

深度学习的金融实证应用： 动态、贡献与展望

苏 治 卢 曼 李德轩

(中央财经大学, 北京 100081; 中国人民大学国际货币研究所, 北京 100872;
中国传媒大学, 北京 100024)

摘 要: 随着智能时代来临以及金融数据分析需求提升, 深度学习已经成为金融领域中的应用前沿。特别是在预测金融市场运动、处理文本信息、改进交易策略方面。深度学习包含深度神经网络、深度信念网络等多种结构, 通过分层结构提取深层特征, 强化重要因素、过滤噪音, 对提升预测准确率具有重要意义; 其应用及由此衍生的优化技术改进了金融领域预测分析方法, 促使实证研究范式从线性向非线性转变、从关注参数显著性向关注模型结构和动态特征转变, 同时为丰富金融经济理论做出贡献。构建结构合适、效果稳健的模型以捕捉金融数据有效特征并进行经济含义阐释是应用深度学习方法的难点与重点; 未来研究可以从挖掘深层经济意义、提炼一般性预测分析框架、探索其对异质信息的适用性等角度展开。

关键词: 深度学习; 金融市场预测; 文本挖掘; 深度神经网络; 深度信念网络

JEL 分类号: C45, G17, B41 **文献标识码:** A **文章编号:** 1002-7246(2017)05-0111-16

一、引 言

伴随着大数据积累、计算机并行计算能力的发展, 人工智能逐渐渗透到各个领域, 在经济社会发展中引导行业进行着深刻变革。而在变革过程中, 深度学习风靡于多样化的人工智能任务中, 成为解决人工智能发展瓶颈的关键技术。深度学习是一种基于多层神经网络的机器学习方法, 其由神经网络输出层和输入层以及两者之间的一系列堆叠隐藏层构成 (Schmidhuber, 2015; Bengio, 2009)。深度学习由于适合解决复杂性问题而引

收稿日期: 2016-12-27

作者简介: 苏 治 (通讯作者) 经济学博士, 教授, 中央财经大学, 中国人民大学国际货币研究所, Email: suzhi1218@163.com.

卢 曼 经济学博士研究生, 中央财经大学, Email: luman1007@sina.cn.

李德轩, 中国传媒大学, Email: sttam@163.com.

* 本文感谢国家哲学社会科学基金重大项目 (15ZDC024)、国家自然科学基金面上项目 (71473279) 资助。感谢匿名审稿人的宝贵意见。文责自负。

起了业界极大的关注,Google、百度、微软等高科技公司纷纷投入大量资金以期占领该项技术制高点。近年来,深度学习被广泛应用到人工智能任务中(如 AlphaGo),并在语音识别、计算机视觉以及自然语言处理方面取得了一系列先进成果(Prieto et al., 2016; Långkvist et al., 2014)。这得益于深度学习可以通过逐层的学习提炼出高抽象水平的、复杂的特征,并以此作为数据的表现形式(Najafabadi et al., 2015),从而最终提升分类或预测的准确性。深度学习有能力在全局路径中生成学习模式、处理数据中非临近数据之间的关系,它不仅提供了适合于人工智能应用的数据特征,还使得人工智能的最终目标“独立于人”成为可能(Najafabadi et al., 2015)。对金融领域而言,智能化是金融行业发展的方向,智能投顾、智能投研与智能投资都将是智能金融的表现形式,而这些形式实现的关键在于如何实现“智能”。深度学习在其他领域的成功应用为智能金融的发展提供了良好的借鉴。因此,随着智能时代的来临以及金融数据分析需求的提升,深度学习成为了金融领域中的应用前沿。

另一方面,金融市场是一个嘈杂的、具有非参数特点的动态系统(Huang and Tsai, 2009; Cavalcante et al., 2016),其往往具有一系列复杂而非线性的特征,对金融数据进行分析与预测一直是金融领域中一项极具挑战性的工作。现有分析预测方法在不同的程度上都体现出一定的“不适性”:传统计量方法或含参数的方程并不适合用于分析复杂、高维度、具有噪音的金融市场数据序列,传统人工神经网络方法的非线性特点尚不具备对如此复杂的数据进行准确建模的能力(Långkvist et al., 2014),而传统机器学习方法的表现很大程度上又需要于依赖人工特征设计(Arel et al., 2010),这将对结果造成一定的干扰。同时,这些方法在应用过程中存在着过拟合、收敛慢等问题,其不再遵循传统金融计量学基础设定,在应用方面与金融研究者存在着一定的距离(Dixon et al., 2016)。而 Hinton and Salakhutdinov(2006)提出的深度学习方法为金融数据分析提供了一个新的思路。深度学习会像人脑处理数据那样考虑信息,通过一系列隐藏层加强传统神经网络进而提升预测能力(Fehrer and Feuerriegel, 2015)。深度学习在金融领域中的应用不仅在一定程度上缓解了上述分析与预测的难题,更带来了金融学乃至经济学实证分析范式的改变。本文将在梳理已有深度学习在金融领域应用研究的基础上,论述其对金融数据的适用性以及对改进传统方法的贡献;深入分析深度学习所带来的金融领域实证分析范式的转变;进而剖析深度学习方法在金融领域应用的难点与重点,并指出其在未来金融领域应用中可能的研究方向。

二、深度学习在金融领域中主要应用研究议题

深度学习方法是一类模型结构的总称,常用的深度学习结构有深度神经网络、栈式自动编码器、深度信念网络、深度卷积神经网络、循环神经网络等。对具体模型结构而言,深度神经网络中每一层的权重和阈值可以由监督学习或者无监督学习过程来决定(Bengio et al., 2013),其他结构的权重决定方法与之相类似。栈式自动编码器和深度信念网络往

往由传统机器学习中的自动编码器和受限玻尔兹曼机串联而成(Najafabadi et al. 2015) , 这种结构具有能够对大量数据进行无监督学习的能力(Hinton and Salakhutdinov 2006) ; 深度卷积神经网络由于具有局部感受野、权值共享、池化层以及层级结构使用上的特点, 被广泛应用于图片识别问题中; 循环神经网络由于可以展示动态时序行为而常被应用在语音处理方面(LeCun et al. 2015) , 长短期记忆模型是循环神经网络中应用较多的模型, 其可以处理极其复杂的长期时序动态信息(Heaton et al. 2016a) 。同时, 伴随着深度学习各种模型结构的发展, 衍生出了一系列的优化技术(Hinton and Salakhutdinov 2006; Qiu et al. 2014; Schmidhuber 2015) 。为了适应深度学习的训练, 解决传统反向传播算法和梯度下降法计算成本较高的问题(特别是在训练深度神经网络时尤为明显) (Dixon et al. , 2016) , Hinton et al. (2006) 发展出贪心逐层算法, 该算法极大地缩短了用于训练学习的时间(Prieto et al. 2016) 。此外, 在优化技术层面, 还发展出 Dropout 方法(Hinton et al. , 2012; Ba and Frey 2013) 用以提高深度学习模型应用的泛化能力(Schmidhuber 2015) ; 正则化(Srivastava et al. 2014) 、快速训练技术(如用分段线性激活函数代替先验函数, Nair and Hinton 2010) 等方法加快深度学习模型的训练速度; Mordvintsev et al. (2015) 发展出了新的可视化技术能够让研究者更加深入地了解神经网络的操作。这一系列算法改进, 不仅仅是改善了预测效果, 对智能计算领域的相关研究也具有重要意义(Kuremoto et al. , 2014) 。已有部分学者将上述深度学习模型引入到金融领域研究中, 其主要的方向大致可以分为三个方面: 一是预测金融市场运动; 二是处理相关文本信息; 三是改进交易策略。

(一) 预测金融市场运动

深度学习在第一类应用中运用较为广泛, 如各类金融资产(股票、期货、债券等) 波动性预测、价格预测等, 研究重点在于深度学习的各种结构如何进行设计及调整以获得更好的预测结果。多数学者都将应用深度学习的预测效果与其他传统模型的预测效果进行了对比, 得出了深度学习方法能够提升模型预测准确率的结论。Xiong et al. (2015) 以可以代表公众情绪和宏观经济因子的谷歌国内趋势为切入点, 利用长短期记忆模型考察了 2004 年 10 月 19 日至 2015 年 7 月 24 日期间上述因素对 S&P500 波动性的影响。Shen et al. (2015) 将经典深度信念网络进行了改进, 使用连续受限玻尔兹曼机构建深度信念网络, 使该模型适用于连续数据, 从而用其对 GBP/USD、INR/USD、BRL/USD 三种汇率的周数据进行预测。Dixon et al. (2015) 将深度神经网络引入期货市场, 利用深度神经网络对芝加哥商品交易所 45 种大宗商品和期货的价格运动方向进行预测, 预测平均准确率达到 73%。在这一研究的基础上, Dixon et al. (2016) 进一步考虑了金融市场时间序列的时滞、移动平均数、相关性等因素的影响, 采用同样的方法预测 43 种大宗商品和外汇期货的 5 分钟间隔中间价的变动方向, 之后他们还进行了一个简单的长期交易回测。结果显示, 平均来看, 他们的模型对价格运动方向的预测准确率达到 42%, 在某些情况下预测率可以达到 68%; 回测结果显示利用这一模型可以获得高达 3.29 的年化夏普率。部分学者针对我国沪深股市进行了相关预测研究。曾志平等(2017) 将沪深股市所有股票的收盘

价特征转化为具有上升趋势、下降趋势以及无明显趋势的三类图片,利用深度信念网络对此进行训练,然后利用训练好的模型对 2000 年至 2014 年期间金融市场日度数据进行预测、筛选出可交易样本,根据信号进行交易,并统计交易成功与失败的样本数。研究表明,利用深度信念网络可以将金融时序数据量化决策分析中的准确率提升到 90.54%。张承钊(2016)也对沪深市场进行了研究,其选取沪深 300 指数为研究对象,构建了深度分合神经网络,将原始时间序列分解为多个级别的分解序列(他将多个分解的级数视为深度),用每一级别的序列训练一个神经网络,最后将所有神经网络(对应不同级别)进行整合,综合成一个整体的预测输出。这种建模方式与 ARIMA 模型、GARCH 族模型以及普通神经网络模型相比具有更低的预测误差。此外,张承钊(2016)还将该模型应用到我国汇率市场以及澳大利亚股指的预测上,也取得了较好的效果。可见,应用深度学习进行预测可以提升模型效果,但是也应该注意到,深度学习并不必然会产生高准确率;其虽然对金融市场数据具有传统方法所不具备的应用优势,但是针对具体的金融问题时,还是需要考虑金融市场环境特性。

在金融市场运动预测的应用中,还有一系列研究是针对限价指令簿展开的。传统报价驱动的股票市场已经由指令驱动交易平台所取代,命令处理时间上的快速下降引起了总体股票交易方式向高交易量、高频交易转化,指令驱动的交易系统特征改变了市场的动态性,因此有必要发展新的可以捕捉短期行为的研究方法(Bovier et al., 2006; Cartea et al., 2014)。Kercheval and Zhang(2015)回顾了针对限价指令簿的建模方法,认为其可以划分为两类:一种是统计模型,另一种是机器学习方法。他们认为由于限价指令簿的统计特性时刻处于变化中,而在应用统计方法进行分析时,往往是在有大量约束的条件下进行建模,所以模型预测结果受到了前提条件的限制(Cont 2011; Cont et al., 2010; Huang and Kercheval 2012; Roşu 2009; Shek 2011);而机器学习相对来说具有较少的约束,并且在预测方面有了一定的应用基础。已有大量学者基于多核学习、支持向量机等方法对限价指令簿进行研究(Fletcher and Taylor 2013; Kercheval and Zhang 2015; Kijima et al., 2016)。Kercheval and Zhang(2015)基于某一特定事件是否发生的预测,利用支持向量机对限价指令簿的股票中间价上升、下降以及不变进行建模预测。Sirignano(2016)受其启发,基于真实事件发生概率,以深度神经网络为基本结构框架构建了一种具有 D 维的空间深度神经网络模型,用以提取纳斯达克股票交易所限价指令簿的深层信息,进而实现价格预测。该空间深度网络模型在当期限价指令簿状态条件下,利用限价指令簿的动态结构特征,通过模拟限价指令簿中价格的局部行为生成一个低维的价格运动模型。该模型不仅可以用于对未来时间的最佳卖出价和最佳买入价的联合分布进行建模,也可以应用到限价指令簿的其他特征中,比如可以考虑最优卖价、最优买价、最优买入规模和最佳卖出规模的联合分布建模,还可以拓展用于任何 D 维数据分布的建模。通过以上研究, Sirignano(2016)进一步指出,由于空间深度神经网络能够更好地利用限价指令簿的深层信息,所以在信息分布尾部处理方面,其研究对风险管理应用来说具有特别的意义。

(二) 处理文本信息

第二类应用是研究如何利用深度学习对海量文本数据(如网页新闻、公司公告等)进行文本挖掘,从而实现了对股票或者事件的预测。金融市场预测分析经常要面对海量的异质性数据和海量的文本内容,后者也同时间序列一样对企业决策具有很大的影响。过去对财经新闻的决策支持主要依赖于传统的机器学习方法从新闻文本中提取简单的特征,如词袋、名词短语以及实体名词等(Nassirtoussi et al. 2014; Minev et al. 2012; Pang and Lee 2008; Fehrer and Feuerriegel 2015; Kogan et al. 2009; Schumaker and Chen 2009)。但是现有依据文本进行预测的框架几乎已经排除了传统机器学习方法:近期研究表明,传统的机器学习方法并不能有效的提取合适的特征,也不能有效捕捉复杂工作所需的非线性特征,而向传统神经网络模型中添加隐藏层来构建新的深度学习模型成为了解决这一问题的有效方法(Fehrer and Feuerriegel 2015)。

深度学习主要被应用于两类文本数据:一种是新闻文本,另一种是公司财务信息(如财务信息披露报告等)。Yoshihara et al. (2014) 首次利用循环神经网络和受限玻尔兹曼机构成的深度循环神经网络模型对新闻事件文本信息进行处理,从而预测股票市场趋势。他们不仅仅关注了与事件相关的语义信息,还考虑了过去重要事件的时间效应。Ding et al. (2014) 利用结构性事件代表新闻文本,然后利用标准前馈神经网络研究了事件和股票价格运动之间所隐藏的复杂关系。虽然这种提取特征的方式起到了一定作用,但是所得特征不能捕捉到语句中的结构性关系,限制了相关模型的应用能力(Ding et al. 2015)。基于此,Ding et al. (2015) 将张量神经网络与深度卷积神经网络相结合,利用所结合的模型处理文本信息,对 S&P500 以及从中选取的个股价格进行预测。在他们的研究中模型分为两部分,首先利用张量神经网络对文本中提取的事件进行事件嵌入训练,然后利用深度卷积神经网络捕捉事件影响、进一步建立事件短期影响和长期影响的模型。Ding et al. (2015) 研究认为张量神经网络的应用考虑到了事件的语义合成性,将词语进行有机结合,而不是像标准前馈神经网络那样只是简单的、不问原因的将它们结合起来;而深度卷积神经网络的应用可以对输入的事件序列执行语义分析,用卷积层提取最具有代表意义的全局特征,量化地分析历史事件的长期影响,在此基础上利用前馈神经网络将全局特征和股票趋势联系起来,从而提升模型效果。他们构建的模型与基准方法(标准前馈神经网络)相比,在 S&P500 和个股的预测准确率上有将近 6% 的提高。Ding et al. (2015) 指出,Luss and d'Aspremont(2015)、Ding et al. (2014) 所构建的基于新闻文本的预测模型在实际应用中,如果前一天没有关于个股的新闻,就不能预测出一天之后股价运动的趋势,因为他们没有考虑中期和长期的新闻,这虽然对结果的准确性影响较小,但是会较大地影响真实利润。与之相比,Ding et al. (2015) 的研究应用这一预测模型之后,盈利能力显著提高。

在处理财务信息方面,Fehrer and Feuerriegel(2015) 利用递归自动编码器预测 2004 年 1 月至 2011 年 6 月期间某只股票对已披露财务信息的反应。他们在 Hagenau et al. (2013)、Muntermann and Guettler(2007)、Pröllochs et al. (2014) 研究基础上,重点关注某

些特殊的新闻和异常回报之间的关系,采用以递归自动编码器为主体的深度学习模型进行预测。还有部分学者对财务报告文本中与风险相关的信息进行分析,试图进行风险预测。Wang and Hua(2014)利用收益报告文本数据中风险相关的语句预测了公司股票的波动性;Lischinsky(2011)利用标准语义分析工具对公司年报中的危机相关内容进行了分析。他们的研究都表明数据驱动的方法有助于提升预测准确性(Rönnqvist and Sarlin, 2016)。但是 Rönnqvist and Sarlin(2016)认为这种传统的文本分析明显忽略了语句顺序,仅仅在象征词义水平上进行操作。为了解决这一问题,他们将新闻文本和基本事件信息这两种不同类型的数据相结合,利用深度神经网络对欧洲银行危机事件进行预测。这样,他们所构建的模型可以区分出与事件相关的文本类型和语言类型,能够从大量文本数据的高度易变、稀疏的语言中推断出抽象特征,从而为预测提供支持。在他们的研究中,重点关注了银行破产、政府干预等 243 个事件,从三个层面(欧洲整体、国家个体、银行个体)考虑文本信息传达出的有关银行危机的信号,利用深度神经网络结构结合分布式记忆方法去刻画这种文本信息与危机事件之间的联系。

(三) 改进交易策略

深度学习在金融领域中的第三类应用是构建资产交易组合或加强已有投资策略(如动量策略)。虽然在股票市场交易中动量效应已经为人们所熟悉,但其仍然具有稳健的表现,并且被应用于国际股票、汇率、商品、债券等多种产品中(Asness et al., 2013),但是由于股票价格运动中含有大量的噪音成分,利用公众信息去预测股票收益是十分困难的。并且,随着投资者自身不断的学习和彼此之间为了交易收益而不断竞争,任何已经存在的交易模式都有必要改变。Takeuchi and Lee(2013)利用由栈式受限玻尔兹曼机构成的编码器提取个股历史特征,对动量效应进行强化。他们将研究限制在普通的股票交易中(纽约证券交易所、美国证券交易所或者纳斯达克市场),使用 Hinton and Salakhutdinov(2006)分析手写数字的方法对投资者使用的历史价格表进行可视化分析,然后采用栈式受限玻尔兹曼机构成的编码器来提取股票价格特征,将这些特征传递到一个正反馈的神经网络分类器中进行分析。通过这种加强动量策略的方法,他们在测试期内(1990 年 - 2009 年)获得了 45.93% 的年化收益率,而同期的基础动量策略收益率只有 10.53%;策略获得的累积财富值是 1919 美元(不考虑成本),而基础动量策略的只有 7.41 美元,整整高出 259 倍。Zhu et al.(2014)利用深度信念网络,结合振荡箱理论构建了一个自动股票决策支持系统。他们根据股票市场的运行情况将市场分为四类:纯粹牛市、纯粹熊市、有波动的牛市以及有波动的熊市。针对每一市场情形,均利用深度信念网络对 S&P500 中的 400 只股票历史交易数据进行学习,并预测下一时期股票价格的最高价和最低价,然后将预测出的最高价和最低价作为振荡箱的上界和下界;设定了一个检验规则以检验股票价格运动是否超过这一界限。其中,振荡箱的应用有助于消除局部低精度的影响,研究中还采用了灰色关联思想以提高精度。当价格突破振荡箱上界时买入股票,突破下界时卖出股票。Zhu et al.(2014)策略测试结果表明,在上述四类市场中,所构建的系统交易情况都优于基本的买入 - 持有策略。Deng et al.(2016)利用循环深度神经网络来处理实

时金融信号特征,试图以这一结构为基础构建交易策略去战胜经验丰富的金融资产交易者。他们的模型由具有相关性的深度学习和强化学习两部分构成,其中,前者用于从动态市场条件中自动学习相关特征、进行特征提取,后者对所学习到的相关特征做出反应,进而在一个未知环境中进行交易决策、积累收益。通过复杂的深度神经网络与循环结构相结合,他们的模型取得了较好的效果。Deng et al. (2016) 还进行了一系列测试,证明了他们的模型在股票市场和商品期货市场上都具有良好的稳健性。

还有部分学者是从策略所选择的资产组合出发,利用深度学习方法对资产组合的构成进行优化,从而起到加强交易策略的作用。Sharang and Rao (2015) 利用由栈式受限玻尔兹曼机构成的深度信念网络设计了一个中频交易策略,用来对 5 年期和 10 年期的美国国债期货进行交易。他们将这个交易策略转化成一个分类问题,将两种国债期货的日度价格均线和双周价格均线等作为原始输入数据,采用主成分分析法对价格序列进行处理、降低维度,生成由这个两种国债期货组成的资产组合价格,然后利用深度信念网络提取这个组合技术指标的特征,进而对资产组合价格的周度运动方向进行预测。其中,深度信念网络是由适用于连续性输入的两层受限玻尔兹曼机构成。Sharang and Rao (2015) 将构建的预测模型应用到交易策略中,采用对冲交易方式,买入一种国债期货、卖出另一种国债期货以规避风险,一段时间之后将交易收益与采用随机分类器模型的交易收益进行对比,在排除了交易费用和税收之后,这个构建的资产组合盈利了 10 个单位,大约是 9 万美元。Heaton et al. (2016b) 基于经典的马克维茨理论构建了一个深度资产组合理论。在给定一个资产组合目标函数的情况下,建立深度分层资产组合,然后将组合与目标函数联合起来、挖掘其中的映射关系,并对此进行平衡优化,进而绘制深度资产组合的有效前沿。简单来说,他们的理论基于首先对市场信息进行编码,然后再解码从而形成一个符合目标的资产组合。为了将深度因素进行量化,Heaton et al. (2016b) 将深度因素看成是以训练中市场数据为表征的看涨期权或看跌期权。在构造出这个深度资产组合理论之后,Heaton et al. (2016b) 应用该理论,从生物技术指数 IBB 的成分股中筛选出 10 只最能代表共性的股票和 X 只最不具有共性的股票,重新构建了一个深度资产组合。与 IBB 指数对比分析来看,他们应用深度资产组合理论构建的组合表现优于 IBB 指数表现。

三、深度学习在金融领域中的应用贡献

虽然相关文献数量上并不是十分丰富,但是学者们已经尝试将各种深度学习结构应用在金融领域中。通过对已有文献的梳理,我们认为深度学习在金融领域中的应用具有两方面的贡献:一方面,深度学习方法更适合金融市场数据,其应用不仅改进了金融领域中的预测分析方法,还带来了算法的优化进步,各种适应深层网络、能够解决无效训练的算法推动了传统金融数据分析方法的改进;另一方面,深度学习的发展也推动了金融领域实证研究范式的改变,其在实证领域的成果在一定程度上助推了相关金融经济理论的发展与完善。

(一) 改进金融领域中的预测分析方法

金融领域数据从类型上看,主要是时间序列数据。这样的数据往往具有三个特点:一是通常具有大量的噪音并且具有高维度特性。虽然存在一些降维技术,但是降维可能会引起信息的损失;二是不能确定所获得的信息是否足以预测未来趋势。比如在金融市场中,单独观察一只股票的信息时,仅仅是观察了金融市场这个复杂系统中的一小部分信息,这很可能导致预测所需信息不足;三是时间序列数据对时间变量具有很强的依赖性(Långkvist et al., 2014)。解决这一问题可以有两种方式,一种是尽可能输入时间跨度较长的时间序列数据,另一种是使用具有记忆性的模型。但是第一种方法会造成数据量过大而难以驾驭,第二种方法会存在无法确定依赖时间长度的问题。第三个特点最直接的体现就是数据的时变性。有研究表明与其他非时间序列数据相比,随着时间的延长,大多数时间序列的特征都会由不变转为可变(Långkvist et al., 2014)。对传统时间序列建模进行分析的方法主要包括计量方法和数据挖掘方法,前者主要是指线性回归模型、随机游走模型、GARCH 族模型(张承钊, 2016)以及向量自回归模型(Lütkepohl, 2005),后者主要是指隐马尔可夫模型(Rabiner and Juang, 1986)、人工神经网络(Atsalakis and Valavanis, 2009)、机器学习(Hamid and Iqbal, 2004; Bollen et al., 2011)等方法。但是结合金融领域数据特征来看,传统建模分析方法在应用时往往存在以下问题:一是传统的模型方法并不能很好地、综合性地描述一个金融市场(如股票市场),因为这种市场会受到众多因素的影响,如政治事件、公司政策、经济发展状况、投资者预期、投资者心理以及其他相关市场的运动(Tan et al., 2007; Zhu et al., 2014),传统模型难以处理如此复杂的数据,而忽略外部环境因素又将不利于揭示金融问题背后的经济逻辑(尚玉皇和郑挺国, 2016);二是传统方法主要依赖人工设计特征(Rönnqvist and Sarlin, 2016; Fehrer and Feuerriegel, 2015),这难免掺杂了研究者的主观因素,并且设计的特征过于具有针对性、具有不完整性,要想在原模型的基础上继续改进或发展模型需要耗费大量人工(Rönnqvist and Sarlin, 2016);三是即使是较为有效的人工神经网络方法,也存在备受诟病的过拟合问题(Huang et al., 2004)以及导致模型无效的梯度消失或梯度爆炸等问题,即使所建模型如实的反映了系统情况,训练过程也存在着易在高维度函数空间中陷入局部最小化的问题(Xiong et al., 2015)。同时,传统线性方法还具有过强的“线性”设定,机器学习方法有对噪音信号反应较差的缺点(Sharang and Rao, 2015)。这些都成为将传统方法应用于金融领域的劣势,制约了对金融市场数据的分析与预测。

而深度学习的发展在一定程度上缓解了上述问题。Najafabadi et al. (2015) 研究指出,深度学习实际上是连续不断的层的深度结构,其通过多层级之间传递数据而学习特征,可以说堆叠的非线性传输层是深度学习方法的基本思想。利用深度学习方法构建的模型最终获得的是输入数据的高度非线性函数,在深度结构中数据通过的层数越多,所构建的非线性结构就会越复杂,这种非线性的转化有助于揭开数据变化的原因(Najafabadi et al., 2015)。面对金融市场嘈杂、动态、复杂、高维度的数据结构,深度学习利用逐层特征进行特征学习:通过简单且非线性的模块获取特征,这些模块将一个较低层次的特征转

换成更高层、更微妙、更抽象的特征; 然后通过足够多的这样的转换组合, 学习非常复杂的函数(LeCun et al. 2015)。更高层次的特征扩大了输入的金融原始数据中对预测或者分类具有重要作用的因素的影响, 而弱化了不相关因素的影响; 这些特征不是由人工设计的, 而是使用学习过程从数据中学习得到的, 所以在一定程度上最为客观, 这也是深度学习最为关键的优势(LeCun et al. 2015; Rönnqvist and Sarlin 2016)。与相对浅层的人工神经网络学习结构相比, 深度学习的各种模型结构更善于从数据中提取出全局数据之间的关系, 它可以从无监督数据中自动提取特征, 对各种信息形式都具有较好的适用性; 并且可以利用更复杂、更抽象的数据特征, 从而使预测或者分类工作更加有效(Najafabadi et al. 2015)。同时, 深度学习还具有强烈的数据驱动特点(Rönnqvist and Sarlin 2016), 善于发现高维度数据中的复杂结构(LeCun et al. 2015), 多种神经网络(如前馈神经网络、递归神经网络等) 都适合进行改进, 形成深度神经网络等模型(Prieto et al. 2016)。Heaton et al. (2016a) 对深度学习在金融领域中应用的优势进行了概括, 认为其优势具体可以体现为三点: 一是只要是与预测问题相关的数据都可以被纳入模型之中, 不受数据维度限制; 二是考虑了输入数据间的非线性和复杂交互作用, 相对于传统模型可以提升样本内拟合性; 三是可以有效避免过拟合问题。这一系列的优点使得深度学习成为金融领域数据分析中很有价值的工具。

部分学者从实证的角度论证了深度学习方法在金融领域中的适用性。Takeuchi and Lee(2013) 的研究证明了如果经过合适的预处理, 栈式受限玻尔兹曼机构成的编码器能够从很弱的信号或者充满噪音的时间序列(如金融资产价格) 中提取有效特征。Xiong et al. (2015) 将其构建的长短期记忆模型与线性岭回归、Lasso 以及 GARCH 自回归等模型的结果进行了对比分析, 结果表明他们的长短期记忆模型能够有效地避免过拟合, 其表现比线性岭回归、Lasso 和 GARCH 自回归模型的表现至少要优化了 31%, 并且在模型评价指标(MAPE 和 RMSE) 上也具有良好的性质, 这一结果显示了在强噪音情况下长短期记忆模型对金融时间序列应用的潜力。Shen et al. (2015) 以汇率为研究对象, 采用 RMSE、MAE、MAPE 等六个评价指标从四个方面(预测误差、方向精度、真实汇率序列与预测汇率序列的相关性、预测稳定性) 分析其所构建的深度信念网络模型的效果。通过与前馈神经网络的对比, 他们发现深度信念网络在五个评价指标上表现都要好于前馈神经网络, 特别是在样本外预测中, 新构建的深度信念网络不仅具有更高的准确率, 而且结果的稳定性也更好, 证明了这种深度信念网络对汇率预测的适用性。Sirignano(2016) 将他们构建的空间深度神经网络模型与标准神经网络模型、朴素经验模型以及含有非线性特征的逻辑回归模型的表现进行了对比, 发现其所构建的空间深度神经网络的表现优于标准神经网络, 而这两种神经网络的表现都优于朴素经验模型和带有非线性特征的逻辑回归模型; 进一步的, 与标准神经网络相比, 空间深度神经网络结构具有对 D 维数据建模的优势, 可以对 D 维数据的整个空间进行分布建模, 同时, 这种优势还体现在更好的泛化能力、更低的计算成本以及利用局部空间结构的能力, 并且就训练过程来看, 空间深度神经网络具有更低的样本外误差, 更短的训练时间以及更强的解释能力。Fehrer and Feuerriegel(2015) 从

准确率、精度、召回率以及 F1 - score 等方面比较了递归自动编码器与随机森林模型基于文本对股市进行预测的效果。他们认为随机森林与递归自动编码器的主要区别在于前者需要传统的特征工程辅助来提取特征,而后者则是基于优化算法自动生成特征。从实证结果上来看,递归自动编码器预测准确率为 0.56,相对于基准随机森林模型的预测准确率提高了 5.66%;其 F1 - score 与基准相比,也从 0.52 提升到了 0.56,相对上升 7.69%。Fehrer and Feuerriegel(2015)认为准确率和 F1 - score 的提升,意味着递归自动编码器优于经典机器学习算法,他们认为这一结果能够直接证明在股票价格运动的预测中深度学习的应用是优于传统机器学习的。

(二) 改变金融领域实证研究范式

深度学习在金融领域中的应用给该领域实证研究提供了新的思路,也促使我们对其背后所隐含的实证原理进行深度思考。深度学习推动了金融领域实证研究的两类转变:一是从线性建模到非线性建模思想的转变;二是从关注模型参数显著性到关注模型结构以及模型动态特征的变化。

第一种转变根植于深度学习自身特性。过去金融领域普遍采用的传统计量方法多是线性模型,简单地处理局部线性最优往往就能推出整体最优,这在金融市场实际数据中具有极大的局限性。而传统的人工神经网络等非线性模型对金融市场分析从准确率上看都不够理想,给出的结果并不令人满意(Zhu et al. 2014; Agrawal et al. 2013)。深度学习的各种结构主要基于浅层人工神经网络的层级结构,继承了人工神经网络的动态非线性、全局连接、自组织等特点,并且能够适时修正网络参数(Prieto et al. 2016)。深度学习由于具有结构上的“深度”,所以可以根据数据训练产生更加多样、更加复杂的非线性结构,从而更加注重并且更加适合非线性特征及模型的构建。同时,计算机设备的不断优化以及衍生出的一系列优化算法都为深度学习的应用奠定了基础,使得建立合理的、具有较好预测效果的非线性模型成为可能。深度学习的成功应用促使我们更多地思考现实问题中的非线性关系,有利于我们对现实产生新的、符合实际的理解。已有学者通过实证证明了深度学习适合用来构建非线性模型的观点。Zhu et al. (2014)检验了所构建的自动股票决策支持系统效果,利用这种非线性系统证明了深度信念网络具有优秀的深度学习能力。Dixon et al. (2016)在构建深度神经网络的过程中,考虑了独立变量与非独立变量之间的复杂非线性关系,以此为出发点对模型结构进行调整优化。Sirignano(2016)构建的空间深度神经网络模型对依据限价指令簿进行的价格预测显示出了良好的适用性,他们认为能够取得优良结果的原因在于空间深度神经网络能够较好地拟合限价指令簿中的非线性特征。Heaton et al. (2016b)在构建深度资产组合理论时指出,他们所用的模型可以视为一个简单的自动编码器,其中编码的映射关系可以存在丰富的非线性结构,这样可以在捕捉更低的价格误差的同时仍然提供有效的样本外资产组合。

另一方面,传统金融计量模型关注的重点往往在于参数的显著性,通过对显著性的分析来判别相应因素在模型中是否产生影响,其反映了变量之间的相关性关系,有助于我们识别关键因素。传统方法的这种特性虽然便于我们对模型进行处理,但是简化的模型往

往会造成信息损失,掩盖其中的弱相关关系。对于复杂的金融市场环境而言,失之毫厘,极有可能导致谬以千里的后果。而利用深度神经网络构建的模型则不再将模型参数的显著性作为重点关注对象,而是更加关注模型自身的结构特征以及由于时间推移或者系统运动而导致的动态变化特征。Takeuchi and Lee(2013)认为应用深度学习可以发现股票价格运动的新模式,并通过实证证明了栈式受限玻尔兹曼机这种深度学习结构在金融市场中应用的潜力。Ding et al.(2015)的研究认为深度卷积神经网络是预测基于事件驱动的股票价格运动的有效工具,而这其中的关键就在于他们所使用的深度卷积神经网络能够考虑时间变迁因素,捕捉新闻事件的长期影响。Rönnqvist and Sarlin(2016)的研究表明深度神经网络模型可以根据时间变动进行动态调整,这种动态调整使得深度神经网络能够帮助识别、理解过去发生的和正在进行的事件,从而提升事件预测准确率。Dixon et al.(2016)在利用深度神经网络预测期货价格变动情况的试验中,证明了深度神经网络同时处理多元信息的特点有助于从数据中获取更为丰富的信息集,从而捕捉价格运动中的时变协同运动特征。

同时,这些研究不仅证明了深度学习方法对金融数据良好的适用性,也为我们重新认识金融市场结构、深入理解金融市场行为提供了一定的依据。实证领域中金融市场新的结构与特征的发现不仅能够帮助我们解决特定的预测分析问题,也为我们对金融领域中的各种理论(如投资组合理论、有效市场假说理论)的理解开辟了新的思考空间,从而助推相关理论发展与完善。

四、进一步研究的启示

从已有文献总结来看,如何利用深度学习提升预测能力仍然是值得探讨的问题;关于深度学习在金融领域的应用研究仍处于探索阶段,其相关应用在丰富金融实证研究与金融理论等方面具有重要意义。深度学习是金融领域进行数据分析与预测的前沿方法,也是更加贴近实际金融市场数据特征的方法,是推动智能金融的关键技术。想要真正实现智能金融中所包含的智能投顾、智能投研与智能投资,可能需要复杂的深度学习模型结构。深度学习模型的构建与优化在技术上具有相当程度的复杂性,研究者对金融市场的正确理解甚为重要,只是单纯为了应用新方法而应用深度学习难以发挥其实际的作用,研究者对模型的深刻理解与经济金融领域的专业知识相结合才有可能实现深度学习的正确应用。具体来看,现有研究的难点往往集中在以下三个方面。

第一个难点在于深度学习模型结构的构建,主要体现为隐藏节点数量和隐藏层数量的确定。节点数和隐藏层数量是深度学习各种模型结构捕捉数据非线性特征的关键,然而现在并没有形成权威的、成熟的理论来提供指导,目前主要还是依靠研究者不断试验进行摸索。第二个难点在于模型稳健性的验证。由于深度学习模型结构可以根据不同的数据实现自身结构重构,所以对于某一特定模型是否一定适合某一特定领域的分析或预测还需要大量的试验进行验证;数据的频率也会对这种模型的稳健适用产生一定的影响。

有学者对不同的数据频率进行过验证,但尚未形成一致的结论(Shen et al. 2015; Xiong et al. 2015)。第三个难点在于对应用深度学习所得出的结论进行经济学意义上的解释。利用深度学习方法所构建的模型带来了金融领域实证研究范式的改变,弱化了参数显著性的影响,这在一定程度上也削弱了利用经济学解释模型结果的能力。传统金融计量方法利用相关性分析提供了经济因果分析的依据,而对于应用深度学习所得到的结果(及其应用过程)似乎难以找到一个合适的切入点,对经济学中的因果关系进行分析。如何利用深度学习所呈现的模型结构及运动特征来进行经济含义分析,如何在享受深度学习所带来的预测高准确率的同时深入地进行经济学原理阐释,这两个问题不仅仅是研究的难点,也是未来深度学习方法在金融领域应用,乃至在更广泛的经济领域中应用所亟需解决的重要问题。

关于未来深度学习在金融领域中的应用,以下方向值得深入研究:一是进一步寻找深度学习方法与金融经济理论的契合点,挖掘深度学习通过应用所体现出的经济学思想,从而丰富金融经济理论;二是在各种深度学习模型结构中,提炼能够适合金融市场分析与预测的一般性框架或规律性方法,以实现该方法的普遍性应用,整体上推动金融智能化发展;三是进一步根据具体金融市场特征,为利用深度学习进行数据分析与预测提供实证证据,如检验某一深度学习结构对不同频率数据的适用性与预测效果差异,探索深度学习利用非同质信息资源的能力;四是探讨深度学习在金融领域的新型应用,借鉴其在图片、自然语言处理等方面的技术来优化对金融市场各种类型数据进行分析与预测的过程,不局限于时间序列数据,如尝试根据深度学习提取深层文本信息的能力探讨金融研报的自动生成,也可以尝试利用深度学习捕捉信息(特别是尾部信息)的能力在风险管理方面进行应用探索;五是对于我国而言,相关研究较为缺乏,即使有国内学者将深度学习应用在金融市场中,其选择研究的对象往往也是国外金融市场,在深度学习应用的理论与实证方面较少与我国金融市场特征相结合,没有针对我国金融市场进行较为深入的剖析,有必要在此方面加强研究。

参考文献

- [1] 尚玉皇和郑挺国 2016,《短期利率波动测度与预测:基于混频宏观-短期利率模型》,《金融研究》第 11 期,第 47 ~ 62 页。
- [2] 曾志平、萧海东和张新鹏 2017,《基于 DBN 的金融时序数据建模与决策》,《计算机技术与发展》第 4 期,第 1 ~ 8 页。
- [3] 张承钊 2016,《一种金融市场预测的深度学习模型:FEPA 模型》,电子科技大学博士论文。
- [4] Agrawal J. G., Chourasia V. S., and Mittra A. K.. 2013. "State - of - the - art in Stock Prediction Techniques" *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, 2(4): 1360 ~ 1366.
- [5] Arel I., Rose D. C., and Karnowski T. P.. 2010. "Deep Machine Learning - A New Frontier in Artificial Intelligence Research [research frontier]" *IEEE Computational Intelligence Magazine*, 5(4): 13 ~ 18.
- [6] Asness C. S., Moskowitz T. J., and Pedersen L. H.. 2013. "Value and Momentum Everywhere" *The Journal of Finance*, 68(3): 929 ~ 985.

- [7] Atsalakis G. S. , Valavanis K. P. . 2009. "Surveying Stock Market Forecasting Techniques – Part II: Soft Computing Methods" *Expert Systems with Applications* , 36(3) : 5932 ~ 5941.
- [8] Ba J. , Frey B. . 2013. "Adaptive Dropout for Training Deep Neural Networks" *Advances in Neural Information Processing Systems* , Working Paper , pp. 3084 ~ 3092.
- [9] Bengio Y. , Courville A. , and Vincent P. . 2013. "Representation Learning: A Review and New Perspectives" *IEEE transactions on pattern analysis and machine intelligence* , 35(8) : 1798 ~ 1828.
- [10] Bengio Y. . 2009. "Learning Deep Architectures for AI" *Foundations and trends® in Machine Learning* , 2(1) : 1 ~ 127.
- [11] Bollen J. , Mao H. , and Zeng X. . 2011. "Twitter Mood Predicts the Stock Market" *Journal of Computational Science* , 2(1) : 1 ~ 8.
- [12] Bovier A. , Černý J. , and Hryniv O. . 2006. "The opinion game: Stock price evolution from microscopic market modeling" *International Journal of Theoretical and Applied Finance* , 9(01) : 91 ~ 111.
- [13] Cartea Á. , Jaimungal S. , and Ricci J. . 2014. "Buy Low , Sell High: A High Frequency Trading Perspective" *SIAM Journal on Financial Mathematics* , 5(1) : 415 ~ 444.
- [14] Cavalcante R. C. , Brasileiro R. C. , and Souza V. L. F. , et al. . 2016. "Computational Intelligence and Financial Markets: A Survey and Future Directions" *Expert Systems with Applications* , (55) : 194 ~ 211.
- [15] Cont R. , Stoikov S. , and Talreja R. . 2010. "A Stochastic Model for Order Book Dynamics" *Operations research* , 58(3) : 549 ~ 563.
- [16] Cont R. . 2011. "Statistical Modeling of High – frequency Financial Data" *IEEE Signal Processing Magazine* , 28(5) : 16 ~ 25.
- [17] Deng Y. , Bao F. , and Kong Y. , et al. . 2016. "Deep Direct Reinforcement Learning for Financial Signal Representation and Trading" DOI: 10.1109/TNNLS.2016.2522401.
- [18] Ding X. , Zhang Y. , and Liu T. , et al. . 2014. "Using Structured Events to Predict Stock Price Movement: An Empirical Investigation" *EMNLP* , pp. 1415 ~ 1425.
- [19] Ding X. , Zhang Y. , and Liu T. , et al. . 2015. "Deep learning for event – driven stock prediction" *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI' 15)* , pp. 2327 ~ 2333.
- [20] Dixon M. , Klabjan D. , and Bang J. H. . 2015. "Implementing Deep Neural Networks for Financial Market Prediction on the Intel Xeon Phi" *Proceedings of the 8th Workshop on High Performance Computational Finance* , ACM , No. 6.
- [21] Dixon M. F. , Klabjan D. , and Bang J. H. . 2016. "Classification – based Financial Markets Prediction Using Deep Neural Networks" Available at SSRN 2756331.
- [22] Fehrer R. , Feuerriegel S. . 2015. "Improving Decision Analytics with Deep Learning: The Case of Financial Disclosures" *arXiv preprint arXiv: 1508.01993*.
- [23] Fletcher T. , Shawe – Taylor J. . 2013. "Multiple Kernel Learning with Fisher Kernels for High Frequency Currency Prediction" *Computational Economics* , 42(2) : 217 ~ 240.
- [24] Hagenau M. , Liebmann M. , and Neumann D. . 2013. "Automated News Reading: Stock Price Prediction based on Financial News Using Context – capturing Features" *Decision Support Systems* , 55(3) : 685 ~ 697.
- [25] Hamid S. A. , Iqbal Z. . 2004. "Using Neural Networks for Forecasting Volatility of S&P 500 Index Futures Prices" *Journal of Business Research* , 57(10) : 1116 ~ 1125.
- [26] Heaton J. B. , Polson N. G. , and Witte J. H. . 2016a. "Deep Learning in Finance" *arXiv preprint arXiv: 1602.06561*.
- [27] Heaton J. B. , Polson N. G. , and Witte J. H. . 2016b. "Deep Portfolio Theory" *arXiv preprint arXiv: 1605.07230*.
- [28] Hinton G. E. , Osindero S. , and Teh Y. W. . 2006. "A Fast Learning Algorithm for Deep Belief Nets" *Neural computation* , 18(7) : 1527 ~ 1554.

- [29] Hinton G. E. , Salakhutdinov R. R. . 2006. "Reducing the Dimensionality of Data with Neural Networks" *Science* , 313 (5786) : 504 ~ 507.
- [30] Hinton G. E. , Srivastava N. , and Krizhevsky A. , et al. . 2012. "Improving Neural Networks by Preventing Co - adaptation of Feature Detectors ," arXiv preprint arXiv: 1207. 0580.
- [31] Huang C. L. , Tsai C. Y. . 2009. "A Hybrid SOFM - SVR with A Filter - based Feature Selection for Stock Market Forecasting" *Expert Systems with Applications* , 36(2) : 1529 ~ 1539.
- [32] Huang H. , Kercheval A. N. . 2012. "A Generalized Birth - death Stochastic Model for High - frequency Order Book Dynamics" *Quantitative Finance* , 12(4) : 547 ~ 557.
- [33] Huang Z. , Chen H. , and Hsu C. J. , et al. . 2004. "Credit Rating Analysis with Support Vector Machines and Neural Networks: A Market Comparative Study" *Decision support systems* , 37(4) : 543 ~ 558.
- [34] Kercheval A. N. , Zhang Y. . 2015. "Modelling High - frequency Limit Order Book Dynamics with Support Vector Machines" *Quantitative Finance* , 15(8) : 1315 ~ 1329.
- [35] Kijima H. , Takada H. , and Tomiya T. . 2016. "SVM - Enhanced Filtering Model for Limit Order Book Dynamics" *Proceedings of the ISCIE International Symposium on Stochastic Systems Theory and its Applications , the ISCIE Symposium on Stochastic Systems Theory and Its Applications* , pp. 181 ~ 188.
- [36] Kogan S. , Levin D. , and Routledge B. R. , et al. . 2009. "Predicting Risk from Financial Reports with Regression" *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics , Association for Computational Linguistics* , pp. 272 ~ 280.
- [37] Kuremoto T. , Kimura S. , and Kobayashi K. , et al. . 2014. "Time Series Forecasting Using A Deep Belief Network with Restricted Boltzmann Machines ," *Neurocomputing* , (137) : 47 ~ 56.
- [38] Långkvist M. , Karlsson L. , and Loutfi A. . 2014. "A Review of Unsupervised Feature Learning and Deep Learning for Time - series Modeling" *Pattern Recognition Letters* , (42) : 11 ~ 24.
- [39] LeCun Y. , Bengio Y. , and Hinton G. . 2015. "Deep Learning" *Nature* , 521(7553) : 436 ~ 444.
- [40] Lischinsky A. . 2011. "In Times of Crisis: A Corpus Approach to the Construction of the Global Financial Crisis in Annual Reports" *Critical Discourse Studies* , 8(3) : 153 ~ 168.
- [41] Luss R. , d' Aspremont A. . 2015. "Predicting Abnormal Returns from News Using Text Classification" *Quantitative Finance* , 15(6) : 999 ~ 1012.
- [42] Lütkepohl H. . 2005. "New Introduction to Multiple Time Series Analysis" Published by Springer Science & Business Media.
- [43] Mineev M. , Schommer C. , and Grammatikos T. . 2012. "News and Stock Markets: A Survey on Abnormal Returns and Prediction Models" Technical Report , UL.
- [44] Mordvintsev A. , Olah C. , and Tyka M. . 2015. "Inceptionism: Going Deeper into Neural Networks" Google Research Blog. Retrieved June , 2015 , 20.
- [45] Muntermann J. , Guettler A. . 2007. "Intraday Stock Price Effects of Ad Hoc Disclosures: The German Case" *Journal of International Financial Markets , Institutions and Money* , 17(1) : 1 ~ 24.
- [46] Nair V. , Hinton G. E. . 2010. "Rectified Linear Units Improve Restricted Boltzmann Machines" *Proceedings of the 27th International Conference on Machine Learning (ICML - 10)* , pp. 807 ~ 814.
- [47] Najafabadi M. M. , Villanustre F. , and Khoshgoftaar T. M. , et al. . 2015. "Deep Learning Applications and Challenges in Big Data Analytics" *Journal of Big Data* , 2(1) : 1 ~ 21.
- [48] Nassiroussi A. K. , Aghabozorgi S. , and Wah T. Y. , et al. 2014. "Text Mining for Market Prediction: A Systematic Review" *Expert Systems with Applications* , 41(16) : 7653 ~ 7670.
- [49] Pang B. , Lee L. . 2008. "Opinion Mining and Sentiment Analysis" *Foundations and trends in information retrieval* , 2

- (1-2): 1 ~ 135.
- [50] Prieto A. , Prieto B. , and Ortigosa E. M. , et al. . 2016. "Neural Networks: An Overview of Early Research , Current Frameworks and New Challenges" *Neurocomputing* , (214) : 242 ~ 268.
- [51] Pröllochs N. , Feuerriegel S. , and Neumann D. . 2014. "Generating Domain – Specific Dictionaries Using Bayesian Learning" Available at SSRN 2522884.
- [52] Qiu X. , Zhang L. , and Ren Y. , et al. . 2014. "Ensemble Deep Learning for Regression and Time Series Forecasting" *CIEL* , pp. 21 ~ 26.
- [53] Rabiner L. , Juang B. . 1986. "An Introduction to Hidden Markov Models" *IEEE ASSP magazine* , 3(1) : 4 ~ 16.
- [54] Rönnqvist S. , Sarlin P. . 2016. "Bank Distress in the News: Describing Events through Deep learning" arXiv preprint arXiv: 1603.05670.
- [55] Roşu I. . 2009. "A Dynamic Model of the Limit Order Book" *Review of Financial Studies* , 22(11) : 4601 ~ 4641.
- [56] Schmidhuber J. . 2015. "Deep Learning in Neural Networks: An Overview" *Neural Networks* , (61) : 85 ~ 117.
- [57] Schumaker R. P. , Chen H. . 2009. "Textual Analysis of Stock Market Prediction Using Breaking Financial News: The AZFin Text System" *ACM Transactions on Information Systems (TOIS)* , 27(2) : 12 ~ 40.
- [58] Sharang A. , Rao C. . 2015. "Using Machine Learning for Medium Frequency Derivative Portfolio Trading" arXiv preprint arXiv: 1512.06228.
- [59] Shek H. H. S. . 2011. "Modeling High Frequency Market Order Dynamics Using Self – excited Point Process" Available at SSRN 1668160.
- [60] Shen F. , Chao J. , and Zhao J. . 2015. "Forecasting Exchange Rate Using Deep Belief Networks and Conjugate Gradient Method" *Neurocomputing* , (167) : 243 ~ 253.
- [61] Sirignano J. A. . 2016. "Deep Learning for Limit Order Books" arXiv preprint arXiv: 1601.01987.
- [62] Srivastava N. , Hinton G. E. , and Krizhevsky A. , et al. . 2014. "Dropout: A Simple Way to Prevent Neural Networks from Overfitting" *Journal of Machine Learning Research* , 15(1) : 1929 ~ 1958.
- [63] Takeuchi L. , Lee Y. Y. A. . 2013. "Applying Deep Learning to Enhance Momentum Trading Strategies in Stocks" CS229 , Stanford.
- [64] Tan T. Z. , Quek C. , and Ng G. S. . 2007. "Biological Brain – inspired Genetic Complementary Learning for Stock Market and Bank Failure Prediction" *Computational intelligence* , 23(2) : 236 ~ 261.
- [65] Wang W. Y. , Hua Z. . 2014. "A Semiparametric Gaussian Copula Regression Model for Predicting Financial Risks from Earnings Calls" *ACL* , (1) : 1155 ~ 1165.
- [66] Xiong R. , Nicholas E. P. , and Shen Y. . 2015. "Deep Learning Stock Volatilities with Google Domestic Trends" arXiv preprint arXiv: 1512.04916.
- [67] Yoshihara A. , Fujikawa K. , and Seki K. , et al. . 2014. "Predicting Stock Market Trends by Recurrent Deep Neural Networks" *Pacific Rim International Conference on Artificial Intelligence* , Springer International Publishing , pp. 759 ~ 769.
- [68] Zhu C. , Yin J. , and Li Q. . 2014. "A Stock Decision Support System based on DBNs" *Journal of Computational Information Systems* , 10(2) : 883 ~ 893.

Deep Learning in Financial Empirical Applications: Dynamics , Contributions and Prospects

SU Zhi LU Man LI Dexuan

(Central University of Finance and Economics;
International Monetary Institute , Renmin University of China;
Communication University of China)

Abstract: With the coming intelligent era and increasing demand for financial data analysis , deep learning has become the foreland in financial filed , especially in the prediction of financial market movements , text information processing and trading strategies improvement. Deep learning contains deep neural networks , deep belief networks and other structures. They extract deep features through the layered structure , strengthen important factors and filter the noise , which is of great significance to improve forecast accuracy. The applications and the variable optimization technologies of deep learning improve the forecast and analysis method in financial filed , make empirical research paradigm shift from linear to nonlinear and the concern from significance of parameters to the structure and dynamic characteristics of models , and also contribute to enriching financial and economic theories. It is important and difficult to build models with suitable structure and steady effect to capture the effective characteristics of financial data and interpret economic meaning with deep learning. The future research about deep learning in financial filed can be from these aspects , such as digging deep economic significance , refining the general analysis framework and exploring the applicability for heterogeneous information.

Key words: Deep Learning , Financial Market Predication , Text Mining , Deep Neural Network , Deep Belief Network

(责任编辑: 李景农) (校对: LN)