市场上涉及贸易往来的资金规模不断增大,互联网与高性能计算机技术的发展也促进了跨国贸易的发展,它们不仅使得国际间的贸易往来越来越频繁,而且使得形式也越来越多样化,汇率作为国际间经济贸易活动的重要纽带之一,发挥着越来越重要的作用。

在人民币汇率改革之后,汇率形成机制的市场化特征不断增强,人民币汇率变动更加剧烈。近些年来,外汇市场受到东南亚金融危机,美国次贷危机以及欧债危机的影响,国际上主要货币的汇率走势依旧复杂难测,外汇市场风险不断增加,汇率预测与风险度量成为汇率研究领域的重要问题之一。同时,加强外汇市场汇率预测的准确度以及对汇率风险的管理对我国跨国企业经营活动十分重要。

深度学习模型自提出以来,在图像识别、语音识别、分类、数据挖掘等方面都有着较好的表现,在处理非线性问题方面有着较高的性能。汇率的历史数据序列呈现显著的非线性特征。目前的经典模型都无法对这其中的内在关系进行较好的描述以及特征提取,本文将深度学习模型引入到金融预测与风险度量领域,对汇率市场历史数据中的复杂非线性特征进行建模分析。

针对汇率走势的精确预测与汇率风险的精确估计问题,本文提出基

于深度学习理论的汇率预测与风险度量模型。首先，本文介绍了深度学习模型近年来在图像识别、语音识别等多个领域的研究现状，然后提出一种新的基于深度学习模型的外汇市场汇率预测模型。之后基于 VaR 风险度量理论，提出一种新的基于深度学习理论的集成 VaR 风险度量模型。该模型将深度学习模型与 ARMA-GARCH 模型相结合，构建了 VaR 估计模型。并且采用 7 个主流外汇市场的汇率数据进行实证研究，实证结果表明本研究中所提出的预测与风险度量模型具有较好的精确性以及可靠性。

关键字：汇率预测，汇率风险度量，Deep Belief Network (DBN) 模型，Long Short Term Memory (LSTM) 模型，Value at Risk (VaR) 模型

RESEARCH ON FORECASTING AND VALUE-AT-RISK MEASUREMENT IN FOREIGN EXCHANGE MARKETS BASED ON DEEP LEARNING

ABSTRACT

With the accelerated development of global integration and the continuous expansion of international trade, the scope of the funds for trade in international financial markets is growing. The development of Internet and high-performance computer technology have also promoted the development of cross-border trade activities, making International trade is becoming more and more frequent, and the forms are becoming more and more diversified. The exchange rate plays an increasingly important role as one of the important links of international economics and cross-border trade activities.

After the reform of the RMB exchange rate, the market-oriented characteristics of the exchange rate formation mechanism are increasing, and the exchange rate of the RMB changes more violently. In recent years, Southeast asian financial crisis, subprime crisis, the euro crisis have a serious negative impact on the foreign exchange market, the international exchange rate of major currencies is still complex and unpredictable, the risk of foreign exchange market continues to increase, exchange rate forecast and risk measurement into exchange rate research field One of the important research . At the same time, it is very important to strengthen the accuracy of exchange

rate forecast in foreign exchange market and the risk management of exchange rate to multinational enterprises in China.

The deep learning model has a good performance in image recognition, speech recognition, classification, data mining and so on, and has high performance in dealing with nonlinear problems. The historical data of the exchange rate also has a very complex linear and non-linear relationship, the current classic models can not be a good description of which the relationship between the intrinsic description and feature extraction, Therefore, it is a new attempt to describe the complex intrinsic function of exchange rate market historical data by applying the depth learning model to the field of financial forecasting.

Aiming at the problem of accurate forecasting of exchange rate movements and more accurate estimation of exchange rate risk, this paper proposes a model of exchange rate forecast and risk measurement based on deep learning theory. Firstly, this paper introduces the research status of the deep learning model in image recognition and speech recognition in recent years, and then proposes a new exchange rate forecasting model in foreign exchange market based on the deep learning model. Then, a new risk management model of integrated VaR is proposed based on the theory of risk measurement. This model combines the depth learning model with the ARMA-GARCH model to construct the VaR model. And exchange rate data of seven mainstream foreign exchange markets is used in the empirical study, the model proposed in this study has good accuracy and reliability.

KEYWORDS: Exchange rate forecast, Exchange rate risk measurement, Deep belief network model (DBN), Long short term memory model (LSTM), Value at Risk model (VaR)

目录

第一章 绪论	1
1.1 研究背景	1
1.2 研究目的及意义	2
1.3 国内外研究综述	3
1.3.1 外汇市场预测的研究综述	3
1.3.2 外汇市场风险度量研究综述	6
1.3.3 VaR 风险度量方法的研究综述	7
1.3.4 集成算法的研究综述	10
1.3.5 深度学习方法的研究综述	11
1.4 本文研究内容以及创新点	14
第二章 深度学习理论	17
2.1 神经网络	18
2.1.1 神经元	18
2.1.2 激活函数	18
2.2 深度学习模型	20
2.2.1 深度置信网络 (Deep Belief Networks, DBNs)	20
2.2.2 深度多层感知器	23
2.2.3 循环神经网络 (Recurrent Neural Networks, RNNs)	24
2.3 优化算法	28
2.3.1 共轭梯度法	28
2.3.2 随机梯度下降算法	29
2.4 本章小结	30
第三章 基于深度学习的外汇市场汇率预测模型	31
3.1 引言	31
3.2 基于深度学习的外汇市场汇率预测模型	31
3.2.1 数据处理	31

3.2.2 预测模型结确定 32

3.2.3 基于深度学习的集成汇率预测 33

3.3 实证研究 33

3.3.1 实验数据 33

3.3.2 实证结果分析 34

3.4 本章小结 36

第四章 基于深度学习的集成 VaR 风险测度模型37

4.1 引言 37

4.2 VaR 风险度量理论 37

4.3 基于深度学习的 VaR 风险测度模型 39

4.3.1 数据处理 39

4.3.2 分风险估计 40

4.3.3 总风险集成 41

4.4 实证研究 42

4.4.1 实验数据 42

4.4.2 评价指标 43

4.4.3 实证结果分析 43

4.5 本章小结 46

第五章 总结和展望47

参考文献49

致谢 错误!未定义书签。

作者和导师简介 错误!未定义书签。

Contents

Chapter 1 Introduction.....	1
1.1 The background on research	1
1.2 Research purposes and significance	2
1.3 Literature review at home and abroad	3
1.3.1 Literature review of forecasting in foreign exchange market	3
1.3.2 Literature review of risk management in foreign exchange market.....	6
1.3.3 Literature review of VaR	7
1.3.4 Literature review of ensemble algorithm	10
1.3.5 Literature review of deep learning	11
1.4 The innovations of this thesis	14
Chapter 2 Deep learning theory	17
2.1 Neural networks	18
2.1.1 Neuronal structure	18
2.1.2 Activate function	18
2.2 Models based on deep learning	20
2.2.1 Deep Belief Networks(DBNs)	20
2.2.2 Deep neural networks (DNN)	23
2.2.3 Recurrent Neural Networks (RNNs)	24
2.3 Optimization algorithm	28
2.3.1 Conjugate Gradient Algorithm (CG)	28
2.3.2 Stochastic Gradient Descent Algorithm (SGD)	29
2.4 Summary	30
Chapter 3 Forecasting of exchange rate in foreign exchange market based on Deep learning model.....	31
3.1 Introduction	31

3.2 Establishment the forecasting of exchange rate in foreign exchange market based on Deep learning models	31
3.2.1 Data for empirical analysis	31
3.2.2 The Construction of Deep Learning Models	32
3.2.3 Forecasting of exchange rate in foreign exchange markets	33
3.3 Empirical analysis	33
3.3.1 Data Collection and Statistical Analysis	33
3.3.2 Results Discussion.....	34
3.4 Summary	36

Chapter 4 VaR measurement study based on deep learning model.37

4.1 Introduction	37
4.2 VaR Theory	37
4.3 Establishment the VaR model based on deep learning.....	39
4.3.1 Data for empirical analysis	39
4.3.2 Risk estimation of IMF.....	40
4.3.3 Test of VaR model	41
4.4 Empirical analysis	42
4.4.1 Data Collection and Statistical Analysis	42
4.4.2 Performance Evaluation Criteria for the Model	43
4.4.3 Results Discussion.....	43
4.5 Summary	46

Charpter 5 Conclusions and Prospect47

Referance49

Acknowledgements 错误!未定义书签。

About Author and Tutor61

第一章 绪论

1.1 研究背景

随着全球一体化的加速发展与国际贸易范围的不断扩大,国际金融市场上涉及贸易往来的资金规模不断增大。互联网与高性能计算技术的发展也促进了跨国贸易的发展,国际间的贸易往来越来越频繁,形式越来越多样化。汇率作为连接国际间经济贸易活动的重要纽带之一,发挥着越来越重要的作用^[1]。在经济贸易全球化的进程中,汇率对国民经济的影响不断加深。汇率不仅可以影响宏观经济,调节国际收支平衡,而且也是影响金融市场中的投资分析、衍生证券定价以及风险管理的重要变量,所以外汇市场中汇率的变动趋势是重要的研究问题^[2]。

如果一个国家需要用本国的货币来换取得到另一个国家的本国货币,汇率作为纽带成为两个国家货币之间互相表示价格的中间量^[3]。在国际上汇率有直接标价法以及间接标价法两种标价方法。直接标价法又可以称为应付标价法,是一个国家若要用本国货币换取一定单位数量的外国货币时,本国货币应付的金额,这也可理解为是购买一定单位数量的外国货币时所支付的本国货币的金额。世界上大多数国家采取的汇率标价方法都是直接标价法。间接标价法又可以称为应收标价法,是一个国家若用一定单位数量的本国货币收取外国货币时,可以换取到外国货币的金额。主流国际汇率市场中的欧元市场、英镑市场、澳元市场应用的是间接标价法^[4]。

上个世纪 70 年代开始,布雷顿森林固定汇率体系开始逐渐崩塌,全世界各个国家开始纷纷放弃与美元相关的固定汇率制度,选择实施浮动汇率制度,汇率的波动变得频繁而且没有规律,国际上的主流货币的汇率浮动幅度不仅影响着国家的经济情况,也因此相应的产生了汇率风险^[5]。2005 年 7 月以后,中国不再单一的关注美元,选择在市场需求的基础上实施浮动汇率制度,汇率制度的改变使得汇率的变动变得富有弹性化以及充满了不可预测的特性。在汇率制度改革之前,人民币汇率处于一个相对稳定的状态,但是在实施管理下的浮动汇率制度以后,人民币对美元汇率在 2007 年的时候降低至 7.8,之后在 2008 年的时候又再次降低,此时降低至 7.0,随后人民币对美元汇率仍有下降趋势,直至突破了 6.0^[6]。在人民币汇率改革之后,市场对汇率的影响与之前相比变得更加的重要,国家对汇率市场进行干预的程度逐渐减弱,人民币汇率的浮动趋势变得更加不可预测。近些年来,外汇市场受到东南亚金融危机,美国次贷危机以及欧债危机的影响,国际上主要货币的汇率走势仍然多变难测,外汇市场风险不断增加,汇率预测与风险度量成为汇率研究领域的重要问题之一。

同时,自从我国在 2001 年加入世贸组织以后,在国际上进行经济贸易活动的比重开始逐渐增加,我国的汇率也会随着外汇市场主流国际汇率的波动而波动。在全球

一体化加速发展的背景之下，越来越多的企业开始将合作范围扩大到国际市场中，与国外企业开展各种各样的合作方式，与此同时也面临着多种多样的竞争方式。随着我国经济的高速发展，以及我国在国际市场中的地位的提高，人民币也逐渐成为了一种国际主流货币，由此其他国家开始要求人民币升值，这使得我国的外汇储备以及外债都将受到外汇市场汇率波动的影响^[7]。如果人民币开始升值，我国所持有的所有外汇储备将会迅速贬值。与其他国家相比，我国的外汇市场起步较晚，在外汇预测、外汇风险度量等方面以及对风险的防范和管理还不完善，加强外汇市场汇率预测的准确度以及对汇率风险管理对我国跨国企业经营活动十分重要。

近几年来，全世界很多学者选择将研究重心放在外汇市场的风险度量上，他们使用多种测量方法对其进行测量。VaR 方法可以考虑到多种市场因素并且对非线性的风险也有着较好的度量效果，逐渐成为了对风险进行度量的主要工具之一。对我国外汇市场进行汇率预测以及风险度量的研究既有理论价值，又有实际意义。

1.2 研究目的及意义

随着全球一体化以及贸易形式多样化的发展，汇率对国家经济以及国际贸易都有着不可替代的影响。汇率会因为受到货币政策、利率水平、通货膨胀、经济增长率以及国家的经济状况等多种因素的影响而变动^[8]。虽然全球的经济情况正在逐步恢复，金融危机所带来的负面影响程度得到了一定程度的缓解，但是随着经济格局的变化，国际间主要货币的汇率浮动，又加上多个国家要求人民币升值，在这种形势之下，我国的跨国企业需要面对更加严峻的汇率风险。因此正确的认识外汇市场汇率的波动，并加强外汇市场汇率预测的准确度以及对汇率风险的管理成为外汇市场研究的一个重要课题。

汇率风险主要分为三种风险，分别是交易风险（Transaction Exposure）、折算风险（Accounting Exposure）以及经济风险（Economic Exposure.），汇率风险中的交易风险是指从交易发生到结束清算期间，汇率不可预测的波动带来的交易损失。交易风险主要是受到在进出口货物或者国际间劳务交换的过程中，从合同签订的那天起到款项结清当日，外汇市场中汇率发生变化时带来的影响，除此之外，交易风险还受到在国际债权债务还清之前外汇市场汇率发生波动所带来的影响，以及受到国际银行在外汇买卖过程中，持有外汇头寸的空头或者多头时，外汇市场汇率发生波动所带来的影响^[9]。为了及时、有效地掌握并且控制外汇市场汇率波动所带来的风险，学者们需要加强对汇率预测的精准度。过度的汇率波动会给国际资产的价值、国债的利率和政策、国际间的债权债务清算的结果等多方面带来负面的影响^[10]。外汇市场中汇率预测的精准度，影响着国际贸易过程中的交易方式以及决策，除此之外，汇率的波动幅度决定

着中央银行利率的高低,所以汇率预测的精准度还影响着银行对于利率的决策^[11]。基于这几方面原因,如何加强对外汇汇率市场的预测精准度以及风险度量的可靠性成为了我国学者关注和研究的重要课题之一。

从历来学者们在汇率方面的研究中可以发现,汇率市场其实是一个多变量非线性系统,所以线性预测方法等传统方法并不能够对汇率市场历史数据之间复杂的非线性关系进行完全的描述^[12]。汇率市场并不是一个相对稳定的市场,有很多因素可以影响到汇率的波动,所以在实际研究中,由于学者们可能无法对汇率预测模型的输入变量进行精确的定义,传统经典的汇率预测模型的性能并不是十分的出色,在这种情况下,学者们也需要考虑将一些新的方法应用到汇率的研究中^[13]。

深度学习模型是继神经网络之后人工智能领域的新兴模型,在图像识别、文本分类、语音识别等方向有着持续的热度,但是目前将深度学习模型应用到金融领域对外汇市场进行汇率预测与风险度量的研究较少,本文的研究可以对外汇市场的汇率预测以及风险度量研究进行文献的补充。与此同时,通过应用深度学习模型能够更完善的描述我国及国际主流汇率市场的汇率波动情况并且加强外汇市场汇率预测的精准度。除此之外,本文将深度学习模型应用到 VaR 风险度量中,通过使用深度学习模型与 ARMA-GARCH 模型相结合的方法来对集成的 VaR 值进行估算,从而提高 VaR 估算的精准度与可靠性。所以,深度学习模型对我国外汇市场的汇率预测以及风险度量研究既有着理论价值,又有着实际意义。

1.3 国内外研究综述

1.3.1 外汇市场预测的研究综述

Deboeck G J (1994) 提出不仅仅是外汇市场,而且几乎所有的金融市场都是动态的市场,金融市场复杂多变,而且这些金融市场都具有着高噪声、非平稳的特质^[14]。这也就是说外汇市场是非线性变化的动态市场,历史数据序列的均值以及方差具有不平稳的特征,而且历史数据序列之间的内在关系也是动态变化的。所以在这种情况下,对外汇市场中的时间序列进行建模并且进行预测对学者们来说是一个比较难的研究课题。Yaser S A M 等学者(1996)认为外汇市场汇率预测是时间序列预测中较为有难度的研究方向^[15]。外汇市场出现以来,外汇市场中的汇率预测对各国学者们来说就成了一个研究的热点课题。在目前得到应用的预测方法中,阈值自回归模型(Threshold Autoregressive, TAR)、差分自回归移动平均模型(Autoregressive Integrated Moving Average Model, ARIMA)、指数自回归模型(Exponential Autoregressive)、多项式自回归模型(Polynomial Regression)等线性的概率统计模型得到了应用,随着研究的发

展,学者们提出了灰色预测、神经网络等新方法,在这些方法中,神经网络重点受到了学者们的关注,并且与其他方法相比它可应用的范围更加广泛。

在 Meese R A 等学者(1983)通过实证研究后得到了与监督随机游走模型相比,多个汇率决定的结构模型在样本外预测的精准度并不高的结论后^[16],学者们在此基础上受到启发,考虑到汇率市场中可能包含了不能被线性结构模型充分描述和解释的非线性成分,在这种情况下,学者们将已知的非线性结构的计量参数模型应用到汇率预测方面,来寻找汇率历史数据序列可能存在的非线性特征的结构类型。Engle R (1982) 以及 Bollerslev T (1986) 等学者们都得出了汇率历史数据序列中例如条件异方差特性等多种特性都可以用 ARCH 类模型来进行相关描述的结论^{[17][18]}。Hall S G 等学者(1989)也同样指出了可以用 ARCH 类模型对汇率数据序列中的条件异方差特性进行描述^[19]。Engle R (1983), Domowitz I 等学者(1985), Hsieh D A (1989), 以及 Hogan K 等学者(1994)等提出了用 ARCH 模型以及 GARCH 模型来对汇率时间序列数据中的非线性依赖性进行解释,实证结果显示这种方法得出了较好的结果^[20-23]。Nelson D B (1991), Glosten L R 等学者(1993)以及 Zakoian J M (1994)通过实证研究得出了 EGARCH 模型和 TARCH 模型能够较好地对比杠杆效应所带来的汇率非对称波动的现象进行描述和解释的结论^[24-26]。Obstfeld M 等学者(1997)通过阈值自回归模型(TAR)对多个国家的汇率历史数据序列内部间存在的非线性特征进行相关描述,并且将阈值自回归模型(TAR)与 AR 模型进行性能的比较,通过这两个模型的实证结果结果来判断 PPP 短期是否成立^[27]。Taylor M P 等学者(2001)以及 Liew V K S 等学者(2003)提出了不同汇率数据序列中的非线性均值调整的过程可以通过自激励阈值自回归模型(STAR)模型来捕捉,实证结果显示线性 AR 模型在汇率历史数据序列特征的描述方面的性能不是太好^[28-29]。Furao Shen 等学者(2015)构建了改进的深度置信网络模型对汇率市场进行预测,并用共轭梯度法对模型的学习速率进行加速,通过对英镑对美元的汇率市场进行实证分析,实证结果显示该模型与传统的预测模型相比性能更好^[30]。Georgios Chortareas 等学者(2011)对欧元汇率日数据进行波动预测,学者们选择使用波动率模型,实证结果显示与传统的模型相比,如果在预测时也考虑高频数据以及长期记忆维度,则预测结果的准确度会更高^[31]。Tomasz Korol (2014)在汇率预测方面构建了一个基于模糊逻辑理论的模型,并用日元对美元,英镑对美元和瑞郎对美元的历史数据进行实证研究,实证结果显示这种方法对汇率预测参数中的预测速率有所改善^[32]。Imad Moosa 等学者(2014)选用随机游走模型对美元、日元、英镑和加拿大美元这四个汇率市场进行汇率预测,实证结果显示在盈利能力方面,传统的预测模型性能比随机游走模型的性能好,但是在均方差误差估计方面,随机游走模型性能比传统预测模型的性能好^[33]。Soojin Yun 等学者(2015)用连续跳跃的分解方法来预测韩元对美元回升后的实现差异率,实证分析结

果显示这种新的分析方法比现有传统的跳跃分解方法有着更好的预测能力^[34]。Michele Ca' Zorzi 等学者（2015）将基于 Dornbusch 方法的贝叶斯 VAR 模型与随机游走模型在汇率预测方面进行了模型性能的比较，学者们提供了一种新的方法来对汇率市场进行汇率预测^[35]。Georgios Sermpinis 等学者（2015）构建了混合滚动遗传算法与支持向量机相结合的模型（RG-SVR）来对最优参数进行选择，对欧元对美元、欧元对英镑和欧元对日元进行汇率预测，与传统预测模型相比，新模型性能优于传统模型的性能^[36]。Jozef Barunik 等学者（2016）与 Georgios Chortareas 等学者（2011）的研究有些相似，同样使用了高频数据，学者们建立了基于 GARCH 模型的多时频分解实现波动率的预测模型来对不同时频对波动率预测的影响进行研究，同时将几种常用的波动预测模型的性能进行比较^[37-38]。Salim Lahmiri（2017）在人工智能神经网络有限的输入值的基础上，在不同的假设条件下比较了 GARCH 模型、GARCH 模型与 BPNN 模型结合的新模型性能，并构建了一个从历史数据中提取新指标作为 BPNN 模型输入值的新模型来对美元对加元以及美元对欧元进行历史波动的预测，与传统的预测模型相比，学者们所构建的新模型更加简单有效^[39]。

随着人工智能计算技术的飞速发展，在总结国外学者们的研究进展后，国内的学者们受到了一定的启发，提出了在汇率市场中应用神经网络模型的新的尝试办法，对历史数据具有非线性特征的汇率市场进行描述和预测。

王松喜（2005）考虑到 EGARCH 模型等基于计量经济学的预测模型在汇率短期预测方面的效果良好，将基于证券技术分析的指标与基于计量经济学的预测模型相结合，对汇率市场进行预测，实证结果显示这种方法提高了外汇交易中获得收益的概率^[40]。余乐安等学者（2005）构建了一个基于 BP 神经网络的外汇市场汇率预测的子系统，又构建了一个基于 Web 的外汇交易决策支持的子系统，然后用集成方法将这两个系统集成起来，构建了一个基于集成方法的外汇市场汇率预测与外汇交易决策支持系统（Integrated forexrolling fore-casting and trading decision support system, IFRFTDSS）^[41]。汪寿阳等学者（2007）构建了一个基于 TEI@I 方法论的外汇市场汇率预测模型，外汇市场汇率的趋势用传统的经济计量模型处理，外汇汇率市场中历史数据的非线性特征用人工神经网络描述，外汇市场中的突发性以及不稳定性则用文本挖掘和专家系统处理^[42]。李艳丽等学者（2016）将指数平滑模型、ARCH 模型和 GARCH 模型，应用到人民币汇率波动拟合和样本外预测的研究中，并在研究中应用了动态时间滚动窗口技术，通过损失函数和 DM 检验等验证方法比较了所应用的模型对人民币汇率波动的样本外预测的性能^[43]。宋博等学者（2017）通过使用 HP 滤波法将外汇市场中的汇率时间序列数据分解为趋势成分序列和波动成分序列两个序列，然后使用自回归 ARMA-GARCH 模型分别对趋势成分序列和波动成分序列这两个成分序列进行拟合和预测，实证结果显示 ARMA-GARCH 模型在这方面的性能较好^[44]。

1.3.2 外汇市场风险度量研究综述

外汇市场的收益率序列具有着非线性的特征,而且外汇市场中汇率的历史数据序列有着偏态以及厚尾等特征,所以非正态分布假设下的投资组合风险度量方法成为了主要方法之一,常应用于构建投资组合模型以及风险度量。

汇率风险管理是企业运营过程中必须掌握的一部分,这也是保障企业稳定运营的手段之一。目前对外汇市场中的汇率风险进行专门的风险估算的研究还较少,但是对外汇市场中汇率收益率序列的波动进行相关分析的研究就比较多。考虑到组成金融市场风险的因素之中就包括着汇率自身,而且外汇市场中汇率的波动与其他金融市场中的历史数据序列的波动类似,所以将金融市场中对风险进行度量的研究方法应用到外汇市场中,用类似的方法对汇率的波动进行研究具有一定的可行性。

在金融时间序列的波动性研究方面,大量实证研究显示金融时间序列数据的波动特性可以用 GARCH (1, 1)、GARCH (1, 2)、GARCH (2, 1) 等模型来描述。考虑到其他金融市场有着不同的金融时间序列,学者们构建了基于 GARCH 模型的单整 GARCH 模型 (Integrated GARCH) [45]、指数 GARCH 模型 (Exponential GARCH) [46] 以及分整 GARCH 模型 (Fractional Integrated GARCH) [47] 等多个模型。

国外这方面的研究相比于国内来说更早,Adler.M 等学者 (1984) 创建了一种从公司价值角度衡量外汇风险的方法 [48]。Sklar A (1959) 在 1959 年提出了 *coupula* 理论,20 世纪 90 年代后 *coupula* 理论在国外取得了飞速的发展 [49]。Levonian (1994) 在针对传统方法进行较为全面的分析之后,总结之前学者们的研究成果,提出了 WAP 法 (weighted aggregate Position) 这种新的计量外币敞口的方法 [50]。Embrechts P 等学者 (1999) 将 Copula 函数应用到了金融研究领域 [51]。Rockinger M 等学者 (2006) 和 Palaro H P 等学者 (2006) 在 GARCH 模型中应用到了 Copula 函数,通过两者结合的方法在金融市场中估算了投资组合的 VaR 值 [52-53]。Patton (2001) 使用两元时变 Copula 方法对日元和马克兑换美元汇率的波动关联性的情况进行了研究 [54]。

学者们应用 GARCH 族模型对外汇市场中的汇率风险进行度量的研究中,国内有许多学者则针对人民币汇率的波动性相关问题进行了这方面的研究。惠晓峰等学者 (2003) 使用 GARCH (1, 1) 模型对人民币对美元的日汇率数据进行了实证研究,实证数据的时间跨度为 1994 年至 1997 年,通过实证研究得到了效果较好的参数估计值。并对实证结果进行相关检验从而得到了人民币对美元汇率确实存在条件异方差性 (ARCH 效应) 的结论 [55]。苏岩等学者 (2007) 将人民币对日元汇率的历史数据作为实证数据,对汇率的波动进行了实证研究,对多种 GARCH 族模型进行性能比较。实证结果显示 GARCH (1, 1) 模型是外汇市场中汇率方面研究的第一选择 [56]。

王德全（2009）应用 GARCH 模型对外汇市场中的四种汇率进行了实证研究，实证数据包括港币、欧元、美元、日元这四种主流货币对人民币的汇率历史数据，实证结果表明这几种汇率的波动率序列都有着非正态平稳的特征，而且也都有着较为突出的 ARCH 效应^[57]。骆珣等学者（2009）使用 GARCH 模型对美元对人民币汇率的日数据进行实证分析，实证数据的时间跨度为 2003 年至 2007 年，实证结果显示 GARCH 模型对汇率历史数据的拟合效果较好^[58]。翟爱梅（2010）基于 GARCH 模型构建了一个 GARCH 族模型，在对汇改以来的人民币汇率波动研究方面应用了人民币对美元汇率中间价的日数据，实证数据的实际跨度为 2005 年至 2009 年。实证结果显示汇率体制改革之后，汇率历史数据的收益序列有着左厚尾的特征，而且汇率历史数据的波动同时有着集聚性的特征^[59]。李志斌等学者（2010）在人民币汇率波动特征研究方面使用了 ARCH 族模型，通过实证分析得出了人民币汇率历史数据序列有着尖峰、厚尾、波动率聚类的特征，而且导致人民币升值的因素会比导致人民币贬值的因素使人民币汇率出现更大程度的波动的相关结论^[60]。曲映（2011）在人民币对美元汇率波动的集群性研究方面使用了 GARCH 模型，实证结果显示 GARCH（1，1）模型在回归方程残差项的异方差性方面的消除效果较好，与此同时 GARCH（1，1）模型具有良好的预测精度^[61]。刘瑾等学者（2008）使用多个 ARCH 类模型对美元对人民币汇率波动风险进行度量，通过比较实证结果发现 TARCH-M（1，1）模型存在过高风险估计^[62]。徐国祥（2016）通过实证研究对四种外汇资产的动态 VaR 值以及有效性进行分析^[63]。徐国祥等学者（2013）针对高置信度下美元、日元、欧元、韩元等多元资产组合期末风险难以精确度量的问题构建了 Copula-DCC-EVT 模型。学者们将 2007 年至 2011 年间我国的外汇资产组合风险作为研究对象，根据 Copula-DCC-EVT 模型在 700 个交易日的样本外滚动回溯测试检验期内计算期末 VaR 值，研究结果对于政府管理层与决策者掌握金融风险有着较好的应用价值^[64]。李鹏举（2016）应用美元、英镑、欧元、日元以及港币这五种货币的日收益数据，构建了基于 AR（2）-GARCH（1，1）-t 模型的各个外汇日收益率的边际分布，对应用正态 Copula 函数以及 t-Copula 函数的投资组合的 VaR 值用蒙特卡洛模拟方法进行估算，实证结果显示这种方法对外汇市场资产风险问题可以提供可操作的管理决策^[65]。

1.3.3 VaR 风险度量方法的研究综述

VaR（Value-at-Risk）是风险估值模型的简称，是近几年来国外兴起的一种度量金融风险的工具。这个方法最早应用在银行内部的资本风险管理中，但是现在已经得到金融领域国内外学者们的重视和应用^[66]。

投资组合管理决策理论是 1952 年 Markowitz 提出的相关理论，投资组合选择模

型基于均值-方差模型的研究比较多,但是这种模型对方差的度量并不是十分准确,这种模型并不十分完善。基于 VaR 方法的风险度量模型是 Baumol 在 1963 年提出的一种新模型。随着 VaR 理论的提出与这种方法逐渐被广泛应用,针对投资组合风险度量方面的研究在理论以及实证研究方面都取得了比较大的进展。

JP Morgan (1996) 提出了 Riskmetrics 体系,对估算 VaR 的方法进行了较为详细的阐述。在这之后, Jorion (1996) 提出在市场不发生异常变化的情况下,在一定的持有期限内和某一个置信水平下,资产或者投资组合可能发生的最大的损失值^[67]。

Engle (1982) 首次提出了 ARCH 模型,在这个模型被提出来之后,学者们在英国通货膨胀指数的波动性研究中应用了这种模型,实证结果显示这种模型的性能较好^[68]。随着 ARCH 模型快速的发展, Bollerslev (1986) 在这个基础上对模型进行了拓展,提出了 GARCH 模型。通过这个模型对收益率序列波动的非对称性进行衡量,在 GARCH 模型的基础上而得到了很多基于 ARCH 模型的波动性研究成果^[69]。Glosten 等学者 (1993) 在条件方差如何对收益率进行影响的研究方面,应用了多种 GARCH 模型,并且将所用到的模型进行对比^[70]。Z R Wang 等学者 (2010) 构建了 GARCH-EVT-Copula 模型对我国外汇资产组合的 VaR 值进行计算,但模型仅用到了较为简单的 Copula 函数,并且没有提供 VaR 值估计有效性的检验结果^[71]。Youssef M 等学者 (2015) 对能源商品的 VaR 值和 ES 值进行估算的研究中应用了三种 GARCH 模型,实证结果显示 DISPATCH-EVT 模型能够很好的拟合在不同时间跨度下,短期和长期的交易头寸能源商品的波动率^[72]。

Nelson (1991) 在前人研究的基础上提出了 EGARCH 模型,并在研究中引入了条件方差的对数形式,也就是说条件的预测值的非负形式就是杠杆效应的指数^[73]。Glasserman (1997) 提出风险测度的有效的模型中的 Monte Carlo 模拟方法^[74]。Alexander 等学者 (2002) 对历史模拟法进行了研究,实证结果显示投资组合中方差越低则出现风险的概率越大^[75]。David (2003) 在估算 VaR 值时使用了半参数方法,并对在极端情况下的 VaR 值的估算方法进行了研究^[76]。Artzner (1997) 提出虽然 VaR 方法简单并且易操作,但是在应用中却存在约束与局限性的结论^[77]。Rockafellar 等学者 (1999) 对现有的 VaR 估算方法进行了优化,针对 VaR 的局限性提出了 CVaR 的概念,构建了 CVaR 投资组合优化模型,实证结果显示 CVaR 方法具有优越性^[78]。Pflug (2004)、Kibzun (2006)、Andersson (2007)、Pownall (1999)、以及 Mulvey 等学者 (2006) 等研究学者在后来的研究中都对 CVaR 的优越性进行了验证,并把这种方法应用到了信用风险管理、证券组合优化管理等方面^[79-83]。Jonathan A. Batten (2014) 使用了资产回报率的多重分形模型对欧元对美元的 VaR 值进行估算,并将研究结果与基准模型进行对比,实证结果显示资产回报率的多重分形模型比基准模型的性能好^[84]。Stavros Degiannakis 等学者 (2017) 提出日间或者日内模型能否进行短期

预测以及能否对 VaR 值进行准确的估算的疑问,使用 AR (1) -GARCH (1, 1) -skT 以及 AR (1) -HAR-RV-skT 两种方法来对日间和日内波动率模型的性能进行比较,实证结果显示日内波动率模型在预测方面的性能不如在 VaR 方面的性能好^[85]。

国内也有很多学者们使用 VaR 这一理论对风险进行估算,顾乃康(1998)将行业中较为常用的资产负债管理方法与所提及的 VaR 方法这两种市场风险管理进行了系统的比较,得出了一些方法与措施的建议^[86]。李亚静等学者(2000)介绍了风险指标、历史模拟、拨靴法及厚尾法四种风险价值估算方法,并将这四种方法应用到了加元对美元市场、马克对美元市场、法郎对美元市场、日元对美元市场、以及英镑对美元市场这五种汇率市场的汇率历史数据中,并且估算了相应汇率市场下的 VaR 值^[87]。吴振翔等学者(2004)基于 Archimedean Copula 函数方法对欧元和日元投资组合的 VaR 值进行估算^[88]。杨艳军等学者(2005)将 GARCH 模型应用到 VaR 方法中来对风险进行度量,对上海期货交易所中铜和铝的合约进行了风险值估算^[89]。龚妮(2006)将 GARCH 模型于 VaR 方法相结合对美元市场进行风险度量^[90]。崔嵬(2003)将 VaR 方法应用到了银行市场中,结合投资主体不同的风险承受能力给出了实例^[91]。娄可元等学者(2009)将 GARCH 模型应用到了三种分布中,这三种分布分别是正态分布、t 学生分布及 GED 三种分布,将在不同的分布假设下对人民币对美元汇率体制改革后的日汇率中间价作为实证数据,对置信水平为 95%以及 99%的汇率日收益率序列的 VaR 值进行估算^[92]。钟波等学者(2007)应用统计方法中的 MCMC 算法来对参数的 Bayes 估计进行求解,从而估算出金融风险值 VaR 值,实证结果显示这种方法可以辅助决策者根据观测得到的数据和自己的经验信息对 VaR 模型进行调整,从而使决策更加有效^[93]。房小定等学者(2013)利用 VaR 模型对风险进行度量,得到了 EGARCH (1, 2) -GED 分布能够较好的刻画 Shibor 对数收益率序列的分布的结论^[94]。陆静等学者(2013)对人民币对四种外汇汇率历史数据序列波动的相关参数进行估计时使用了 GARCH 模型,学者们根据各个货币在组合中所占的比重以及资产组合的整体条件方差,分别估算了在置信水平为 95%以及 99%时资产组合的 VaR 值^[95]。苟红军等学者(2015)对通过运用 GARCH-EVT-Gaussian-COPULA 模型、GARCH-EVT-t-COPULA 模型以及应用历史模拟法等多个经典理论进行估算的 VaR 值进行比较,实证结果显示 GARCH-EVT-Gaussian-COPULA 模型和 GARCH-EVT-t-COPULA 模型的预测可靠性优于其他在研究中所使用的四种度量方法^[96]。刘飞等学者(2015)运用正态分布和非对称拉普拉斯分布对三种汇率历史数据序列进行实证研究,实证数据的汇率种类包括美元对日元,美元对英镑,美元对欧元,时间跨度为 1991 年至 2012 年。实证结果显示非对称拉普拉斯分布比正态分布更合适在 VaR 方法中风险控制的检验^[97]。李洛源(2016)将 VaR 方法应用到上海银行间同业拆借利率的风险度量中,在 kupiec 检验的基础上,基于各个 GARCH 模型估算出 VaR 的均值和标准差,增强风险度量的完

整性和可靠性^[98]。

1.3.4 集成算法的研究综述

集成算法是 Bates 等学者（1969）提出的一种算法，这种方法最早应用在预测方面，集成算法结合预测模型提高了预测的准确度，使得集成算法逐渐受到了研究学者们的关注^[99]。传统的模型算法是通过单一的模型从数据中寻找或者提取相关的特征，然后寻找数据内部真正存在的线性或者非线性关系，从而对数据进行预测或者分类。但是集成算法是相当于将多个传统模型组合集成起来，然后将多个模型的预测或者分类结果进行一定的组合，从而得出一个最终的结果。

集成算法主要分为输入变量重构^[100]、输出变量重构^[101]、样本重取样^[102]、以及参数选择法^[103]四种。与传统模型的算法相比，集成算法的性能要优于传统单一的模型，得到了很多研究学者们的重视，集成算法不再只应用在提高模型预测准确度方面，研究学者们也将集成算法应用到了语音识别、图像识别、分类等领域中，都得到了较好的效果^[104]。

在预测方面，Sharkey A J C 等学者（1998）在医学乳腺癌的诊断方面将集成方法以及神经网络模型相结合，实证结果显示模型的预测较好^[105]。

在图像识别方面，Hansen L K 等学者（1992）提出用集成方法将投票法以及多个神经网络模型进行结合的方法，对手写体的数字进行识别，实证结果显示相比于单一的传统神经网络识别准确度，集成方法的识别准确度要更高^[106]。Gutta S 等学者（1996）在人脸识别方面的应用中应用集成方法将支持向量机和决策树方法进行结合，对多人种、多姿态的正面人脸进行识别，实证结果显示这种集成方法提高了模型的识别准确度^[107]。Schwenk H 等学者（2000）在之前研究学者们的研究基础上，提出了一种将 Ada Boost 集成方法与神经网络模型结合的方法，对手写体的字符进行识别，实证结果显示这种方法得到了较好的效果^[108]。

在语音识别方面，Hampshire J 等学者（1990）将集成算法与延时神经网络进行结合，对单扬声器以及多扬声器中的语音进行识别研究，实证结果显示这种方法大幅减少了识别错误率^[109]。

在分类方面，Shim shoni Y 等学者（1998）在地震波分类方面应用了集成方法，研究学者们将三种表示地震波信号的数据作为输入值应用到与集成方法相结合的三种不同网络结构的模型中，然后应用加权平均方法对这三个集成的分类值进行集成得到最后的分类结果^[110]。Schapire 和 Singer（2000）将神经网络集成方法中的 Boosting 方法应用到了文本分类方面。实证结果显示 Boosting 方法的分类效果要优于这方面分类的常用技术^[111]。

在风险测度方面，目前集成算法的应用相比于这种方法在语音识别、图像识别、分类等领域中的广泛应用来说较少。Himmat Singh 等学者（2014）在差分方法中，将集成方法和突变与交叉参数方法相结合，计算电力市场的 VaR 值，实证结果表明这种方法比古典差分方法以及自适应差分方法的性能要好^[112]。

国内也有研究学者们将在股市预测以及汇率预测方面效果较好的集成算法应用到了预测、图像识别、语音识别、分类等方面。

在预测方面，张秀艳等学者（2003）将神经网络集成技术应用到了股市预测方面，建立了宏观分析模型、基本数据模型等多个股市预测模型^[113]。宋星光等学者（2004）在水库调度方面应用了神经网络集成方法中的 Bgagnig 方法，来预测水库的平均值从而对水库调度进行决策^[114]。

在语音识别方面，孙冰等学者（2006）在语音识别方面应用了集成方法，在汉语汉语孤立数码语音识别系统的研究中加入了覆盖思想的设计，实证结果显示语音识别的效果与其他基准模型相比有了提高^[115]。钱博等学者（2008）为解决语音识别过程中话语意义以及说话者思想这两个特征相互影响的问题，提出了一种神经网络集成方法，从而从语音特征中及时区分开这两个互扰信息，为语音识别方面提供了一种新的方法^[116]。

在分类方面，周水庚等学者（2000）应用了神经网络集成中的 Boosting 算法，对中文文档的分类进行了实证研究^[117]。刁力力等学者（2002）同样在文本分类方面将神经网络集成中的 Boosting 算法与决策树方法进行结合，实证结果显示这种方法取得了较好的效果^[118]。

在风险测度方面，Kaijian He 等学者（2009）将小波分解方法与集成算法相结合，构建了一个小波分解非线性集成 VaR 测算方法，将纽约交易所的迪拜、美国西德克萨斯轻质原油以及布伦特原油市场的数据作为实证数据，实证结果表明这种算法提高了风险估计的准确度以及可靠性^[119]。同样，Kaijian He 等学者（2012）在之前研究的基础上将集成算法与多分辨率分析方法相结合，对伦敦金属市场以及纽约商品交易所的铜、金和银金属市场进行 VaR 值的测算，实证结果表明这种方法比传统的 VaR 估计方法性能更好^[120]。

1.3.5 深度学习方法的研究综述

深度学习（Deep Learning）是指通过多层神经网络拟合来训练样本分布的一种新的机器学习方法，这种方法在一定程度上解决了传统神经网络算法在训练 2 层以上的神经网络时容易出现的局部最优问题。深度学习的这个特征使得它比较适合处理非线性自然信号，例如自然语言处理、图像识别、语音识别、大数据特征提取等方面，

这为正面临着瓶颈情况的多个信号处理问题提供了可尝试的新方法。Jürgen Schmidhuber (2015) 对深度学习的发展进行了一个回顾^[121]。深度学习出现的时间还不算特别长, 目前来说, 国内外研究中较为常用的深度学习模型主要有深度置信网络模型 (Deep Belief Networks, DBNs)、深度卷积神经网络模型 (Deep Convolutional Neural Networks, DCNNs)、堆叠自编码器模型 (Stack Auto-Encoder, SAE)、深度递归神经网络模型 (Deep Recurrent Neural Networks, DRNNs) 等模型, 大部分模型都是以卷积神经网络模型 (Convolutional Neural Networks, CNNs)、受限玻尔兹曼机模型 (Restricted Boltzmann Machine, RBM) 递归神经网络模型 (Recurrent Neural Networks, RNNs)、自编码器模型 (Auto-Encoder, AE) 等模型的基础上拓展构建的模型。

深度学习在实际应用中取得了较为成功的应用。谷歌公司利用近两万个处理器构建了一个节点数量超过十亿的大规模深度网络结构, 目前谷歌公司已经在 Google Glass、图像搜索等项目中应用了所取得的研究成果。

在语音识别方面, 学者们提出了深度神经网络模型 (Deep Neural Network, DNN), LING Z 等学者 (2013) 提出了将 HMM-DNN 模型和 HMM-DNN 模型进行语音合成方法。除了 HMM-DNN 模型外, CNN 模型和 RNN 模型也被应用到了语音识别的模型构建中^[122]。ABDEL-HAMID O 等学者 (2012) 研究了 CNN 模型应用到语音识别方面的声学建模方法。CNN 模型提高了识别不同的人说话的稳定性^[123]。GRAVES A 等学者 (2013) 尝试了将 RNN 模型应用于语音识别的模型构建中, 在 TIMIT 语料库上取得了较好的识别效果^[124]。KOVACS G 等学者 (2015) 把 CNN 模型和 Relus 激活函数相结合, 将模型应用于语音识别方面, 这种结合方法使得 CNN 模型的性能得到看进一步提升^[125]。

随着深度学习模型的广泛应用, 学者们将其应用范围扩大到了预测方面。Mohammad Hossein Rafiei 等学者 (2015) 将玻尔兹曼机与遗传算法相结合对房地产价格进行预测, 这个方法具有一定的有效性和准确性^[126]。Quoc V. Le (2015) 结合简单的例子对非线性分类、反向传播算法进行了解释, 将多层感知器模型和循环神经网络模型与深度学习理论进行结合^[127]。Xiao Ding 等学者 (2015) 用卷积神经网络模型对短期和长期的股票价格进行预测, 实证结果显示卷积神经网络模型的预测结果比国家基准线的预测方法的预测性能更好^[128]。Haoyang Zeng 等学者 (2016) 将卷积神经网络应用到了生物方面, 对 DNA-蛋白质的转录进行了预测^[129]。Chengzhang ZHU 等学者 (2014) 在振荡箱理论背景下, 应用深度置信网络根据历史的时间序列数据对股票未来的价格进行预测, 实证证明这种方法可以对股票交易的决策提出可行性建议^[130]。Takaomi HIRATA 等学者 (2015) 将时间序列模型以及深度置信网络相结合, 实证结果显示这种结合方法具有可行性^[131]。Aleksandra Dedinec 等学者 (2016) 应用

深度置信网络模型对 2008 年至 2014 年的马其顿短期电力负荷进行预测, 实证结果显示深度置信网络模型相比于传统的预测方法有着明显的优势^[132]。Ryo Akita 等学者 (2016) 将深度学习模型、段落向量、长短期记忆模型结合, 构建了一个应用到金融时间序列预测方面的新模型, 所预测的结果在实证方面得到了证实^[133]。

随着深度学习模型的发展, 国内学者们对深度学习模型的研究也有了进展。

在语音识别方面, 王龙等学者 (2015) 针对循环网络模型计算过程较为复杂, 从而使得模型语言训练效率很低的问题, 构建了一个基于批处理的并行优化训练算法, 这个模型可以并行处理多个语音样本, 实证结果表明该方法有效提高了模型的效率, 并且同时不影响模型的识别性能^[134]。孙志远等学者 (2016) 对目前语音识别领域的研究模型以及应用进行了介绍^[135]。张春云等学者 (2017) 为了更准确的获取语句的信息, 构建了基于卷积神经网络的自适应权重 multi-gram 语句系统应用在自然语言处理方面, 在电影评论正负倾向和分类中应用, 实证结果显示模型的性能较好^[136]。

在图像识别方面, 何炳金等学者 (2017) 针对现如今计算机对图像数据进行检索时有一定困难的问题, 应用卷积神经网络对图像进行自动标注, 获得了较好的效果^[137]。李传朋等学者 (2017) 针对目前网络模型在图像提取特征时会受到噪声影响问题, 构建了一种基于深度卷积神经网络模型的去噪声方法, 实证结果表明这个方法快速并且效果好^[138]。严严等学者 (2017) 对近年来图像识别领域中的人脸识别方向的研究进行了综述, 对人脸识别的任务进行了分类并对数据库进行了介绍^[139]。迟元峰等学者 (2017) 为了更好更全面的提取视频中的信息, 来对人体行为进行识别, 将卷积神经网络模型和循环神经网络模型进行结合, 对视频中的人体行为进行分类, 实证结果显示模型的分类结果较好^[140]。

国内学者们也将深度学习方法应用到了预测方面, 田文慧等学者 (2015) 提出现有的工厂工作环境有着温度高、多粉尘的特点, 烟轮机的工作环境以及状态无法准确预测, 所以构建了一个基于深度置信网络的预测失效模型, 实证结果表明该模型在预测烟轮机退化趋势方面准确度较高^[141]。周文杰等学者 (2016) 将深度置信网络模型应用到电信客户投诉预测方面, 提高了企业客户的满意度以及企业的竞争力, 同时证明了深度置信网络模型具有自动学习的特征, 并且在电信大数据领域有着可行性^[142]。王声柱等学者 (2017) 认为深度学习的思想与入侵检测研究中通过学习来提高检测概率以及降低误报率的思想较为相似, 所以尝试将深度学习理论与入侵检测相结合, 实证研究结果显示这种结合方法可以提高入侵系统的分类准确度^[143]。罗向龙等学者 (2017) 认为现有的预测研究方法并不能充分的表述交通流内部的本质规律, 从而将深度置信网络模型和支持向量回归方法结合, 对交通流量进行预测, 与传统预测方法相比较, 这种方法提高了预测的准确度, 是一种有效可行的预测方法^[144]。

深度学习模型是近年来新兴的神经网络模型, 应用在语音识别、图像识别方面较

多，而应用在外汇市场方面较少，所以将深度学习模型应用于外汇市场，进行汇率预测、风险度量的相关研究具有着极高的研究价值。因此，本研究将在目前学者们研究进展的基础上，构建能够预测外汇市场汇率以及度量外汇市场风险的深度学习模型，对国际主流汇率市场中的 7 个国际汇率市场进行汇率预测并对相应的外汇市场进行风险度量，将实验结果与基准模型作对比，找到精准度最高和可靠性最好的模型。

1.4 本文研究内容以及创新点

在全球一体化飞速发展的形势下，如何加强对外汇市场的汇率预测的精准度，以及如何增强汇率风险测度的可靠性并且找出高效可操作的风险管理方法是学者们高度关注的课题。因为汇率的波动幅度以及波动趋势都对汇率风险的产生有着很大的影响，而且汇率对国际贸易活动有着至关重要的影响。所以本研究将外汇市场的汇率预测以及汇率市场风险度量作为研究内容，如图 1-1 所示，对包括人民币（RMB）、欧元（Euro）、日元（Jap）、加拿大美元（Cad）、澳洲美元（Aud）、英镑（Gbp）、瑞士法郎（Chf）在内的 7 个外汇市场国际主流汇率作为研究对象，应用深度学习模型对数据进行建模分析，得到汇率预测值、风险度量值以及相关参数的估计值。然后根据实验结果对模型进行优化，提高模型在汇率预测以及风险度量方面的精准度与可靠性。

本研究的论文框架分为以下五部分。

第一章为绪论。其中绪论这部分主要对本研究中的背景、目的、意义进行相关介绍，而后总结国内外外汇市场汇率预测、外汇市场风险度量、深度学习模型、风险度量以及集成算法方面研究的综述。在外汇市场汇率预测综述部分，对国内外外汇市场汇率预测的发展情况进行介绍，对目前外汇市场汇率预测的研究方法进行了总结。在外汇市场风险度量综述部分，对汇率风险理论进行了简单的介绍，并对国内外汇率风险测量的方法进行了归类。深度学习模型综述部分，对深度学习的分类作了简单介绍，并且针对目前应用的方向以及相关研究进行了综述。在风险度量综述部分，对 VaR 方法的国内外研究现状进行了综述。在集成算法综述部分，对国内外应用在预测、图像识别、语音识别以及分类方面的集成方法进行了综述。除此之外，在第一章还提出了本文的研究方法和思路，并对论文框架和创新点做了解释说明。

第二章为理论概论。对深度学习模型的基本定义、分类和计算方法进行介绍。并对深度多层感知器网络模型、循环神经网络模型和深度置信网络的基本定义，基本结构以及主要算法进行了介绍。

第三章为基于深度学习模型的外汇市场汇率预测。在这部分介绍汇率预测的主要方法以及结合 7 个国际主流汇率市场的历史数据进行实证研究。包括对数据的选择和

来源进行描述，然后对深度置信网络模型、循环神经网络模型的参数进行了估计，并且对数据集进行实证研究。

第四章为基于深度学习模型的风险度量。对 VaR 算法的定义和计算方法进行了说明，并且用集成算法将 ARMA-GARCH 模型与深度学习模型相结合，将这种新的方法应用在了外汇市场风险度量方面，并对模型的参数进行估计，对数据集进行实证研究，将度量结果与传统的基准模型进行比较，然后对实证结果进行分析并进行检验。

第五章为总结与展望。总结和展望部分是对本研究论文中所提及的理论以及实证的全部内容进行总结和评价，而后在此研究状况的基础上指出下一步研究的方向，并对基于深度学习模型的外汇市场预测与风险度量研究部分的未来研究做出展望。

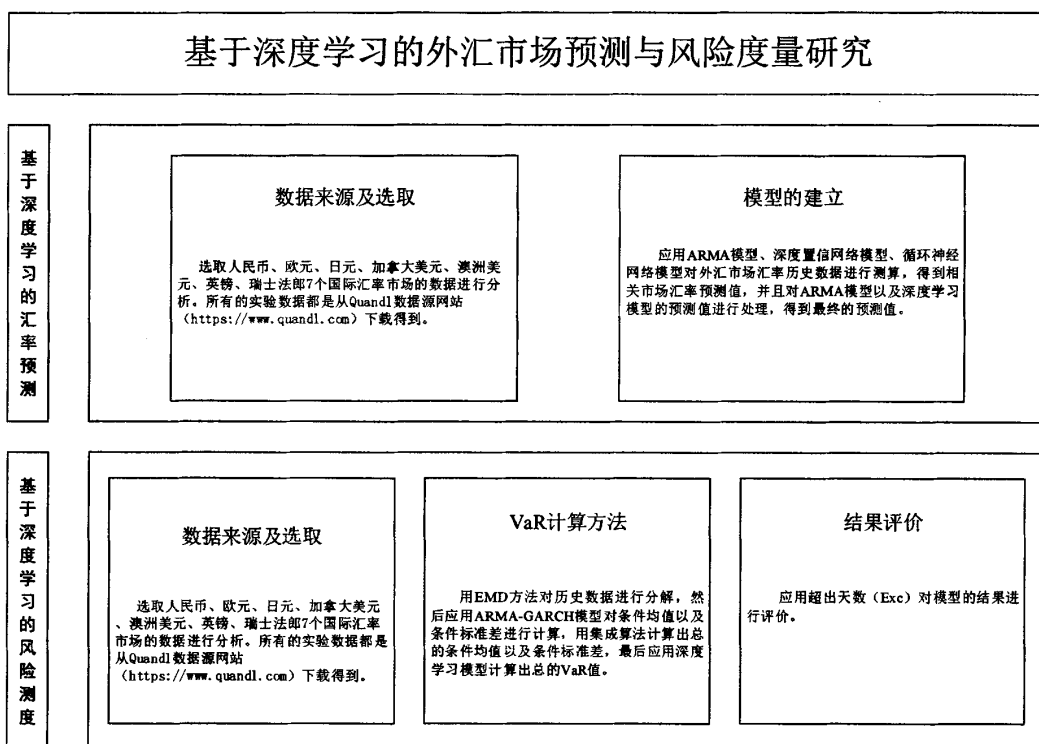


图 1-1 基于深度学习的外汇市场预测与风险度量研究内容

Fig1-1 Research content of exchange rate and VaR in 7 foreign exchange markets based on deep learning

本研究的创新点主要有以下两个方面：

1、在外汇市场汇率预测方面，本研究提出一种新的基于深度学习理论的汇率预测模型。同时考虑历史数据的线性以及非线性特征，使用对 ARMA 模型的预测结

果以及深度学习模型的预测结果线性集成的方法来对预测结果进行处理。本研究所提及的模型对包括人民币、欧元、日元、加拿大美元、澳洲美元、英镑、瑞士法郎在内的主流外汇市场的国际汇率进行实证研究。实证结果发现基于深度置信网络模型和循环神经网络模型构建的汇率预测模型具有较好的预测精度。

2、在外汇市场风险度量方面，本研究提出了一种新的基于深度学习理论的集成 VaR 风险测度模型，该模型采用 EMD 方法对外汇市场的历史数据进行不同时间尺度的分解，然后应用 ARMA-GARCH 模型对条件均值以及条件标准差进行估算，用集成算法计算出总的条件均值以及总的条件标准差，最后应用深度学习模型计算总的 VaR 值。本研究采用包括人民币、欧元、日元、加拿大美元、澳洲美元、英镑、瑞士法郎等主流外汇市场国际汇率进行实证研究。实证研究表明所提及的模型具有较高的可靠性与精确性。

第二章 深度学习理论

本章包括以下三部分，第一部分是介绍神经网络的主要结构以及较为常用的几种激活函数，第二部分是对本研究中应用的三种深度学习模型进行介绍，介绍深度学习模型的相关结构以及主要算法，第三部分是对深度学习模型中较为常用的优化算法进行简单介绍。

深度学习（Deep Learning）是 Hinton 在 2006 年提出的一种新的模型，这种模型是一种通过多层神经网络对样本进行训练的新的机器学习方法。这种方法是通过模拟人类大脑对事物进行学习分析的方式来构建一个神经网络，主要用来对历史数据之间的关系进行描述和解释，近年来，深度学习成为机器学习领域新兴的研究方向^[145]。深度学习这种新的机器学习方法在相比于传统神经网络模型来说，如果传统神经网络模型具有 2 层及 2 层以上的网络结构，则在训练过程中比较容易出现局部最优化的缺点，但是同样网络结构的深度学习模型则不会出现这样的问题。深度学习模型是人工智能领域的新的一个研究方向，它与传统的神经网络相比来说更适用于处理具有非线性特征的数据，例如语音数据、图像数据等，通过多层的特征映射从图片、声音、文字等输入数据中提取有助于理解数据本身所代表的含义的有效特征表示，为许多目前研究进度停滞的信号处理问题提供了可尝试的新方法^[146]。对人类大脑皮层结构的模拟是研究学者们对深度学习模型进行构建的灵感来源。人类大脑皮层对接收到的信息进行逐层的分解与理解，深度学习模型的结构一般来说也是分层的结构，也对网络的输入数据进行逐层的处理，这与人类大脑的结构很相似。深度学习模型每一层提取数据的一个或者多个不同方面的特征，并且将提取到的特征作为下一层的输入数据，通过不断组合这些低层的数据特征从而形成更加抽象的高层特征。深度学习模型与浅层神经网络模型相比较而言，它的优势在于深度学习模型能够更好的表示复杂的高维函数^[147]。而深度学习方法的目的是去寻找历史数据内部真正的关系，对他们进行描述和解释。目前，国内外研究中常用的深度学习模型主要有受限玻尔兹曼机模型（Restricted Boltzmann Machine, RBM）、卷积神经网络模型（Convolutional Neural Networks, CNNs）、循环神经网络模型（Recurrent Neural Networks, RNNs）、自编码器（Auto-Encoder, AE）等模型，而在这些模型的基础上，学者们又构建了深度置信网络模型（Deep Belief Networks, DBNs）、深度递归神经网络（Deep Recurrent Neural Networks, DRNNs）、深度卷积神经网络模型（Deep Convolutional Neural Networks, DCNNs）、堆叠自编码器模型（Stack Auto-Encoder, SAE）等深度学习模型^[148]。

2.1 神经网络

学者们通过对人类大脑的组织结构和运行机制进行模拟，从而构建出类似于人类大脑结构和智能行为的一种智能仿生网络，这种网络就是人工神经网络（Artificial Neural Network, ANN）使得网络可以学习以及识别信息。神经网络已经被广泛的应用到时间序列分析、生物医学、语音识别、图像处理、性能材料分类识别等多个领域中^[149]。

2.1.1 神经元

神经网络最基本的单位是神经元，神经元之间通过权重来连接彼此。神经网络通过多个通过权重互相连接的神经元来对数据信息进行计算和学习，然后对输出值以及输入值进行描述，对数据之间复杂的非线性内在关系进行表达。

神经元结构如图 2-1 所示。神经元的输入值由 $x_1, x_2 \cdots x_m$ 来表示，连接权重由 $\omega_1, \omega_2 \cdots \omega_k$ 来表示，实际输出值由 $y_1, y_2 \cdots y_k$ 来表示。输出值与输入值的定义如下^[149]：

$$u_k = \sum_{j=1}^m \omega_{kj} x_j \tag{2-1}$$

$$y_k = \varphi(u_k) \tag{2-2}$$

其中， $\varphi(\cdot)$ 表示激活函数。

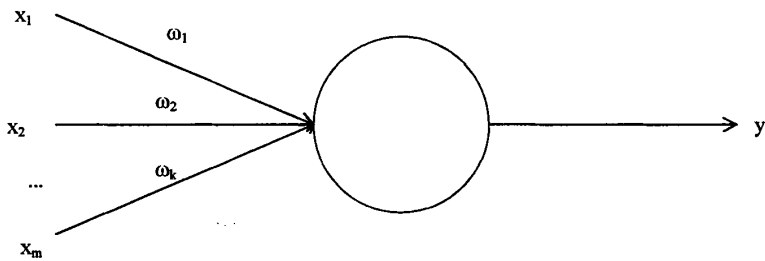


图 2-1 神经元结构图

Fig2-1 Neuronal structure

2.1.2 激活函数

神经元通过激活函数神经网络的输出值以及其他区域进行定义。激活函数主要有以下几种^[150]：

- 1、阶梯函数

阶梯函数又可以称为阈值函数，如果自变量的值大于等于 0，则阶梯函数相应的输出值是 1；如果自变量的值小于 0，则阶梯函数相应的输出值是 0。阶梯函数定义如下：

$$\varphi(v)=\begin{cases} 1 & \text{if: } v \geq 0 \\ 0 & \text{if: } v < 0 \end{cases} \quad (2-3)$$

阶梯函数也被称作是赫维赛德函数（Heaviside function），在使用阶梯函数作为激活函数的时候，此时的神经元 k 的输出函数表示为：

$$y_k = \begin{cases} 1 & \text{if: } v_k \geq 0 \\ 0 & \text{if: } v_k < 0 \end{cases} \quad (2-4)$$

其中， v_k 表示着局部感应区域：

$$v_k = \sum_{j=1}^m w_{kj} x_j + b_k \quad (2-5)$$

2、Sigmoid 函数

Sigmoid 函数又可以称为 s 型函数，是构建神经网络模型时较为常用的激活函数。可以用来对函数进行拟合、优化以及分类。

Sigmoid 函数定义如下：

$$\varphi(v) = \frac{1}{1+e^{-av}} \quad (2-6)$$

其中， a 表示着斜率参数。

3、线性校正函数（Rectified Linear Units）

线性校正函数也是神经元模型中一个典型的模型。它的结构可以看做在所有 Sigmoid 函数具有相同权重以及偏差的情况下，多个 Sigmoid 函数搭建而成。偏差有着类似于 -0.5, -1.5, -2.5... 的不同的固定偏移。概率之和定义如下：

$$\sum_{i=1}^N \varphi(v-i+0.5) \approx \log(1+e^v) \quad (2-7)$$

其中， $v=w^T x+b$ ，这个函数需要 S 形函数多次运算来获得所需概率正确的采样整数值。可以近似定义为：

$$\varphi(v) = \max(0, v + N(0, \varphi(v))) \quad (2-8)$$

其中， $N(0, \sigma)$ 表示着具有零均值以及方差 σ 的高斯噪音。

4、分段线性函数

分段线性函数在 $[-1, 1]$ 的区间上有着同向的缩放性，常被当做非线性的缩放工具，分段性函数定义如下：

$$\varphi(v) = \begin{cases} 1 & \text{if: } v > 0 \\ 0 & \text{if: } v = 0 \\ -1 & \text{if: } v < 0 \end{cases} \quad (2-9)$$

2.2 深度学习模型

2.2.1 深度置信网络 (Deep Belief Networks, DBNs)

1、受限的 Boltzmann 机 (Restricted Boltzmann Machine, RBM)

玻尔兹曼机 (Boltzmann Machine) 是由 Hinton 和 Sejnowski 在网络模型的基础上加入了随机机制从而提出的一种概率神经网络模型, 这可以看作是随机生成形式的网络^[151]。玻尔兹曼机是一种无监督学习的神经网络, 能够有效的对数据进行学习。如图 2-2 所示, 受限的玻尔兹曼机是一个由可视层 v 和隐藏层 h 组成的具有两层神经网络结构的模型, 网络结构中的节点都是二值变量节点, 不仅可视层中的节点是没有任何连接的, 而且隐含层中的节点也是没有任何连接的, 也就是说它们之间都是没有任何连接的。如果可视层 v 可知, 则将满足玻尔兹曼分布的全概率分布定义为 $P(v, h)$, $P(h|v) = \prod_{i=1}^n P(h_i|v)$, 同样的情况下, 如果隐藏层 h 可知, 则根据 $P(v'|h)$ 就可以得到相应的可视层, 其中, v' 表示着这种情况下的相对应的可视层单元^[152]。

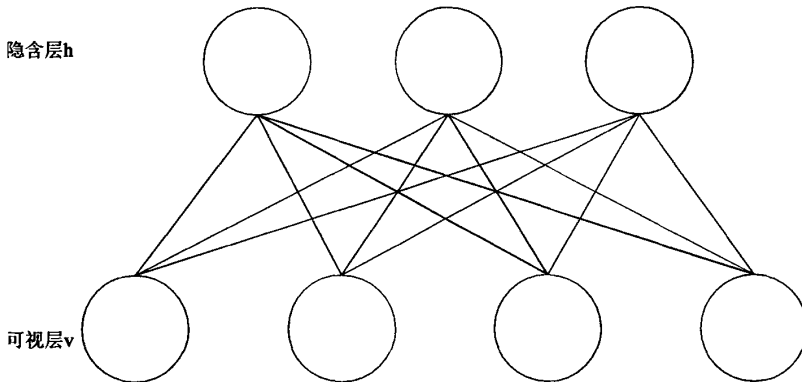


图 2-2 受限的玻尔兹曼机模型结构图

Fig2-2 Structure of Restricted Boltzmann Machine

网络中可视节点和隐藏节点的联合概率分布定义如下:

$$P_{\theta}(v, h) = \frac{1}{Z(\theta)} e^{-E(v, h, \theta)} = \frac{1}{Z(\theta)} \prod_{i,j} e^{w_{ij} v_i h_j} \prod_i e^{b_i v_i} \prod_j e^{a_j h_j} \quad (2-10)$$

其中, $Z(\theta) = \sum_{v, h} e^{-E(v, h, \theta)}$

RBM 能量函数定义如下:

$$E(v, h, \theta) = -a^T h - b^T v - v^T w h \quad (2-11)$$

$$= - \sum_{i,j} W_{ij} v_i h_j - \sum_i b_i v_i - \sum_j a_j h_j$$

其中, $\theta = \{W, a, b\}$ 表示着模型参数, E 表示着期望值。 a_j 和 b_i 表示着输入变量 v_i 与隐藏变量 h_j 的偏差, W_{ij} 表示着在可视层与隐含层中 i 和 j 两者之间相互作用的权重, 数被向量 v 的边际概率分布 P 定义如下^[153]:

$$\begin{aligned} P(v; \theta) &= \sum_h \frac{e^{-E(v, h, \theta)}}{\sum_{v, h} e^{-E(v, h, \theta)}} \quad (2-12) \\ &= \frac{1}{Z(\theta)} \sum_h e^{-E(v, h, \theta)} \\ &= \frac{1}{Z(\theta)} \sum_h e^{a^T h + b^T v + v^T W h} \\ &= \frac{1}{Z(\theta)} e^{b^T v} \prod_{j=1}^F \sum_{h_j \in \{0, 1\}} e^{(a_j h_j + \sum_{i=1}^D W_{ij} v_i h_j)} \\ &= \frac{1}{Z(\theta)} e^{b^T v} \prod_{j=1}^F (1 + e^{a_j + \sum_{i=1}^D W_{ij} v_i}) \end{aligned}$$

在这个等式中, v 表示着输入向量, h 表示着隐含单元的向量。这是可见单位的边际分布函数。具有低能量的系统 (v, h) 被给予了高概率, 而具有高能量的系统则被给予了低概率。根据能量函数, 则可以定义以下概率:

$$P(v|h; \theta) = \prod_i P(v_i|h) \quad (2-13)$$

$$P(v_i=1|h) = \phi(b_i + \sum_j h_j W_{ij}) \quad (2-14)$$

$$P(h|v; \theta) = \prod_j P(h_j|v) \quad (2-15)$$

$$P(h_j=1|v) = \phi(a_j + \sum_i v_i W_{ij}) \quad (2-16)$$

其中, ϕ 表示着 sigmoid 函数 $\phi(x) = \frac{1}{1+e^{-x}}$ 。能量函数是在二值变量上进行定义, 所以这并不适用连续数值, 但是通过修改能量函数使其变为高斯伯努利 RBM, 则依然可以表示连续数值:

$$E(v, h, \theta) = \sum_i \frac{(v_i - b_i)^2}{2\sigma_i^2} - \sum_j a_j h_j - \sum_{i,j} W_{ij} \frac{v_i}{\sigma_i} h_j \quad (2-17)$$

其中, 向量 $\theta = \{W, a, b, \sigma^2\}$ 和 σ_i 表示输入变量 v_i 的方差。向量 a 表示着可视单位的偏差, b 表示着隐藏单位的偏差。可视向量 v 的边缘分布为^[153]:

$$P(v; \theta) = \sum_h \frac{e^{-E(v, h, \theta)}}{\sum_h e^{-E(v, h, \theta)}} \quad (2-18)$$

$P(v|h)$ 变成平均值为 $a_i + \sigma_i \sum_j W_{ij} h_j$ 的多元高斯分布和对角协方差矩阵, 可视的条件分布以及隐藏单位变成了如下形式:

$$P(v_i = x | h) = \frac{1}{\sigma_i \sqrt{2\pi}} e^{-\frac{(x - b_i - \sigma_i \sum_j W_{ij} h_j)^2}{2\sigma_i^2}} \quad (2-19)$$

$$P(h_j = 1 | v) = \phi \left(b_j + \sum_i \frac{v_i}{\sigma_i} W_{ij} \right) \quad (2-20)$$

其中, ϕ 表示着 sigmoid 函数 $\phi(x) = \frac{1}{1+e^{-x}}$ [154], 如果数据被处理成标准化的 $[0, 1]$ 区间, 则伯努利 RBM 也可以处理连续数据。

2、深度置信网络 (Deep Belief Networks, DBNs)

深度玻尔兹曼机 DBM (Deep Boltzmann Machine, DBM) 是在 RBM 的基础上不断增加隐藏层的层数 [155], 如果将贝叶斯信念网络模型应用到相对靠近可视层的部分, 同时将 RBM 模型应用到相对远离可视层的部分, 通过这种方法则可以得到深信度网络模型。

如图 2-3 所示, DBN 的结构可以看作是多个 RBM (Restricted Boltzmann Machine, RBM) 组合而成, 每个 RBM 都独立训练自身的参数同时学习上一个 RBM 的参数得到输出值, 上一个 RBM 学习到的数据特征作为下一个 RBM 训练的输入值, 进行下一个 RBM 的训练, 由上到下进行训练, 从而构成了含有多个隐含层的 DBN [156]。

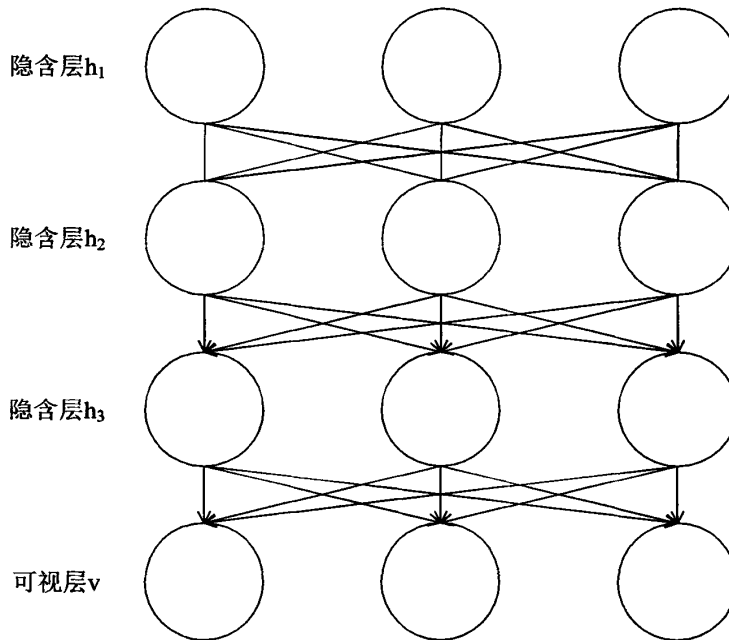


图 2-3 深度置信网络结构图

Fig2-3 Deep Belief Networks structure

2.2.2 深度多层感知器

1、多层感知器

多层感知器（Multilayer Perceptron, MLP）是一种有监督学习的前馈神经网络模型，被研究学者们广泛应用到了优化计算、图像处理、模式识别、最优预测等领域中。多层感知器由输入层、隐含层和输出层三部分组成。同一层中的神经元相对独立，而相邻两层中的神经元则是相互连接的。多层感知器的结构如图 2-4 所示^[157]。

设输入向量为 $x_1, x_2 \cdots x_m$ ，激活函数为 sigmod 函数，隐含层第 k 个隐含节点的输出定义如下：

$$z_k = f(\sum_{j=1}^m \omega_{kj} x_j + \omega_{k0}) \quad (2-21)$$

其中， ω_{kj} 表示着权重，这个权重连接着第 j 个输入节点以及第 k 个隐含节点， ω_{k0} 则表示着第 k 个隐含节点的偏差。

输出层第 i 个节点的网络输出值定义如下：

$$\begin{aligned} y_k &= \sum_{k=1}^n \vartheta_{ik} z_k + \vartheta_{i0} \\ &= \sum_{k=1}^n \vartheta_{ik} \frac{1}{1 + \exp\{-\{\sum_{j=1}^m \omega_{kj} x_j + \omega_{k0}\}\}} + \vartheta_{i0} \end{aligned} \quad (2-22)$$

其中， ϑ_{i0} 表示着连接第 i 个输出节点和第 k 个隐含节点的权重， ϑ_{i0} 表示着偏差。

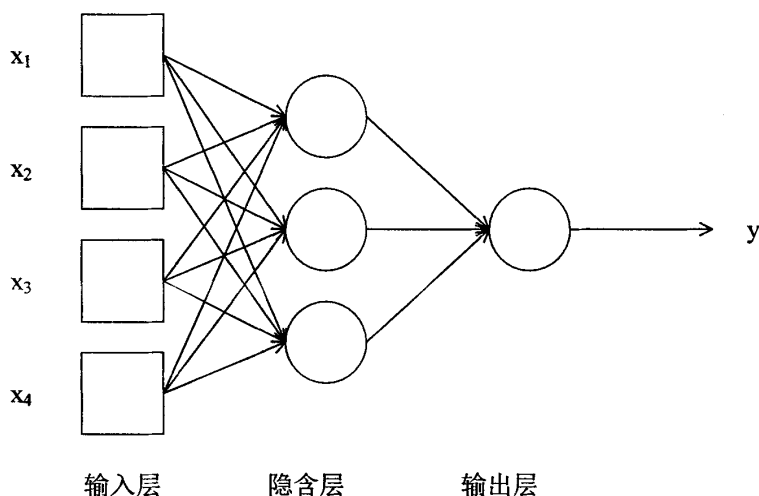


图 2-4 多层感知器结构图

Fig2-4 Multilayer Perceptron structure

2、深度多层感知器

如果多层感知器的基础上增加隐含层的层数，就得到了深度多层感知器。深度多层感知器也可以称为深度神经网络（Deep Neural Network, DNN）。也就是说深度多层感知器有着输入层、输出层和一层以上的隐含层三个部分^[158]。深度多层感知器的结构如图 2-5 所示：

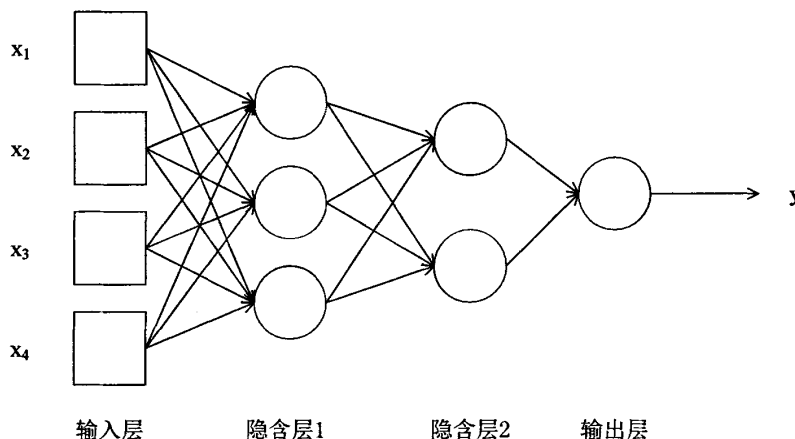


图 2-5 深度多层感知器结构图

Fig2-5 Deep Neural Network structure

2.2.3 循环神经网络（Recurrent Neural Networks, RNNs）

1、循环神经网络（Recurrent Neural Networks, RNNs）

循环神经网络模型与多层感知器模型都是前馈神经网络模型，但是循环神经网络与多层感知器并不是完全相同，循环神经网络不仅有着前馈连接还有着内部的反馈连接，因此反馈神经网络模型是循环神经网络模型的又一个名称。传统的前馈神经网络模型中输入的数据信息通过输入层传到第一个隐含层，然后再传到下一个隐含层，一直到传到输出层。每一层只接受来自上一层的相关信息，换句话说就是传统的前馈神经网络模型中每一层的输出结果只受到它之前的隐含层的影响，并不受到后面隐含层数的影响。但是循环神经网络不仅存在前馈连接也存在着网络内部的反馈连接。如图 2-6 所示，隐含层的反馈不仅进入到输出层，而且也会进入到下一个隐含层中，从而会对下一个隐含层的输出结果有所影响^[159]。

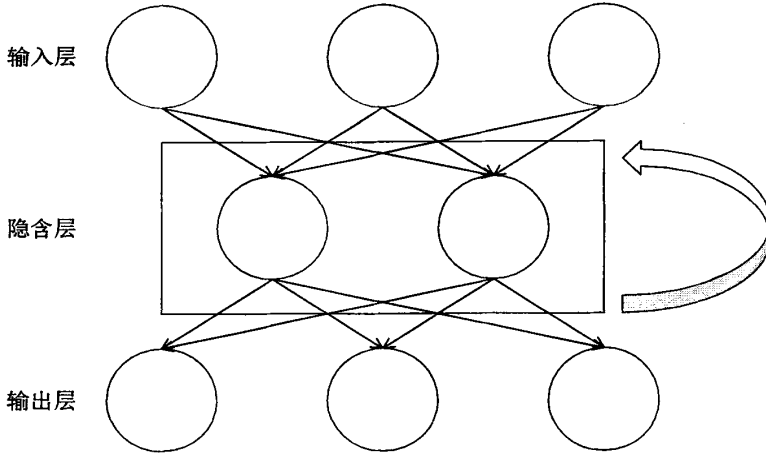


图 2-6 循环神经网络结构图

Fig2-6 Recurrent Neural Networks structure

循环神经网络模型定义如下：

$$x_j(n) = \sum_{i=0}^m \omega_{ij}^a(n) e(n-i) + \omega_j^c(n) y_j(n-1) \quad (2-23)$$

$$y_j(n) = f[x_j(n)] \quad (2-24)$$

$$u_k(n) = \sum_{j=1}^s \omega_{jk}^b(n) y_j(n) \quad (2-25)$$

$$v_k(n) = f[u_k(n)] \quad (2-26)$$

其中， $\omega_{ij}^a(n)$ 表示着隐含层的输入权重， i 表示着输入层的第 i 个节点， j 表示着隐含层的第 j 个节点， ω_{jk}^b 表示着输出层的输入权重， k 表示着输出层的第 k 个节点， $e(n-i)$ 表示着输入层的数值。 $\omega_j^c(n)$ 表示着循环神经网络的输入权重， $x_j(n)$ 表示着隐含层的输入加权和， $y_j(n)$ 表示着输出， $u_k(n)$ 表示着输出层的加权和， v_k 表示着输出。

2、长短期记忆网络模型（Long Short Term Memory Networks, LSTM）

长短期记忆网络模型，简称为 LSTM 模型，它是在循环神经网络模型的基础上，构建的一种新型的深度机器学习神经网络。如图 2-7 所示，长短期记忆网络模型包括包括许多通过递归连接的子网，称为记忆模块。每个记忆模块都包含一个或多个自连接的记忆单元和三个乘法计算单元。三个乘法计算单元分别用输入门、输出门和遗忘门来表示，用来控制神经元的读、写及重置存储的信息^[160]。

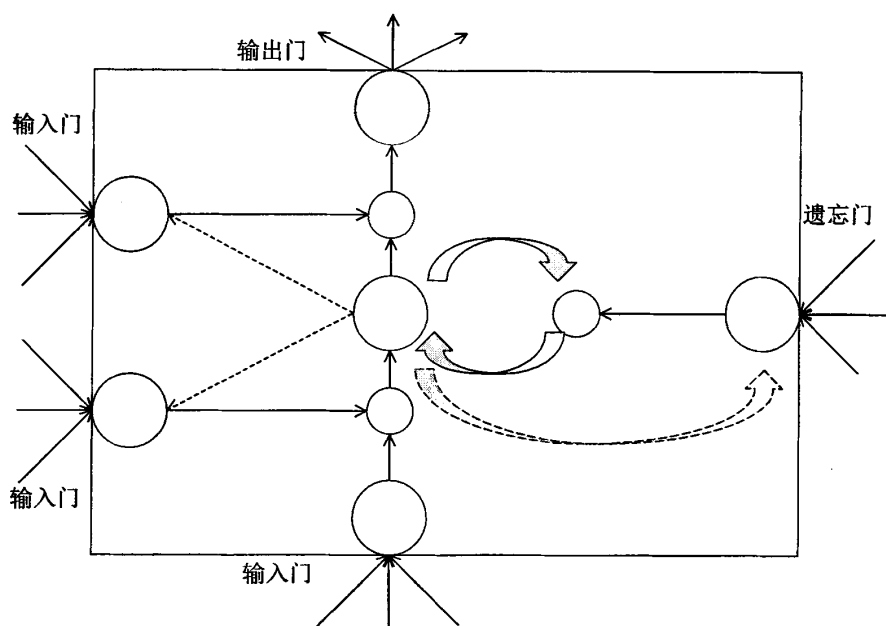


图 2-7 长短期记忆网络模型的结构图

Fig2-7 Structure of Long Short Term Memory Networks Model

长短期记忆网络模型中每个单元中包括输入、输入激励函数处理、输入门控制、自反馈、遗忘门控制、输出门控制、输出激励函数处理、输出等结构。

长短期记忆网络模型的前向计算的输入假设为长度为 T 的序列 x ，从 $t=1$ 时刻开始进行训练，随着 t 不断的增加，利用函数进行计算与递归循环，直到 $t=T$ 时刻结束。

假设长短期记忆网络模型为标准的循环神经网络模型，那么这种情况下所有的状态信息和激励函数都设置为 0， $t=0$ ，而且所有的 δ 在 $t=t+1$ 时是 0。长短期记忆网络模型的定义如下^[161]：

向前传播：

输入门：

$$a_i^t = \sum_{i=1}^I \omega_{il} x_i^t + \sum_{h=1}^H \omega_{hl} b_h^{t-1} + \sum_{c=1}^C \omega_{cl} s_c^{t-1} \quad (2-27)$$

$$b_i^t = f(a_i^t) \quad (2-28)$$

遗忘门：

$$a_\phi^t = \sum_{i=1}^I \omega_{i\phi} x_i^t + \sum_{h=1}^H \omega_{h\phi} b_h^{t-1} + \sum_{c=1}^C \omega_{c\phi} s_c^{t-1} \quad (2-29)$$

$$b_\phi^t = f(a_\phi^t) \quad (2-30)$$

单元内部:

$$a_c^t = \sum_{i=1}^I \omega_{ic} x_i^t + \sum_{h=1}^H \omega_{hc} b_h^{t-1} \quad (2-31)$$

$$s_c^t = b_{\emptyset}^t s_c^{t-1} + b_1^t g(a_c^t) \quad (2-32)$$

输出门:

$$a_{\omega}^t = \sum_{i=1}^I \omega_{i\omega} x_i^t + \sum_{h=1}^H \omega_{h\omega} b_h^{t-1} + \sum_{c=1}^C \omega_{c\omega} s_c^{t-1} \quad (2-33)$$

$$b_{\omega}^t = f(a_{\omega}^t) \quad (2-34)$$

单元输出:

$$b_c^t = b_{\omega}^t h(s_c^t) \quad (2-35)$$

向后传播:

$$\epsilon_c^t = \frac{\partial o}{\partial b_c^t} \quad (2-36)$$

$$\epsilon_s^t = \frac{\partial o}{\partial s_c^t} \quad (2-37)$$

单元输出:

$$\epsilon_c^t = \sum_{k=1}^K \omega_{ck} \delta_k^t + \sum_{h=1}^H \omega_{ch} b_h^{t+1} \quad (2-38)$$

输出门:

$$\delta_{\omega}^t = f'(a_{\omega}^t) \sum_{c=1}^C h(s_c^t) \epsilon_c^t \quad (2-39)$$

状态值:

$$\epsilon_s^t = b_{\omega}^t h'(s_c^t) \epsilon_c^t + b_{\emptyset}^{t+1} \epsilon_s^{t+1} + \omega_{c1} \delta_1^{t+1} + \omega_{c\omega} \delta_{\omega}^t + \omega_{c\emptyset} \delta_{\emptyset}^{t+1} \quad (2-40)$$

单元内部:

$$\delta_c^t = b_1^t g'(a_c^t) \epsilon_s^t \quad (2-41)$$

遗忘门:

$$\delta_{\emptyset}^t = f'(a_{\emptyset}^t) \sum_{c=1}^C s_c^{t-1} \epsilon_c^t \quad (2-42)$$

输入门:

$$\delta_1^t = f'(a_1^t) \sum_{c=1}^C g(a_c^t) \epsilon_c^t \quad (2-43)$$

其中, I 表示着输出层单元的数量, K 表示着输出层单元的数量, H 表示着隐含层单元的数量。

2.3 优化算法

2.3.1 共轭梯度法

共轭梯度法 (Conjugate Gradient Algorithm, CG) 是 Hestenes 以及 Stiefel 提出的一种优化算法^[162]。共轭梯度法不仅在求解大型线性方程组方面有着广泛的应用, 而且研究学者们也提出无约束优化问题方面共轭梯度法也是一种较为有效的方法^[163]。

共轭梯度法主要是用于对无约束优化问题进行求解:

$$\min_{x \in \mathbb{R}^n} f(x) \quad (2-44)$$

迭代公式定义如下:

$$x_{k+1} = x_k + \alpha_k d_k \quad (2-45)$$

$$d_k = \begin{cases} -g_k, & k=0 \\ -g_k + \beta_k d_{k-1}, & k \geq 1 \end{cases} \quad (2-46)$$

其中, x_0 表示着初始点, x_k 表示着当前的迭代点, α_k 表示着步长因子, d_k 表示着第 k 次迭代的搜索方向, g_k 表示着 $f(x)$ 在当前迭代点 x_k 处的梯度向量, β_k 表示着共轭参数。

较为常用的共轭梯度方法主要有以下几种:

Hestenes-Stiefel(HS)方法:

$$\beta_{k-1}^{HS} = \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}} \quad (2-47)$$

Fletcher-Reeves(FR)方法:

$$\beta_{k-1}^{FR} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2} \quad (2-48)$$

Ploak-Ribiere-Polyak(PRP)方法:

$$\beta_{k-1}^{PRP} = \frac{g_k^T y_{k-1}}{\|g_{k-1}\|^2} \quad (2-49)$$

Conjugate -Descent(CD)方法:

$$\beta_{k-1}^{CD} = \frac{\|g_k\|^2}{-d_{k-1}^T g_{k-1}} \quad (2-50)$$

Liu-Storey(LS)方法:

$$\beta_{k-1}^{LS} = \frac{g_k^T y_{k-1}}{-d_{k-1}^T g_{k-1}} \quad (2-51)$$

Dai-Yuan(DY)方法:

$$\beta_{k-1}^{DY} = \frac{\|g_k\|^2}{d_{K-1}^T y_{k-1}} \quad (2-52)$$

其中, $y_{k-1} = g_k - g_{k-1}$

2.3.2 随机梯度下降算法

随机梯度下降法 (SGD) 英文全称是 Stochastic Gradient Descent Algorithm, 它又可以称为渐进梯度下降法, 与共轭梯度法一样是一种优化算法。随机梯度下降法在求解矩阵因子分解问题方面应用广泛, 而且随机梯度下降法不仅实现较为简单而且收敛速度较快^[164]。

随机梯度下降法主要是应用在对具有总和形式的目标函数进行最优化问题的求解:

$$f(\omega) = \sum_{i=1}^n f_i(\omega) \quad (2-53)$$

迭代公式如下:

$$\omega_{t+1} = \omega_t - \eta_{t+1} \nabla f(\omega) = \omega_t - \frac{\eta \sum_{i=1}^n f_i(\omega)}{n} \quad (2-54)$$

其中, ω_t 表示着第 t 次迭代的权重, η_t 表示着步长因子。

较为常见的机器学习中的随机梯度下降法优化损失函数有以下几种:

感知器:

$$\min_{\omega} \frac{1}{n} \sum_{i=1}^n \max \{0, -y_i \omega^T x_i\} \quad (2-55)$$

迭代公式如下:

$$\omega_{t+1} = \omega_t + \eta_t \begin{cases} y_t x_t & y_t \omega^T x_t < 0 \\ 0 & \text{otherwise} \end{cases} \quad (2-56)$$

支持向量机:

$$\min_{\omega} \frac{\lambda}{2} \|\omega\|^2 + \frac{1}{n} \sum_{i=1}^n \max \{0, 1 - y_i \omega^T x_i\} \quad (2-57)$$

迭代公式如下:

$$\omega_{t+1} = \omega_t - \eta_t \begin{cases} \lambda \omega_t - y_t x_t & y_t \omega^T x_t < 0 \\ \lambda \omega_t & \text{otherwise} \end{cases} \quad (2-58)$$

套索 (Least absolute shrinkage and selection operator, Lasso) 模型:

$$\min_{\omega} \lambda \|\omega\|_1 + \frac{1}{n} \sum_{i=1}^n \frac{1}{2} (\omega^T x_i - y_i)^2 \quad (2-59)$$

其中, $\omega = u - v$, $u \geq 0, v \geq 0$

迭代公式如下:

$$u_{t+1} = \max \{0, u_t - \eta_t (\lambda e + \omega^T x_t - y_t) x_t\} \quad (2-60)$$

2.4 本章小结

本章主要为深度学习方面的理论概论, 对本研究中所用到的深度学习中的深度置信网络、深度多层感知器以及循环神经网络进行相关理论的介绍, 介绍了这三种深度学习模型的结构、主要算法以及在训练过程中较为常用的加速算法。

第三章 基于深度学习的外汇市场汇率预测模型

本章包括以下三个部分，第一个部分是对外汇市场汇率预测方面的所应用的模型的流程以及相关参数进行估计，第二个部分是对样本数据的选取以及所应用的软件方法进行表述，第三个部分是根据模型所得到的不同外汇市场的汇率预测结果进行分析，对比得出模型的性能差异。

3.1 引言

汇率的波动幅度以及波动趋势影响跨国业务企业的经营决策、银行利率决策以及国家的货币政策等方面，对外汇市场的汇率进行高精度的预测有着重要的意义。本章针对外汇市场的汇率预测问题，提出了基于深度学习模型建立适合外汇市场汇率预测模型，将多个深度学习模型的实证预测结果进行对比，确定更加适合应用到汇率预测方面的模型，提高汇率预测方面模型的精准度。

3.2 基于深度学习的外汇市场汇率预测模型

本研究应用两种基于深度学习理论的模型构建外汇市场汇率预测模型，这两种模型分别是深度置信网络模型以及循环神经网络模型中的长短期记忆模型。并且将 ARMA 模型作为基准模型。该模型对不同主流国际汇率市场的历史数据进行预测，将结果与 ARMA 模型的预测结果进行集成处理，从而计算出相应外汇市场的汇率预测值。基于深度学习模型的外汇市场汇率预测模型包括数据处理、模型结构的确定、集成汇率预测三个步骤。

3.2.1 数据处理

在本研究中，使用深度学习模型之前，首先要对数据进行预处理，我们将原始数据分为两个部分，其中整体数据中的 70% 数据作为训练集来对深度置信网络、循环神经网络进行参数的估计，剩下的 30% 数据作为测试集来对模型的参数以及性能进行检验。

通过使用 TensorFlow 框架来实现深度置信网络、循环神经网络进行参数的估计，通过 AIC 准则函数对模型进行定阶，从而使得对模型的参数估计值进行合理的选择。

3.2.2 预测模型结确定

1、深度置信网络模型结构的确定

(1)应用 $r_t = \ln(P_t/P_{t-1})$ 来对历史数据进行预处理,其中 P_t 表示着 t 时期的日收盘价, r_t 表示着 t 时期的日收益率。从而得到 r_t 的时间序列数据。

(2)应用 ARMA 模型以及 DBN 模型对训练集数据进行训练。

对于时间序列 X_t 来说, ARMA 模型定义为^[165]:

$$\Phi(L)X_t = \Theta(L)\varepsilon_t \quad (3-1)$$

其中, $\Phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p$, $\Theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q$ 。L 表示着延迟算子, $L^p X_t = X_{t-p}$, $\{\varepsilon_t\}$ 表示着白噪声。

(3)通过 AIC 准则函数对模型进行定阶, AIC 定义为如下形式:

$$AIC = \ln \hat{\sigma}_a^2 + \frac{2(p+q)}{N} \quad (3-2)$$

通过 AIC 准则函数对模型进行定阶,得到滞后阶数为 2,也就是说将过去 1 天、过去 2 天的历史数据作为模型的输入数据,来对当日的汇率进行预测。

(4)将 DBN 模型的隐含层层数设置为 2 层,输入节点设置为 100 个,神经元的激活函数选用线性校正函数(Rectified Linear Units),将模型中的学习速率设置为 0.01,通过 2000 次的循环训练得到权重等相关参数值,从而得到 DBN 模型的网络结构。

2、循环神经网络模型结构的确定

(1)与 DBN 模型相似,同样应用 $r_t = \ln(P_t/P_{t-1})$ 来对历史数据进行预处理,从而得到 r_t 的时间序列数据。

(2)应用 ARMA 模型以及循环神经网络中的长短时间记忆网络模型(Long Short Term Memory Networks, LSTM)对训练集数据进行训练。

(3)通过 AIC 准则函数对模型进行定阶,得到滞后阶数为 2,也就是说将过去 1 天、过去 2 天的历史数据作为模型的输入数据,来对当日的汇率进行预测。

(4)本研究中,将 LSTM 模型的输入层层数设置为 1 层,节点设置为 2 个,输出层层数设置为 1 层,节点设置为 1 个,隐含层层数设置为 2 层,每层节点设置为 4 个,通过 2000 次的循环训练得到权重等相关参数值,选用自适应矩估计(Adaptive Moment Estimation, ADAM)方法作为模型的优化算法,ADAM 算法定义如下:

$$(\theta_i)_{t+1} = (\theta_i)_t - \frac{\alpha \bar{W}_t}{\sqrt{\hat{V}_t + \varepsilon}} \quad (3-3)$$

其中, $W_t = \rho_1 W_{t-1} + (1 - \rho_1)g_t$, $\bar{W}_t = \frac{W_t}{1 - \rho_1^t}$, $V_t = \rho_2 V_{t-1} + (1 - \rho_2)g_t^2$, $\hat{V}_t = \frac{V_t}{1 - \rho_2^t}$, $(\theta_i)_{t-1}$ 是 t 步的参数, W_t 表示着梯度的带权平均, V_t 表示着带权有偏方差, ρ_1 和 ρ_2 表示着衰减因

子。

选用均方误差方法 (Mean Squared Error, MSE) 作为模型的目标损失函数, MSE 方法定义如下:

$$MSE = \frac{1}{N} \sum_{i=1}^N [\widehat{f_i(x)} - f_i(x)]^2 \quad (3-4)$$

其中, N 表示着观测值的数量, $\widehat{f_i(x)}$ 表示着预测值, $f_i(x)$ 表示着真实值, MSE 预测精度为 10^{-5} 。

3.2.3 基于深度学习的集成汇率预测

1、深度置信网络的汇率预测值

将已知结构的 ARMA 模型以及 DBN 模型用于测试集数据的预测, 从而得到预测值。直接应用 DBN 模型对测试集进行预测, 发现预测模型出现过度拟合现象, DBN 网络呈现过度饱和。考虑到数据集可能同时存在非线性和线性的特征, 所以选择将集成算法与 ARMA 模型与 DBN 模型相结合的方法对结果进行处理, 从而得到最后的预测值。本研究中假设历史数据中非线性特征以及线性特征的影响程度是一样的, 基于 ARMA 模型与 DBN 模型的集成公式定义如下:

$$Y_t = \frac{\alpha_t + \beta_t}{2} \quad (3-5)$$

其中, Y_t 表示着最终的预测值, α_t 表示着 ARMA 模型输出的预测值, β_t 表示着 DBN 模型输出的预测值。

2、循环神经网络模型的汇率预测值

将已知结构的 ARMA 模型以及 LSTM 模型用于测试集数据的预测, 从而得到预测值。本研究中设置历史数据中非线性特征以及线性特征的影响程度是一样的, 所以集成方法下的预测结果定义如下:

$$Y_t = \frac{\alpha_t + \delta_t}{2} \quad (3-6)$$

其中, Y_t 表示着最终的预测值, α_t 表示着 ARMA 模型输出的预测值, δ_t 表示着 LSTM 模型输出的预测值。

3.3 实证研究

3.3.1 实验数据

本研究选取了国际主要的 7 个汇率市场的历史数据对深度学习模型在外汇市场的汇率预测能力进行研究，所有的实验数据都是从 Quandl 数据源网站 (<https://www.quandl.com/>) 下载得到。

这 7 个国际汇率市场分别是人民币 (RMB)、欧元 (Euro)、日元 (Jap)、加拿大美元 (Cad)、澳洲美元 (Aud)、英镑 (Gbp)、瑞士法郎 (Chf) 国际汇率市场。每个汇率市场的数据时间跨度是从 2007 年 7 月 23 日到 2017 年 2 月 24 日为止，每个市场中一共包括 2409 个观测数据作为实验数据。

在本研究中使用 7 个汇率市场的原始数据进行深度学习模型的训练和预测工作，应用 $r_t=\ln(P_t/P_{t-1})$ 来对历史数据进行预处理，其中 P_t 表示着 t 时期的日收盘价， r_t 表示着 t 时期的日收益率。

在本研究中，通过使用 Google (谷歌) 提出的 TensorFlow 框架来实现深度学习模型进行外汇市场汇率预测的分析过程，通过不同国际汇率市场的预测结果对比，从而确定预测性能最好的深度学习模型的结构。并且用均值、标准差、偏度、峰度等描述性统计量以及 Jarque-Bera 正态性检验与 BDS 检验等对 7 个国际主流汇率市场的历史数据进行描述性统计分析。

3.3.2 实证结果分析

表 3-1 7 个外汇市场描述性统计分析

Fig3-1 Descriptive statistical analysis of 7 foreign exchange markets

	Aud	Cad	Chf	Euro	Gbp	Jap	RMB
均值($\times 10^{-5}$)	2.993	2.9253	-17.883	-0.48736	-12.791	-9.4861	-11.698
方差	0.0104	0.0071	0.0076	0.0069	0.0069	0.0073	0.0012
偏度	-0.6732	-0.0677	0.7606	0.2156	-0.3729	-0.2817	0.0118
峰度	14.3532	8.5443	16.444	5.988	9.0811	7.7138	17.4163
JB 检验	0.001	0.001	0.001	0.001	0.001	0.001	0.001
BDS 检验	0	0	0	0	0	0	0

通过表 3-1 的汇率市场描述性统计分析，可以得到以下结论：

通常情况下，汇率预测的历史数据时间序列因为内部有着非常复杂的线性以及非线性的特征，所以有着非正态分布的特征，Jarque-Bera 测试表明历史数据时间序列是非正态分布的，而 BDS 测试表明历史数据时间序列不是独立分布的。

通过汇率一阶差分的均值以及标准差可知，澳洲美元市场历史数据的变动幅度较

大，数据较为分散。加拿大美元市场、瑞士法郎市场、欧元市场、英镑市场、日元市场历史数据的历史数据较为集中，离散程度小。在 7 个汇率市场中人民币市场的历史数据标准差最小，说明该市场汇率变动幅度较小，数据最为集中。

通过偏度和峰度指标可知，当偏度为 0，峰度为 3 时数据具有正态分布的特征，由此可知，澳洲美元市场、加拿大美元市场、英镑市场以及日元市场的历史数据具有尖峰左偏特征，而瑞士法郎市场、欧元市场以及人民币市场的历史数据具有尖峰右偏特征。

通过 JB 测试以及 BDS 测试可知，7 个外汇市场的 JB 值接近于 0，并且 BDS 值为 0，说明本研究中所应用的汇率市场历史数据均拒绝原假设，都是具有非正态分布、非独立分布、非线性的特征。

通过描述性统计分析的结果可知，本研究中所选择的历史数据不是独立分布的，具有非线性、非正态分布的特征。

表 3-2 7 个外汇市场汇率预测结果

Fig3-2 Results of exchange rate forecasting in 7 foreign exchange markets

汇率预测模型	Aud	Cad	Chf	Euro	Gbp	Jap	RMB
ARMA	4.9356	3.1017	6.3702	3.6471	4.2829	4.3479	0.34035
DBN	4.9391	3.1009	6.2669	4.2764	4.2661	4.3490	0.33356
LSTM	4.9897	3.0836	6.3020	3.6844	4.9176	4.4498	0.34703

通过表 3-2 的预测结果，可以得到以下结论：

与基准模型 ARMA 模型相比，DBN 模型以及 LSTM 模型在各个外汇市场的汇率预测结果大都低于 ARMA 模型的预测结果，说明两个模型在国际汇率市场汇率预测的性能都优于基准模型。

对同一国际汇率市场来说，在澳洲美元市场、瑞士法郎市场、英镑市场、日元市场以及人民币市场这五个国际汇率市场中，DBN 模型的预测结果低于 LSTM 模型的预测结果，说明在这五个国际汇率市场中 DBN 模型的预测性能要优于 LSTM 模型。在加拿大美元市场、欧元市场这两个国际汇率市场中，LSTM 模型的预测结果低于 DBN 模型的预测结果，说明在这两个国际汇率市场中，LSTM 模型显示了较为出色的预测性能。

就同一预测模型来说，DBN 模型在不同国际汇率市场的预测结果与 LSTM 模型的预测结果整体上较为接近，但是在欧元市场中，DBN 模型的预测结果高于 ARMA 模型，说明在这个国际汇率市场中 LSTM 模型的预测性能更加出色。在澳洲美元市场

和日元市场中, DBN 模型与 LSTM 模型的预测结果均高于 ARMA 模型的预测结果, 说明这两个模型在这两个国际汇率市场中的预测性能不是十分出色。在欧元市场、英镑市场以及人民币市场这三个国际汇率市场中, LSTM 模型的预测结果高于 ARMA 模型, 说明在这三个国际汇率市场中 DBN 模型的预测性能更加出色。在瑞士法郎市场中, 两个模型的预测结果均低于 ARMA 模型的预测结果, 说明在这个市场中两个模型的预测性能均较好, 但是 DBN 模型的预测性能更好。

整体来说, 总结 7 个汇率市场中 DBN 模型和 LSTM 模型的预测结果, 整体上 DBN 模型的预测性能优于 LSTM 模型, DBN 模型更加适合对外汇市场的汇率进行预测。

3.4 本章小结

本章在第二章深度学习模型的基础上, 构建了基于深度学习模型的外汇市场汇率预测模型。在分别用深度置信网络、循环神经网络以及 ARMA 模型对主流国际外汇市场的历史数据进行汇率预测之后, 用集成方法对深度学习模型的预测结果以及 ARMA 模型的预测结果进行处理, 从而得到最终的外汇市场汇率预测值。本章提出的汇率预测模型在不同的国际外汇市场中对汇率进行预测, 与传统的汇率预测模型相比, 本章所提出的汇率预测模型性能更好。实证结果显示所提出的基于深度学习模型的汇率预测模型在外汇市场汇率预测方面有着较好的性能。

第四章 基于深度学习的集成 VaR 风险测度模型

本章包括以下三个部分，第一部分对 VaR 理论及其度量方法进行了简要的介绍，第二部分对基于深度学习模型的 VaR 风险测度过程进行了介绍，第三部分对深度学习模型的实证结果进行对比，对模型的性能进行评价。

4.1 引言

汇率市场历史数据内部的关系错综复杂，同时存在着线性以及非线性的内在关系，所以汇率的波动幅度以及变化趋势难以预测，国际间的经济贸易活动因为汇率的不可预测的波动有了不确定性，金融领域决策者们的决策也以为汇率的波动有所改变，从而产生了相应的风险。常用的风险度量模型难以对历史数据的内部关系进行完全的描述，所以传统模型对外汇市场的风险估计的精准度还是有着提高空间，所以对外汇市场进行风险的相关估算是十分具有研究价值的研究方向。本章内容中针对外汇市场风险度量问题，提出了应用集成算法将 ARMA-GARCH 模型与深度学习模型进行结合的方法，通过数据处理、分风险估计、总风险集成三个步骤对 7 个国际上主流外汇市场的汇率风险进行相关的度量研究，用实证数据对模型的性能进行验证寻找出更加适合外汇市场风险度量方面研究的模型，从而构建出一个能够提高风险估计精准度的模型。

4.2 VaR 风险度量理论

P. Jorion 认为 VaR 所度量的风险是在市场不发生异常变化的情况下，在一定的持有期限内和某一个置信水平下，资产或者投资组合可能发生的最大的损失值。VaR 近些年来被研究学者们广泛用到风险测度领域中，在金融机构内部对风险评价、风险控制和风险管理方面都有所帮助，成为风险测度的一个常用方法^[67]。

VaR 定义为：在给定的某一个置信水平下以及一定的持有期限内，资产可能发生的最大损失值。这也就是说一定持有期限 t 和概率 α 条件下，在已经给定的持有期限 t 内，资产或者投资组合可能发生的损失值超过 VaR 的概率为 α 。VaR 值的估算公式定义如下^[86]：

$$\text{VaR} = [-\mu + \sigma Z_{\alpha}] X \sqrt{h} \quad (4-1)$$

其中， X 表示着相关资产的价值， σ 表示着收益率的标准差， μ 表示着收益率的均值。 Z_{α} 表示着标准正态分布下的分位数， h 表示着持有期， $cl = 1 - \alpha$ 表示着置信

水平。

VaR 的计算方法目前可以分为参数方法 (Parametric approach)、半参数方法 (Semi-parametric approach) 以及非参数方法 (Non-parametric approach) 三种方法。VaR 计算方法中的非参数方法在计算过程中并不需要对任何参数进行估计;而参数方法则需要在计算过程中估计样本的均值、样本方差或者估计其他相关的参数;半参数法与参数法相似,都是需要对相关参数进行估计,参数法是需要把所有的参数都估计出来,但是半参数法则只需要估计部分参数^[166-167]。

参数法对 VaR 值进行估算的前提是假设汇率的历史数据的收益率序列是服从某种分布的,然后估算概率分布相应中的参数值。这种方法也可以认为是在对汇率历史数据序列的均值以及方差进行估算。VaR 的参数方法中应用较为广泛的参数方法为方差-协方差法、指数加权移动平均法 (Exponentially Weighted Moving Average, EWMA) 以及自回归条件异方差 (GARCH) 模型几种^[166]。

应用方差-协方差的方法对 VaR 值进行估计的前提是汇率历史数据序列的收益率服从正态分布,然后估计收益率的均值、标准差以及相关参数,然后通过估计出的相关参数值对 VaR 值进行估算。

非参数方法,也就是一种在估算 VaR 值的过程中并不需要对任何参数进行估计的方法。历史数据模拟法 (Historical Simulation) 和蒙特卡洛模拟法 (MonteCarlo Simulation) 这两种方法是非参数法中较为经典并且应用广泛的方法^[169]。

半参数法方法是介于参数法和非参数法之间的一种计算 VaR 的方法。半参数法试图在参数法以及非参数法之间寻求一种计算的平衡模式,是研究学者们近来较为常用的一种研究方法。与其他两种方法相比,半参数法更适合对具有着偏峰厚尾特征以及数量较多的极端观测点的汇率历史数据进行相关描述,这在极端情况下对 VaR 值进行估算有着十分重要的影响。非参数法方法在估算过程中引入了非参数技术,这种方法不仅打破了参数法估算过程中对假设的限制,而且也同时保留了参数法中数学稳健、便捷的特性。

GARCH 模型是目前主流的参数 VaR 估计模型。GARCH 模型假设误差项服从某种分布,在这种假设条件下应用不同的 GARCH 类模型对波动率进行估计。GARCH (q, p) 模型可以定义为^[168]:

$$\sigma_t^2 = \omega + \alpha_1 x_{t-1}^2 + \dots + \alpha_p x_{t-p}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_q \sigma_{t-q}^2 \quad (4-2)$$

其中, $\omega, \alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_q$ 的取值范围均大于 0。 Σ_{t-1}^2 表示着第 t-1 期的方差, σ_t^2 表示着第 t 期的方差, x_{t-1} 表示着第 t-1 期的收益率, x_{t-p} 表示着第 t-p 期的收益率。

GARCH 模型中较为常用的模型是 GARCH (1,1) 模型, GARCH (1,1) 模型定

义为:

$$\sigma^2 = \omega + \alpha_1 x_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \quad (4-3)$$

其中, $\omega \geq 0$, $\alpha, \beta \geq 0$, $\alpha + \beta < 1$ 。

VaR 的半参数方法是 VaR 估计方法的前沿研究, 具有很大的发展空间, 对于进一步提高 VaR 估计的准确性有很大的潜力。但是这部分的研究还不是十分的集中, 半参数法的主要方法包括极值理论、分位数回归方法、混合密度神经网络等方法。

4.3 基于深度学习的 VaR 风险测度模型

本研究将外汇市场的总风险转换成不同时间尺度上的分风险, 然后对所有时间尺度的分风险进行集成, 从而计算出外汇市场的总风险值。基于深度学习的 VaR 风险测度主要有数据处理、分风险估计以及总风险集成三部分。

4.3.1 数据处理

1、经验模态分解法 (Empirical Mode Decomposition, EMD)

经验模态分解法 (Empirical Mode Decomposition, EMD) 是通过数据本身的时间尺度特征将原信号逐级进行分解, 在这个过程中并不需要预先设定任何基函数。EMD 算法是将原信号分解成多个若干个相互独立的本征模函数 (Intrinsic Mode Function, IMF)。一个 IMF 具有的条件如下^[170]:

(1) 在整个时间范围内, 序列的局部极值点和过零点的数目必须相等, 或者数量相差 1;

(2) 在任意一个时刻点上, IMF 的局部最大值的包络线 (上包络线) 和局部最小值的包络线 (下包络线) 的平均值必须为零。

2、对于原始数据 $X(t)$, EMD 方法的分解过程可以表示为:

(1) 找出原始数据 $X(t)$ 的所有极值点;

(2) 用三次样条插值函数 (cubic spline function) 拟合形成原始数据 $X(t)$ 的最大值的包络 (上包络线) 和局部最小值的包络 (下包络线):

$$m(t) = \frac{m_{\max} + m_{\min}}{2} \quad (4-4)$$

其中, m_{\max} 表示着上包络线, m_{\min} 表示着下包络线, $m(t)$ 表示着上包络线和下包络线的平均值。

(3) 将原始数据 $X(t)$ 减去上包络线和下包络线的平均值 $m(t)$:

$$X(t)-m(t)=h(t) \quad (4-5)$$

其中, $h(t)$ 表示着原始数据 $X(t)$ 和上包络线和下包络线的平均值 $m(t)$ 的差额。

- (4) 判断 $h(t)$ 是否符合 IMF 的条件,若符合,则将 $h(t)$ 作为第 i 个 IMF, 记为 $c_i(t)$,用剩余序列 $r(t)=X(t)-h(t)$ 来代替 $X(t)$; 若不符合, 则用 $h(t)$ 代替 $X(t)$;
- (5) 一直重复步骤(1)到步骤(4), 直到剩余序列 $r(t)$ 称为单调函数或者小于预设值。

此时原始数据 $X(t)$ 可以用所有 IMF 以及剩余序列 $r(t)$ 的和来表示:

$$X(t)=\sum_{i=1}^n c_i(t)+r_n(t) \quad (4-6)$$

其中, $r_n(t)$ 表示着原始数据的平均值或者是平均趋势。

3、数据处理

将 $r_t=\ln(P_t/P_{t-1})$ 作为日收益率序列, 其中 P_t 和 r_t 分别代表 t 时期的日收盘价和日收益率, 应用 EMD 算法将时间收益率序列分解成若干组互相独立的 IMF, 记为 r_i 。

4.3.2 分风险估计

1、估计在不同时间尺度上每组 IMF 函数和残余序列的条件均值矩阵和条件标准差矩阵。

(1) 假设条件均值服从多尺度 ARMA 过程, 假设 x_1, x_2, \dots, x_N 是平稳序列 $\{X_t\}$ 的观测值, 则 ARMA 模型的均值 μ 定义为:

$$\mu=EX_t \quad (4-7)$$

其中, $\bar{\mu}=\frac{1}{N}\sum_{k=1}^N x_k$ 。

则条件均值 μ 可以由下式估计得到:

$$\hat{\mu}_{t,r}=a_{0,r}+\sum_{i=1}^s a_{i,r} r_{t-i,r}+\sum_{j=1}^m b_{j,r} \epsilon_{t-j,r} \quad (4-8)$$

其中, $\hat{\mu}_{t,r}$ 表示着 t 时期的条件均值, r 表示着 IMF 尺度, $r_{t-i,r}(i=1 \dots s)$ 表示着参数 $a_{i,r}$ 下的滞后为 s 时的返回值, $\epsilon_{t-j,r}(1 \dots m)$ 表示着 $b_{j,r}$ 参数下的滞后阶为 m 时的上一期残差。

(2) 假设条件标准差服从 GARCH 过程, 则条件标准差 σ 可以由下式估计得到:

$$\hat{\sigma}_{t,r}^2=\omega_r+\sum_{i=1}^p a_{i,r} \sigma_{t-i,r}^2+\sum_{j=1}^q \beta_{j,r} \epsilon_{t-j,r}^2 \quad (4-9)$$

其中, $\hat{\sigma}_{t,r}^2$ 表示着 t 时期的条件方差, r 表示着 IMF 尺度, $\sigma_{t-i,r}^2(i=1 \dots p)$ 表示着参数 $a_{i,r}$ 下的滞后阶为 p 时的方差, $\epsilon_{t-j,r}^2(j=1 \dots q)$ 表示着回报率为 0 期间参数 $\beta_{j,r}$ 下的滞后阶为 q 时的平方误差。

2、条件均值与条件标准差的集成

根据所得出的不同尺度下的 IMF 条件均值与条件标准差的值,应用基于深度学习模型的非线性集成算法计算出总的条件均值与条件标准差。

非线性集成方法定义为:

$$\hat{y}_t = \sum_{i=1}^N \omega_i \hat{x}_{i,t} \quad (4-10)$$

其中, \hat{y}_t 表示着 t 时刻的输出重构值, ω_i 表示着不同模型预测值的权重, $\hat{x}_{i,t}$ 表示着 t 时刻的预测值, n 表示着预测值的数量。

将集成方法应用到汇总条件均值和条件标准差方面, 则

条件均值定义如下:

$$\hat{\mu}_t = \sum_{j=1}^J \omega_j \hat{\mu}_{c_j(t)} \quad (4-11)$$

条件标准差定义如下:

$$\hat{\sigma}_t^2 = \sum_{j=1}^J \omega_j \hat{\sigma}_{c_j(t)}^2 \quad (4-12)$$

应用以 DBN 模型、MLP 模型以及 LSTM 模型为代表的深度学习模型, 可以计算出非线性集成算法中的权重参数。

4.3.3 总风险集成

假设持有期为一天, 汇率市场的时间序列服从均值为 μ , 方差为 σ 则 t 时期的总 VaR 值可以定义为:

$$\text{VaR}_t = \mu_t + \sigma_t G^{-1} \quad (4-13)$$

其中, G^{-1} 表示着累积正态分布的倒数。

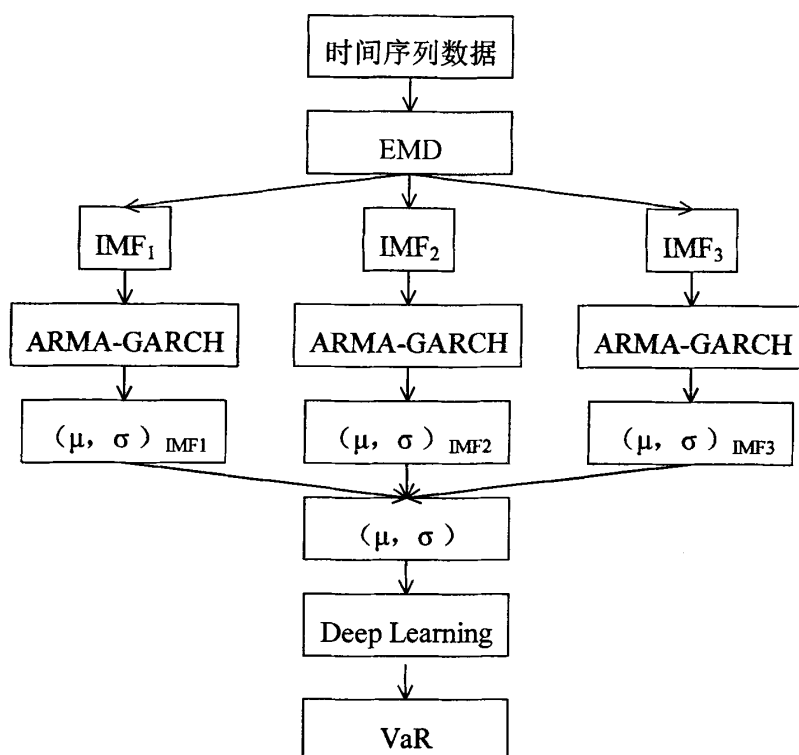


图 4-1 VaR 值计算的流程图

Fig4-1 Calculation of VaR values

4.4 实证研究

4.4.1 实验数据

选取了国际主流外汇市场中的 7 个汇率市场，选取的国际外汇市场包括人民币（RMB）、欧元（Euro）、日元（Jap）、加拿大美元（Cad）、澳洲美元（Aud）、英镑（Gbp）、瑞士法郎（Chf）国际汇率市场。所有的实验数据都是从 Quandl 公开数据源网站（<https://www.quandl.com/>）下载得到。

每个汇率市场的数据时间跨度是从 2007 年 7 月 23 日到 2017 年 2 月 24 日为止，每个市场中一共包括 2409 个观测数据作为实验数据。

在本研究中使用 7 个汇率市场的原始数据进行深度学习模型的训练和预测工作，应用 $r_t = \ln(P_t/P_{t-1})$ 来对历史数据进行预处理，其中 P_t 表示着 t 时期的日收盘价， r_t 表示着 t 时期的日收益率。

在本研究中, 先将测试集中的 70% 的数据设置为滚动窗口的长度, 然后将持有期假设为 1 天, 时间窗口每一步向前滚动一天, 剩余 30% 的数据作为样本外检测, 将历史数据用 EMD 方法分解成 8 组相互独立的 IMF。由此得到 218 个样本外检测值, 来对模型的估计效果进行评价。

4.4.2 评价指标

在本研究中用超出天数 (Exc) 以及 Kupiec 检验这两种较为常用的准确度测量方法对结果进行评价。

Exc 是预测得到的 VaR 值低于实际损失的观测点, 超出点的个数即为超出天数, 因此在理论上来说, 如果 Exc 越小, 则说明遭受损失的可能越小。但是在实际中, Exc 并不是越小越好, 因为这样的情况下就会高估了风险, 决策者需要预留更多资金, 就可能会造成资金的闲置与浪费。

Kupiec 检验法是目前来说在对模型的风险覆盖能力进行评价的一种常用的方法, 在这种方法中假设置信水平为 CL 下的样本区间为 T, 将超出天数记为 N, 可以得到超出率为 $\hat{p}=N/T$, 由此可知相对应的概率为 $p=1-CL$ 。超出数服从二项分布, 此时就可以认为每一个观测点都是独立的伯努利实验, 在这个方法中选用似然比检验, 则表达式定义如下^[17]:

$$LR=-2\ln\left[(1-p)^{T-N}p^N\right]+2\ln\left[(1-p)^{T-N}p^N\right] \quad (4-14)$$

如果 Kupiec 检验中的 P 值越大则可以表示研究中所用模型的风险覆盖能力越强。

4.4.3 实证结果分析

根据表 4-1 中的数据可以得出以下结论:

相对于基准模型 ARMA-GARCH 模型来说, MLP 模型、DBN 模型以及 LSTM 模型 7 个国际汇率市场的超出天数在不同置信度 95%, 97.5%, 99% 水平下更小, 对风险测度的预测更加保守。但是超出天数也不是越小越好, MLP 模型和 DBN 模型在瑞士法郎市场中的风险估计高估了风险, 使得过多的资本闲置, 造成经营成本的不必要增加。而且使得决策者在决策时的判断受到了影响。其中 MLP 和 DBN 模型的超出天数较为接近, 模型的性能也相似。

对同一国际汇率市场来说, 在澳洲美元市场、英镑市场中, MLP 模型 VaR 预测超出天数的平均值均小于 DBN 模型以及 LSTM 模型的超出天数的平均值, 说明在这两个国际汇率市场中 MLP 的风险预测性能高于另外两个模型。在加拿大美元市场中,

LSTM 模型 VaR 预测超出天数的平均值均小于 MLP 模型和 DBN 模型的超出天数的平均值，说明在该国际汇率市场中 LSTM 模型的风险预测性能高于另外两个模型。在瑞士法郎市场中，DBN 模型的 VaR 预测超出天数均小于 MLP 模型以及 LSTM 模型的超出天数的平均值，说明在该汇率市场中 DBN 模型的风险预测性能高于另外两个模型。在欧元市场、日元市场中，MLP 模型以及 DBN 模型的 VaR 预测超出天数相同，均小于 LSTM 模型，说明在这两个国际汇率市场中 MLP 模型以及 DBN 模型的风险预测性能较为相近，并且高于 LSTM 模型。在人民币市场中，DBN 模型以及 LSTM 模型 VaR 预测超出天数相同，均小于 MLP 模型，说明在这两个国际汇率市场中 DBN 模型以及 LSTM 模型的风险预测性能较为相近，并且高于 MLP 模型。

表 4-1 7 个外汇市场 VaR 风险度量结果

Fig4-1 Result of VaR in 7 foreign exchange markets

VaR 度量模型	CL	Aud	Cad	Chf	Euro	Gbp	Jap	RMB
ARMA-GARCH	95%	42	32	20	35	48	36	34
	97.5%	23	14	11	20	26	30	20
	99%	15	5	5	12	15	19	15
	平均值	26.67	17.00	12.00	22.33	29.67	28.33	23.00
MLP	95%	8	14	5	5	26	22	8
	97.5%	5	8	1	5	19	16	6
	99%	3	2	0	3	12	11	3
	平均值	5.33	8.00	2.00	4.33	19.00	16.33	5.67
DBN	95%	9	14	2	5	30	21	7
	97.5%	6	8	0	5	20	16	4
	99%	3	3	0	3	13	12	3
	平均值	6.00	8.33	0.67	4.33	21.00	16.33	4.67
LSTM	95%	12	10	8	7	34	21	7
	97.5%	8	4	5	4	22	17	4
	99%	4	4	3	4	14	12	3
	平均值	8.00	6.00	5.33	5.00	23.33	16.67	4.67

就同一模型来说，MLP 模型在澳洲美元市场、英镑市场、欧元市场、日元市场中风险预测性能较好，而 DBN 模型在瑞士法郎市场、欧元市场、日元市场以及人民币市场中的风险预测性能较好，，LSTM 模型只在加拿大美元以及人民币市场中的风险

预测性能较好。

整体来说，MLP 模型与 DBN 模型的风险预测性能较为接近，均在四个国际汇率市场中预测性能表现良好，说明这两个模型的风险预测性能均高于只在两个国际汇率市场性能较好的 LSTM 模型。但是 MLP 模型与 DBN 模型相比来说，DBN 模型的整体 VaR 预测超出天数更小，说明 DBN 模型在风险预测方面的性能高于 MLP 模型，也就是说 DBN 模型与另外两个模型相比更加适合风险预测方面。

表 4-2 7 个外汇市场 VaR 风险度量 P 值

Fig4-2 P value of VaR in 7 foreign exchange markets

VaR 度量模型		CL	Aud	Cad	Chf	Euro	Gbp	Jap	RMB
ARMA-GARCH	95%	0.0000	0.0000	0.0109	0.0000	0.0000	0.0000	0.0000	0.0000
	97.5%	0.0000	0.0019	0.0340	0.0000	0.0000	0.0000	0.0000	0.0000
	99%	0.0000	0.1005	0.1005	0.0000	0.0000	0.0000	0.0000	0.0000
	平均值	0.0000	0.0341	0.0485	0.0000	0.0000	0.0000	0.0000	0.0000
MLP	95%	0.3451	0.3552	0.0411	0.0411	0.0001	0.0023	0.3451	0.3451
	97.5%	0.8431	0.3005	0.0179	0.8431	0.0000	0.0002	0.8143	0.8143
	99%	0.5975	0.9011	0.0363	0.5975	0.0000	0.0000	0.5975	0.5975
	平均值	0.5952	0.5189	0.0318	0.4939	0.0000	0.0008	0.5856	0.5856
DBN	95%	0.5432	0.3552	0.0007	0.0411	0.0000	0.0051	0.1959	0.1959
	97.5%	0.8143	0.3005	0.0009	0.8431	0.0000	0.0002	0.5094	0.5094
	99%	0.5975	0.5975	0.0363	0.5975	0.0000	0.0000	0.5975	0.5975
	平均值	0.6517	0.4177	0.0126	0.4939	0.0000	0.0018	0.4343	0.4343
LSTM	95%	0.7364	0.7768	0.3451	0.1959	0.0000	0.0051	0.1959	0.1959
	97.5%	0.3005	0.5094	0.8431	0.5094	0.0000	0.0001	0.5094	0.5094
	99%	0.2672	0.2672	0.5975	0.2672	0.0000	0.0000	0.5975	0.5975
	平均值	0.4347	0.5178	0.5952	0.3242	0.0000	0.0017	0.4343	0.4343

根据表 4-2 中的 P 值可以得出以下结论：

与基准模型 ARMA-GARCH 模型的 P 值相比，MLP 模型、DBN 模型以及 LSTM 模型 7 个国际汇率市场的 P 值均大于基准模型，整体来说，MLP 模型，DBN 模型和 LSTM 模型的 VaR 风险度量模型预测性能都优于基准模型。

就同一国际汇率市场来说，在澳洲美元市场、日元市场中，DBN 模型的 P 值平均值均大于 MLP 模型以及 LSTM 模型，说明在这两个国际汇率市场中 DBN 模型的性能优于另外两个模型。在加拿大美元市场、人民币市场中，MLP 模型的 P 值平均值

均大于 DBN 模型以及 LSTM 模型，说明在这两个国际汇率市场中，MLP 模型的性能优于另外两个模型。在瑞士法郎市场中，LSTM 模型的 P 值平均值均大于 MLP 模型以及 DBN 模型，说明在该国际汇率市场中，LSTM 模型的性能优于另外两个模型。在欧元市场中，MLP 模型以及 DBN 模型的 P 值平均值均大于 LSTM 模型，说明在这该国际汇率市场中 MLP 模型以及 DBN 模型的性能较为接近，优于 LSTM 模型。在英镑市场中，三个模型的 P 值平均值都为 0，说明这三个模型在该国际汇率市场中的预测性能不是十分出色。

就同一模型来说，在澳洲美元市场、欧元市场、日元市场中，DBN 模型的预测性能较好，在加拿大美元市场、欧元市场、人民币市场中，MLP 模型的预测性能较好，在瑞士法郎市场中，LSTM 模型的预测性能较好，说明 DBN 模型以及 MLP 模型的风险预测性能较为接近，均在三个市场上性能较好。但是这三个模型均在英镑市场中的预测性能不是十分出色。

4.5 本章小结

本章构建了基于深度学习了外汇市场风险度量模型。在用 EMD 方法对历史数据进行不同时间尺度的分解后，将集成算法与 ARMA-GARCH 模型和深度学习模型相结合，通过用 ARMA-GARCH 模型计算不同时间尺度下的条件均值和条件标准差，然后用集成算法计算出总的条件均值和条件标准差，最后用深度学习模型计算出总体的 VaR 值。因为本章提出的风险度量方法在不同的国际汇率市场中对不同时间尺度的 IMF 分别计算风险，使得风险的度量更加准确，这与传统基准模型的结果相比，本章提出的风险度量方法性能更好。实验结果也说明深度学习模型在金融领域也能够有较好的性能。

第五章 总结和展望

近年来兴起的深度学习模型语音识别、图像识别、大数据特征提取自然语言处理等领域展现出了良好的性能和应用前景,但尚未在金融预测与风险测度领域得到足够的重视和应用。本研究将深度学习模型引入到汇率预测与风险测度领域,对国际主流外汇市场的汇率进行预测并且对风险进行测度。

本研究对外汇市场的汇率预测以及风险度量、深度置信网络模型、深度多层感知器模型、循环神经网络模型进行了综述,介绍了深度置信模型、深度多层感知器模型、循环神经网络模型的理论框架,介绍了这几种模型的学习算法,并且通过实验设计出了性能最好的网络结构。在此基础上,提出了基于深度学习的汇率预测方法,进行了实证,并对实证结果以及模型进行了分析。

汇率预测方面,本研究中用到了包括人民币、欧元、日元、加拿大美元、澳洲美元、英镑、瑞士法郎在内的7个国际主流汇率市场的历史数据。在训练阶段,利用训练集数据对深度置信网络模型以及循环神经网络模型中的长短期记忆模型进行训练,从而确定了预测模型的隐含层层数、结点数、学习速率等参数,构建出了预测性能较好的深度学习网络结构。在预测阶段,将测试集作为输入数据使深度学习模型对测试集进行预测,将结果与基准模型 ARMA 模型进行对比,从而对模型的预测性能进行判断。实证结果证明深度置信网络在汇率预测方面的性能优于循环神经网络。

在 VaR 风险度量方面,本研究中同样应用了7个国际主流汇率市场的历史数据。在数据处理阶段,运用 EMD 方法将历史数据分解成八组相对独立的 IMF。之后用 ARMA 模型对条件均值进行计算,用 GARCH 模型对条件标准差进行计算,最后应用深度置信网络模型、深度多层感知器模型以及循环神经网络中的长短时间记忆模型对不同时间尺度下的条件均值以及条件标准差进行汇总,从而计算出总的 VaR 风险值。实证结果表明深度置信网络模型的 VaR 风险度量模型预测性能优于另外两个模型。

在汇率预测方面还是在风险测度方面,深度置信网络模型均表现出稳定的性能和较好的预测效果。

本文将深度学习模型应用到了汇率预测与风险度量的应用中,取得了较好的效果,但是仍存在以下不足之处:

随着学者们对深度学习研究领域研究的不断深入,深度学习模型可应用到的领域也越来越多,研究学者们构建出的深度学习模型种类也逐渐丰富起来,自编编码器 (Auto-Encoder, AE)、循环神经网络 (Recurrent Neural Networks, RNNs)、卷积神经网络 (Convolutional Neural Networks, CNNs) 等模型都是目前较为常用的深度学习模型,然而在这些模型的基础上,研究学者们又构建出了深度置信网络 (Deep Belief

Networks, DBNs)、堆叠自编码器 (Stack Auto-Encoder, SAE)、长短时间记忆模型 (Long Short Term Mmemory, LSTM)、深度递归神经网络 (Deep Recurrent Neural Networks, DRNNs) 以及深度卷积神经网络 (Deep Convolutional Neural Networks, DCNNs) 等多种模型。

与现在广泛应用在其他领域的深度学习模型种类相比, 在本研究中应用到汇率预测与风险度量方面的只有深度置信网络模型, 循环神经网络模型以及深度多层感知器模型这三种模型, 还有种类繁多的深度学习模型没有应用到金融领域中。在以后的研究中将深度学习模型中较为常见的模型例如卷积神经网络模型等多种模型应用到金融领域是一个研究方向也是学者们关注的热点课题。

参考文献

- [1] 欧阳亮.基于小波分析与神经网络的汇率组合预测研究[D].湖南,湖南大学, 2008.
- [2] Chortareas G,Jiang Y,Nankervis J C.Forecasting exchange rate volatility using high-frequency data: Is the euro different?[J]. International Journal of Forecasting, 2007, 27(4):1089-1107.
- [3] 徐翔.从宏观调控角度看人民币升值对产业结构调整的作用[D].北京,对外经济贸易大学, 2010.
- [4] 王世飞.基于神经网络的汇率预测[D].大连,大连理工大学, 2009.
- [5] 左玮.汇率风险度量与管理[D].上海,华东政法大学, 2012.
- [6] 肖瑶.基于 VaR 模型的外贸企业外汇风险测算及对策研究[D]. 重庆,重庆理工大学, 2015.
- [7] 张海波,陈红.人民币汇率风险度量研究——基于不同持有期的 VaR 分析[J]. 宏观经济研究, 2012(12):25-31.
- [8] 王楠.基于 NDF 与 NARX 网络的人民币汇率预测研究[D]. 大连,大连理工大学, 2013.
- [9] 李亚南.汇率风险管理相关研究综述[J]. 北京金融评论, 2014(4).
- [10] 晁静.基于 DBN 的汇率预测研究[D]. 南京,南京大学, 2012.
- [11] Morales-Arias L,Moura G V.Adaptive forecasting of exchange rates with panel data[J]. International Journal of Forecasting, 2013, 29(3):493-509.
- [12] Shen F,Chao J,Zhao J.Forecasting exchange rate using deep belief networks and conjugate gradient method[J]. Neurocomputing, 2015, 167:243-253.
- [13] Korol T.A fuzzy logic model for forecasting exchange rates[J].Knowledge-Based Systems, 2014, 67(3):49-60.
- [14] Deboeck G J.Trading on the edge: neural, genetic, and fuzzy systems for chaotic financial markets[M]. John Wiley & Sons, Inc. 1994:54-61.
- [15] Abu-Mostafa Y S,Atiya A F.Introduction to financial forecasting[J]. Applied Intelligence, 1996, 6(3):205-213.
- [16] Meese R A,Rogoff K.Empirical exchange rate models of the seventies: do they t out-of-sample[J]. Journal of International Economics. 1983:3--24.
- [17] Engle R F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom Inflation[J]. Econometrica, 1982, 50(4):987-1007.
- [18] Bollerslev T. Generalized autoregressive conditional heteroskedasticity[J]. Journal of Econometrics, 1986, 31(3):307-327.
- [19] Hall S G,Miles D K,Taylor M P.Modelling asset price with time-carying betas [J].The Manchester School, 1989, 57(4):340-356.
- [20] Domowitz I,Hakkio C S.Conditional variance and the risk premium in the foreign exchange market[J]. Journal of International Economics, 1985, 19(1-2):47-66.
- [21] David A. Hsieh.Modeling heteroscedasticity in daily foreign-exchange rates[J]. Journal of Business & Economic Statistics, 1989, 7(3):307-317.
- [22] Engle R F. Estimates of the variance of U. S. Inflation based upon the ARCH model[J].Journal of the Malayan Branch of the Royal Asiatic Society. The Branch, 1988:286-301.

- [23] Jr K C H,Melvin M T.Sources of meteor showers and heat waves in the foreign exchange market[J]. Journal of International Economics, 1994, 37(3-4):239-247.
- [24] Nelson D B. Conditional Heteroskedasticity in asset returns: A new approach[J]. Econometrica, 1991, 59(2):347-370.
- [25] Glosten L R,Jagannathan R,Runkle D E.On the relation between the expected value and the volatility of thenominal rxcess teturn on dstocks[J].The Journal of Finance, 1993, 48(5):1779-1801.
- [26] Zakoian J M.Threshold heteroskedastic models[J].Journal of Economic Dynamics & Control, 1994, 18(5):931-955.
- [27] Obstfeld M,Taylor A M.Nonlinear aspects of goods-market arbitrage and adjustment: Heckscher's commodity points revisited,[J]. 1997, 11(4):441-479.
- [28] Mark P. Taylor, David A. Peel, Lucio Sarno. Nonlinear mean-reversion in real exchange rates: Toward a solution to the purchasing power parity puzzles[J]. International Economic Review, 2001, 42(4):1015–1042.
- [29] Venus Khim-Sen Liew, Terence Tai-Leung Chong, Kian-Ping Lim. The inadequacy of linear autoregressive model for real exchange rates: empirical evidence from Asian economies[J]. Applied Economics, 2003, 35(12):1387-1392.
- [30] Shen F,Chao J,Zhao J.Forecasting exchange rate using deep belief networks and conjugate gradient method[J]. Neurocomputing, 2015, 167:243-253.
- [31] Chortareas G,Jiang Y,Nankervis J C. The random-walk behavior of the Euro exchange rate[J]. Finance Research Letters, 2011, 8(3):158-162.
- [32] Hernandez Medina M A, Mendez G M. Modelling and Prediction of the MXNUSD Exchange Rate Using Interval Singleton Type-2 Fuzzy Logic Systems [Application Notes][J]. Computational Intelligence Magazine IEEE, 2007, 2(1):5-8.
- [33] Moosa I,Burns K.The unbeatable random walk in exchange rate forecasting: Reality or myth?[J]. Journal of Macroeconomics, 2014, 40(4):69-81.
- [34] Yun S,Dong W S.Forecasting the realized variance of the log-return of Korean won US dollar exchange rate addressing jumps both in stock-trading time and in overnight[J]. Journal of the Korean Statistical Society, 2015, 44(3):390-402.
- [35] Zorzi M C,Kocięcki A,Rubaszek M.Bayesian forecasting of real exchange rates with a Dornbusch prior [J]. Economic Modelling, 2015, 46:53-60.
- [36] Sermpinis G,Stasinakis C,Theofilatos K,et al.Modeling, forecasting and trading the EUR exchange rates with hybrid rolling genetic algorithms—Support vector regression forecast combinations[J]. European Journal of Operational Research, 2015, 247(3):831-846.
- [37] Barunik J,Kocenda E,Vacha L.Asymmetric volatility connectedness on forex markets[J]. Social Science Electronic Publishing, 2016.
- [38] Chortareas G,Jiang Y,Nankervis J C.The random-walk behavior of the Euro exchange rate[J]. Finance Research Letters, 2011, 8(3):158-162.
- [39] Lahmiri S.Wavelet transform,neural networks and the prediction of S&P price index: A comparative study of backpropagation numerical algorithms[J]. Business Intelligence Journal, 2012, 5(2).

- [40] 王松喜. 外汇市场有效性分析与 EGARCH 汇率预测[J]. 湖南人文科技学院学报, 2005(4):39-41.
- [41] 余乐安,汪寿阳,黎建强,等.基于 BPNN 和 Web 的智能外汇滚动预测与交易决策支持系统的开发[C].海峡两岸信息管理发展策略研讨会. 2005:109-114.
- [42] 汪寿阳,余乐安,黎建强.TEI@I 方法论及其在外汇汇率预测中的应用[J]. 管理学报,2007, 4(1):21-27.
- [43] 李艳丽,邓贵川,李辰阳.人民币汇率波动的预测——基于损失函数和 DM 检验的比较分析[J]. 国际金融研究, 2016, 346(2):84-96.
- [44] 宋博,陈万义.基于 HP 滤波和 ARMA-GARCH 模型的人民币汇率趋势预测[J].数学的实践与认识, 2017(1):70-78.
- [45] Engle R F,Bollerslev T.Modeling the persistence of conditional variances[J].Econometric Reviews, 1986, 5(1):1-50.
- [46] Nelson D B.Conditional heteroskedasticity in asset returns:A new approach[J].Econometrica, 1991, 59(2):347-370.
- [47] Baillie R T,Bollerslev T,Mikkelsen H O.Fractionally integrated generalized autoregressive conditional heteroskedasticity[J]. Journal of Econometrics, 1996, 74(1):3-30.
- [48] Adler M, Dumas B.Exposure to Currency Risk: Definition and measurement[J]. Financial Management, 1984, 13(2):41-50.
- [49] Sklar M. Fonctions de Répartition À N Dimensions Et Leurs Marges [J]. Publ.inst.statist.univ.paris, 1959, 8: 229-231.
- [50] Levonian M E. Market risk and bank capital: part 2[J]. Frbsf Economic Letter, 1994.
- [51] Paul Embrechts Ph.,Sidney I. Resnick,Gennady Samorodnitsky Extreme value theory as a risk management tool[J]. North American Actuarial Journal, 1999, 3(2):30-41.
- [52] Jondeau E, Rockinger M.The Copula-GARCH model of conditional dependencies: An international stock market application[J]. Journal of International Money & Finance, 2006, 25(5):827-853.
- [53] Palaro H P,Hotta L K. Using Conditional Copula to estimate value at risk[J].Social Science Electronic Publishing, 2006, 4(1):93-115.
- [54] Patton A J. Modelling Time-varying exchange rate dependence using the conditional Copula[J]. Ssrn Electronic Journal, 2001.
- [55] 惠晓峰,柳鸿生,胡伟,等. 基于时间序列 GARCH 模型的人民币汇率预测[J]. 金融研究, 2003(5):99-105.
- [56] 苏岩,杨振海. GARCH(1,1)模型及其在汇率条件波动预测中的应用[J]. 数理统计与管理, 2007, 26(4):615-620.
- [57] 王德全. 外汇风险度量研究——基于 GARCH 类模型及 VaR 方法[J]. 南方金融, 2009(8):11-15.
- [58] 骆珣,吴建红.基于 GARCH 模型的人民币汇率波动规律研究[J]. 数理统计与管理, 2009, 28(2):295-300.
- [59] 翟爱梅. 基于 GARCH 模型对人民币汇率波动的实证研究[J]. 技术经济与管理研究, 2010, 2010(2):20-23.
- [60] 李志斌,刘园. 基于 ARCH 类模型的人民币汇率波动特性分析[J]. 统计与决策,

2010(2):145-147.

- [61] 曲昶.人民币/美元汇率波动集群性研究[J]. 特区经济, 2011(11):89-91.
- [62] 刘瑾,施建淮. 基于 ARCH 类模型的 VaR 方法在外汇风险计量中的应用[J]. 国际金融研究, 2008(8):35-43.
- [63] 徐国祥. 金融统计学[M]. 格致出版社, 2016.
- [64] 徐国祥, 葛陈亮. 基于 Copula-DCC-EVT 模型的我国多元外汇资产组合风险精确度量[J]. 经济统计学:季刊, 2013(1).
- [65] 李鹏举. 基于 Copula-VaR 模型的外汇投资组合风险管理[J]. 苏州市职业大学学报, 2016, 27(1):40-45.
- [66] Degiannakis S, Potamia A. Multiple-days-ahead value-at-risk and expected shortfall forecasting for stock indices, commodities and exchange rates: Inter-day versus intra-day data[J]. International Review of Financial Analysis, 2016.
- [67] Jorion P. Risk 2 : Measuring the risk in value at risk[J]. Financial Analysts Journal, 1996, 52(6):47-56.
- [68] Engle R F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom Inflation[J]. Econometrica, 1982, 50(4):987-1007.
- [69] Bollerslev T. Generalized autoregressive conditional heteroskedasticity[J]. Journal of Econometrics, 1986, 31(3):307-327.
- [70] Glosten L R, Jagannathan R, Runkle D E. On the relation between the expected value and the volatility of the nominal excess return on stocks[J]. The Journal of Finance, 1993, 48(5):1779-1801.
- [71] Wang Z R, Chen X H, Jin Y B, et al. Estimating risk of foreign exchange portfolio: Using VaR and CVaR based on GARCH-EVT-Copula model[J]. Physica A Statistical Mechanics & Its Applications, 2010, 389(21):4918-4928.
- [72] Youssef M, Belkacem L, Mokni K. Value-at-Risk estimation of energy commodities: A long-memory GARCH-EVT approach[J]. Energy Economics, 2015, 51:99-110.
- [73] Nelson D B. Conditional heteroskedasticity in asset Returns: A new approach[J]. Econometrica, 1991, 59(2):347-370.
- [74] Boyle P, Broadie M, Glasserman P. Monte Carlo methods for security pricing[J]. Journal of Economic Dynamics & Control, 1997, 21(8-9):1267-1321.
- [75] Alexander G J, Baptista A M. Economic implications of using a mean-VaR model for portfolio selection: A comparison with mean-variance analysis[J]. Journal of Economic Dynamics & Control, 2002, 26(7-8):1159-1193.
- [76] Cabedo J D, Moya I. Estimating oil price 'Value at Risk' using the historical simulation approach[J]. Energy Economics, 2003, 25(3):239-253.
- [77] Artzner P, Delbaen F, Eber J H, et al. Thinking Coherently :Risk[J], Mathematical Finance, 1997, (10):33-49.
- [78] Rockafellar R T, Uryasev S. Optimization of Conditional Value-at-Risk[J]. Journal of Risk, 2010, 29(1):1071-1074.
- [79] Pflug G C. Some remarks on the Value-at-Risk and the conditional Value-at-Risk[M]. Probabilistic Constrained Optimization. Springer US, 2000:272-281.

- [80] Kibzun A I, Kuznetsov E A. Analysis of criteria VaR and CVaR [J]. Journal of Banking & Finance, 2006, 30(2):779-796.
- [81] Andersson F, Mausser H, Dan R, et al. Credit risk optimization with conditional Value-at-Risk criterion[J]. Mathematical Programming, 2001, 89(2):273-291.
- [82] Pownall R A J, Koedijk K G. Capturing downside risk in financial markets: the case of the Asian crisis[J]. Journal of International Money & Finance, 1999, 18(6):853-870.
- [83] Mulvey J M, Erkan H G. Applying CVaR for decentralized risk management of financial companies[J]. Journal of Banking & Finance, 2006, 30(2):627-644.
- [84] Batten J A, Kinateder H, Wagner N. Multifractality and value-at-risk forecasting of exchange rates [J]. Physica A Statistical Mechanics & Its Applications, 2014, 401(5):71-81.
- [85] Degiannakis, Stavros, George. Forecasting oil prices[J]. Mpra Paper, 2017.
- [86] 顾乃康. VaR:市场风险测定和管理的新工具[J]. 南方金融, 1998(1):9-10.
- [87] 李亚静, 朱宏泉, 何跃. 基于 VaR 的风险分析理论与计算方法[J]. 预测, 2000, 19(5):36-39.
- [88] 吴振翔, 叶五一, 缪柏其. 基于 Copula 的外汇投资组合风险分析[J]. 中国管理科学, 2004, 12(4):1-5.
- [89] 杨艳军, 王永锋. 期货市场交易保证金的设计[J]. 统计与决策, 2005(24):104-105.
- [90] 龚妮. GARCH 模型与 VaR 法在外汇风险度量中的应用[J]. 对外经贸, 2006(6):29-30.
- [91] 崔嵬. VaR 理念在银行间市场的运用[J]. 中国货币市场, 2003(8):50-51.
- [92] 娄可元, 周圣武, 胡素敏, 等. 基于非对称 GARCH 模型的美元/人民币汇率 VaR 分析[J]. 经济研究导刊, 2009(5):102-105.
- [93] 钟波, 汪青松. 基于 Bayes 估计的金融风险值——VaR 计算[J]. 数理统计与管理, 2007, 26(5):881-886.
- [94] 房小定, 吕鹏. 基于 GARCH 模型的上海同业拆借利率风险度量[J]. 西安电子科技大学学报(社会科学版), 2013(4):18-26.
- [95] 陆静, 杨斌. 商业银行汇率风险的 VaR—GARCH(1,1)模型计量[J]. 重庆大学学报(社会科学版), 2013, 19(5):66-72.
- [96] 苟红军, 陈迅, 花拥军. 基于 GARCH-EVT-COPULA 模型的外汇投资组合风险度量研究[J]. 管理工程学报, 2015, 29(1):183-193.
- [97] 刘飞, 郑晓亚. 商业银行汇率风险量化研究——基于正态分布与非对称拉普拉斯分布的在险价值测度[J]. 东北财经大学学报, 2015(4):83-89.
- [98] 李洛源. 基于 VAR 方法的上海银行间同业拆借利率风险度量[J]. 经营管理者, 2016(1):26-27.
- [99] J. M. Bates, C. W. J. Granger. The combination of forecasts[J]. Journal of the Operational Research Society, 1969, 20(4):451-468.
- [100] Cherkauer K J. Human expert-level performance on a scientific image analysis task by a system using combined artificial neural networks[J]. Working Notes of the Aaai Workshop on Integrating Multiple Learned Models, 1999:15--21.
- [101] Dietterich T G, Bakiri G. Solving multiclass learning problems via error-correcting output codes[J]. Journal of Artificial Intelligence Research, 1995, 2(1):263-286.
- [102] Breiman L. Bagging predictors[J]. Machine Learning, 1996, 24(2):123-140.
- [103] Parmanto B, Munro P W, Doyle H R. Improving committee diagnosis with resampling

- techniques[J]. *Revista De Libros*, 1995, 882-888(43-44):882-888.
- [104] 王丽丽. 集成学习算法研究[D]. 广西,广西大学, 2006.
- [105] [105]Sharkey A J C, Sharkey N E, Cross S S. Adapting an ensemble approach for the diagnosis of breast cancer[J]. 1998:281--286.
- [106] Hansen L K, Liisberg C, Salamon P. Ensemble methods for handwritten digit recognition[C].*Neural Networks for Signal Processing*. IEEE, 1992:333-342.
- [107] Gutta S, Wechsler H. Face recognition using hybrid classifier systems[C].*IEEE International Conference on Neural Networks*. IEEE, 1996:1017-1022 vol.2.
- [108] Schwenk H, Bengio Y. Boosting neural networks.[J]. *Neural Computation*, 2000, 12(8):1869-1887.
- [109] Hampshire J B, Waibel A H. The Meta-Pi network: connectionist rapid adaptation for high-performance multi-speaker phoneme recognition[C]. *International Conference on Acoustics, Speech, and Signal Processing*. IEEE, 1990:165-168 vol.1.
- [110] Shimshoni Y, Intrator N. Classification of seismic signals by integrating ensembles of neural networks[J]. *IEEE Transactions on Signal Processing*, 1998, 46(5):1194-1201.
- [111] Iyer R D, Lewis D D, Schapire R E, et al. Boosting for document routing[C]. 2000:70-77.
- [112] Singh H, Srivastava L. Modified differential evolution algorithm for multi-objective VAR management[J]. *International Journal of Electrical Power & Energy Systems*, 2014, 55(2):731-740.
- [113] 张秀艳, 徐立本. 基于神经网络集成系统的股市预测模型[J]. *系统工程理论与实践*, 2003, 23(9):67-70.
- [114] 宋星光, 夏利民. 基于 Bagging 算法的水库水沙联合智能调度[J]. *计算机工程与应用*, 2004, 40(25):218-219.
- [115] 孙冰, 宫宁生, 朱梧櫚. 基于覆盖的神经网络集成在语音识别中的应用[J]. *南京大学学报 (自然科学)*, 2006, 42(3):331-336.
- [116] 钱博, 李燕萍, 唐振民,等. 基于神经网络集成的说话人识别算法仿真研究[J]. *系统仿真学报*, 2008, 20(5):1285-1288.
- [117] 周水庚, 胡运发, 汪保友. 基于 Boosting 方法的中文文档层次分类[C].第 7 届中国机器学习学术会议. 2000.
- [118] DIAO Lili, HU Keyun, LU Yuchang,等. Improved stumps combined by boosting for text categorization 用 Boosting 方法组合增强 Stumps 进行文本分类[J]. *软件学报*, 2002, 13(8):1361-1367.
- [119] He K, Xie C, Chen S, et al. Estimating VaR in crude oil market: A novel multi-scale non-linear ensemble approach incorporating wavelet analysis and neural network[J]. *Neurocomputing*, 2009, 72(16-18):3428-3438.
- [120] He K, Lai K K, Yen J. Ensemble forecasting of value at risk via multi resolution analysis based methodology in metals markets[J]. *Expert Systems with Applications*, 2012, 39(4):4258-4267.
- [121] Srivastava R K, Greff K, Schmidhuber J. Training very deep networks[J]. *Computer Science*, 2015.

- [122] Ling Z H, Deng L, Yu D. Modeling spectral envelopes using restricted Boltzmann machines and Deep Belief Networks for statistical parametric speech synthesis[J]. IEEE Transactions on Audio Speech & Language Processing, 2013, 21(10):2129-2139.
- [123] Abdel-Hamid O, Mohamed A, Jiang H, et al. Applying convolutional neural networks concepts to hybrid NN-HMM model for speech recognition[C]. IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, 2012:4277-4280.
- [124] Weng C, Yu D, Watanabe S, et al. Recurrent deep neural networks for robust speech recognition[C]// IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, 2014:5532-5536.
- [125] Kovács G, Tóth L, Compermolle D V. Selection and enhancement of Gabor filters for automatic speech recognition[J]. International Journal of Speech Technology, 2015, 18(1):1-16.
- [126] Rafiei M H, Adeli H. A Novel Machine Learning model for estimation of sale prices of real estate units[J]. Journal of Construction Engineering & Management, 2015, 142(2):04015066.
- [127] Le Q V, Jaitly N, Hinton G E. A simple way to initialize recurrent networks of rectified linear units[J]. Computer Science, 2015.
- [128] Ding X, Zhang Y, Liu T, et al. Deep learning for event-driven stock prediction[C]. International Conference on Artificial Intelligence. AAAI Press, 2015:2327-2333.
- [129] Zeng H, Edwards M D, Liu G, et al. Convolutional neural network architectures for predicting DNA-protein binding[J]. Bioinformatics, 2016, 32(12):i121.
- [130] Zhu C, Yin J, Li Q. A stock decision support system based on ELM[J]. Journal of Computational Information Systems, 2014, 10(2):67-79.
- [131] Hirata T, Kuremoto T, Obayashi M, et al. Time series prediction using DBN and ARIMA[C]. International Conference on Computer Application Technologies. IEEE, 2015:24-29.
- [132] Dedinec A, Filiposka S, Dedinec A, et al. Deep belief network based electricity load forecasting: An analysis of macedonian case[J]. Energy, 2016, 115:1688-1700.
- [133] Akita R, Yoshihara A, Matsubara T, et al. Deep learning for stock prediction using numerical and textual information[C]. Ieee/acis, International Conference on Computer and Information Science. IEEE, 2016:1-6.
- [134] 王龙,杨俊安,陈雷,等.基于循环神经网络的汉语语言模型并行优化算法[J]. 应用科学学报, 2015, 33(3):253-261.
- [135] 孙志远,鲁成祥,史忠植,等.深度学习研究与进展[J]. 计算机科学, 2016, 43(2):1-8..
- [136] 张春云,秦鹏达,尹义龙.基于卷积神经网络的自适应权重 multi-gram 语句建模系统[J]. 计算机科学, 2017, 44(1):60-64.
- [137] 何炳金,宋海玉,孙东洋,侯建新,牛军海.基于深度学习的自动图像标注研究与实现[J].中国高新技术企业,2017,03:7—9.
- [138] 李传朋,秦品乐,张晋京.基于深度卷积神经网络的图像去噪研究[J]. 计算机工程, 2017, 43(3):253-260.
- [139] 严严,陈日伟,王菡子.基于深度学习的人脸分析研究进展[J]. 厦门大学学报:自然科学版, 2017, 56(1):13-24.
- [140] 迟元峰,顾敏.基于深度学习的人体行为识别研究[J]. 工业控制计算机, 2017, 30(1).

- [141] 田文慧,张来斌,胡瑾秋. 基于 DBN 的烟气轮机退化失效预警模型研究[J]. 设备管理与维修, 2015(21):332-334.
- [142] 周文杰,严建峰,杨璐.基于深度学习的用户投诉预测模型研究[J].计算机应用研究.2016,33.
- [143] 王声柱, 李永忠. 基于深度学习和半监督学习的入侵检测算法[J]. 信息技术, 2017(1):101-104.
- [144] 罗向龙, 焦琴琴, 牛力瑶,等. 基于深度学习的短时交通流预测[J]. 计算机应用研究, 2017, 34(1):91-93.
- [145] 邱元阳. 从 AlphaGo 看深度学习[J]. 中国信息技术教育, 2016(7):22-22.
- [146] Bengio Y. Learning deep architectures for AI[J]. Foundations & Trends® in Machine Learning, 2009, 2(1):1-127.
- [147] 郭丽丽, 丁世飞. 深度学习研究进展[J]. 计算机科学, 2015, 42(5):28-33.
- [148] 乔林,张雄伟,史海宁,贾冲.深度学习理论及其应用专题讲座(二)[J].军事通信技术,2016,37(1):98-104.
- [149] 王之仓. 多层感知器学习算法研究[D]. 苏州,苏州大学, 2006.
- [150] G.E.Hinton, T.J. Sejnowski. Optimal perceptual inference.Proc.IEEE Conf.on computer vision and pattern recognition[J],Washington DC,1983:488-453.
- [151] Gilberto Batres-Estrada ,Deep larning for multivariate financial time series[D],瑞典,瑞典皇家理工学院,2015.
- [152] Salakhutdinov,Ruslan. Deep Generative Models[D]. Toronto,University of Toronto,2009.
- [153] Geoffrey, Hinton. To recognise shapes, first learn to generate images. UTML TR 2006-004[D], Toronto ,University of Toronto,2006.
- [154] Cho K H, Raiko T, Ilin A. Gaussian-bernoulli deep boltzmann machine[C].International Joint Conference on Neural Networks. IEEE, 2013:1-7.
- [155] Chen, Murray, A.F. Continuous restricted boltzmann machine with an implementable training algorithm[J]. Vision, Image and Signal Processing, IEE Proceedings , 2003, 150(3):153-158.
- [156] 邓一沙. 多层感知器神经网络构建交易策略[D]. 成都,西南财经大学, 2012.
- [157] Krauss C, Xuan A D, Huck N. Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500[J]. European Journal of Operational Research, 2016.
- [158] Kohonen T. Self-organization and associative memory[J]. Springer-Verlag Berlin Heidelberg New York. Also Springer Series in Information Sciences, volume 8, 1988, 8(1):3406-3409.
- [159] 周国兵,最少门结构循环神经网络及其应用[D],南京,南京大学,2016.
- [160] 孙瑞奇. 基于 LSTM 神经网络的美股股指价格趋势预测模型的研究[D].北京,首都经济贸易大学, 2015.
- [161] 王腾飞. LSTM 神经网络在 IOS 平台上的实现[D].天津,南开大学, 2015.
- [162] Hestenes M R, Stiefel E. Methods of conjugate gradients for solving linear systems[J]. Journal of Research of the National Bureau of Standards, 1952, 49(6):409-436.
- [163] 周瑜. 非线性优化方法及在四维变分资料同化系统中的应用[D].上海, 华东师范大学, 2008.
- [164] 赵明旺. 基于遗传算法和最速下限法的函数优化混合数值算法[J]. 系统工程理论与实践, 1997, 17(7):59-64.

-
- [165] Zwicker C. Systematic relations existing between the properties of solid materials[J]. Physica, 1948, 14(1):35-47.
- [166] Engle R F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom Inflation[J]. Econometrica, 1982, 50(4):987-1007.
- [167] 张世法, 崔新文. 径流的趋势分析和概率预测[J]. 水科学进展, 1995, 6(1):22-28.
- [168] Bollerslev T. A Conditionally heteroskedastic time series model for speculative prices and rates of return[J]. Review of Economics & Statistics, 1987, 69(3):542-547.
- [169] He K, Lai K K, Xiang G. Portfolio value at risk estimate for crude oil markets: A multivariate wavelet denoising approach[J]. Energies, 2012, 5(4):1018-1043.
- [170] Lu X F, Lai K K, Liang L. Portfolio value-at-risk estimation in energy futures markets with time-varying copula-GARCH model[J]. Annals of Operations Research, 2014, 219(1):333-357.
- [171] Kupiec P H. Techniques for verifying the accuracy of risk measurement models[J]. THE J. OF DERIVATIVES, 1995, 3 (2) .

致谢

三年的硕士生活转瞬即逝，这期间我收获颇丰，母校浓厚的学术氛围以及良好的学习环境让我充满了感激与不舍。此次硕士论文从开始撰写到定稿历时一年多，最终能够顺利完成，我要感谢的人真的很多。

首先要感谢我的导师贺凯健教授。贺老师认真的工作方式、严谨的学术态度，使我由衷地佩服。在跟随老师一起进行的科学研究的过程中，老师渊博的知识和全面的思维也让我学到了很多，受益匪浅。硕士刚入学时，老师就与我讨论并规划了我的硕士论文研究方向。那之后我在学术方面就有了明确的目标，并循序渐进地为硕士论文做好了充分的准备。通过与老师不断地讨论与交流，我逐步完成了文献综述、模型构建以及实证研究等工作，在遭遇研究瓶颈时，老师根据我研究的程度以及未来的时间规划，及时给了我建议，让我转换方向，从而使得我脱离研究瓶颈，进入到另一个研究方向中，使得我的研究顺利进行，最终达到毕业标准。另外，除了学习方面，贺老师也教会了我们许多生活中做人的道理。总而言之，贺老师对我们而言亦师亦友。我非常庆幸能够成为贺老师的学生，老师的教导让我终身受益。

感谢我的父母，从小到大都给我精神上的支持，而且鼓励我多学习多提升自己，所以在我学生生涯中给了我莫大的支持，而且在我决定本科毕业后继续读硕士的时候，父母也给了我无限的支持与鼓励，让我一直可以做自己想做的事情，学习自己感兴趣的方向，感谢父母给了我学习的机会，让我更加珍惜现在的生活。

感谢工商管理专业所有的老师们，让我在课堂上收获了丰富知识，让我在专业实力方面更加扎实。

感谢和我一起上课、生活的朋友们，你们的陪伴让我的硕士生活更加丰富多彩，感谢你们给予我所有的支持和鼓励。

作者和导师简介

作者简介：

代文慧，女，蒙古族，1992年5月出生，内蒙古赤峰市人。2010年9月至2014年6月就读于北京化工大学经济管理学院国际经济与贸易专业，获得经济学学士学位。2014年9月至2017年6月在北京化工大学经济与管理学院工商管理专业攻读硕士学位。

联系方式：yizhiduxiu126@126.com

导师简介：

贺凯健，男，1977年出生，湖南科技大学商学院教授，硕士生导师，获香港城市大学商学院管理科学系哲学博士学位。主要研究领域为经济与金融预测、风险管理、金融工程、能源经济与决策支持、系统应急管理。

联系方式：kaijian.he@my.cityu.edu.hk

北京化工大学

硕士研究生学位论文答辩委员会决议书

研究生姓名： 代文慧 专业： 工商管理

论文题目： 基于深度学习的外汇市场预测和风险度量研究

指导教师姓名： 贺凯健 职称： 教授

论文答辩日期： 2017年5月18日 地点： 化纤楼508

论文答辩委员会成员

姓名	职称	工作单位	本人签名
吴军	教授	北京化工大学	吴军
王璇	副教授	北京化工大学	王璇
任继勤	副教授	北京化工大学	任继勤
吴卫红	副教授	北京化工大学	吴卫红
李小燕	副教授	北京化工大学	李小燕

注：此表用于存档，除本人签名务必用钢笔填写外，其余处必须用计算机打印。

答辩委员会对论文的评语（选题意义、文献综述、论文所取得的成果及水平、学风和论文写作水平、论文的不足之处）：

代文慧同学的硕士论文《基于深度学习的外汇市场预测与风险预测研究》选题合理，具有一定的理论及现实意义。

该论文作者在查阅大量相关文献的基础上，对外汇市场的汇率预测以及风险度量、深度置信网络模型、深度多层感知器模型、循环神经网络模型等进行了归纳分析，文献综述较为清晰全面。针对汇率走势的精确预测与汇率风险更精确估计，本文提出一种新的基于深度学习理论模型的外汇市场汇率预测模型，基于风险度量理论，将深度学习模型与 ARMA-GARCH 模型相结合，构建了集成的 VaR 模型，并应用 7 个主流外汇市场的汇率数据进行实证研究，证明本研究所提出的模型具有较好的精确性及可靠性。论文不足之处在于：第一，论文创新点归纳总结不足；第二，论文中总结部分语言表述不够清晰；第三，论文中少量序号标注、表格、图示不规范。

该论文立意较好，中心较为突出，结构合理，文献较为丰富，思路清晰，语言简洁，达到了硕士研究生水平。此外，代文慧同学的答辩表现较为良好，在答辩过程中较好地回答了评委提出的问题。

同意授予代文慧同学硕士学位。报院学位委员会讨论。

对学位论文水平的总体评价	优秀	良好	一般	较差
			√	

答辩委员会表决结果：

同意授予硕士学位 5 票，不同意授予硕士学位 0 票，弃权 0 票。根据投票结果，答辩委员会做出建议授予该同学硕士学位的决议。

答辩委员会主席签字：



2017 年 5 月 18 日