人工智能对经济金融等社会科学现象预测的拓展

向 巨

摘 要:本文从经济金融对象的特点,说明社会科学领域预测的复杂性。结合人工智能在该领域预测的尝试及已有成果,进一步说明人工智能相比传统分析和研究方法的优势,阐述了人工智能可以成为金融经济等社会科学预测的突破性工具。分析了从数据建立人工智能模型得到结果,进而检验和提升模型的预测效果,以及人工智能提炼因子从而帮助形成理论的全过程。并且对深度网络、集成学习、迁移学习、自动学习等主要发展方向进行了初步探讨。

关键词: 人工智能 社会科学 预测 深度网络 集成学习

《科学》(Science)杂志在其2017年2月的"预测及其限制"(Prediction and its limits)特刊中探讨了人工智能对科学可预测性的促进。^①该特刊中大多数内容谈及的是运用机器学习方法对战争冲突等政治、社会事件,乃至人类行为和文明发展等社会科学课题的预测及取得的显著成果。经济金融本身作为重要的社会现象,以及作为其他社会现象的重要驱动因素,几乎无时无刻影响着人类社会的各个方面,对它们的预测具有特殊意义。

然而,到目前为止学界对于经济金融现象的预测,总体而言还难 以令人满意。2008年金融危机后,英国女王伊丽莎白二世质问经济 学家为什么没预测到此次危机。总体而言,经济金融系统很难通过控

① 包括科学各领域未来发展趋势的可预测性。

制住一些变量,来研究另外一些变量间的关系,或者通过简单拆分影响因子的传统分析方法来进行。因为这些变量之间的关系是错综复杂的,并且很难简单的归结为若干独立的因子。英国国家学术院士 Tim Besley 和 Peter Hennessy, 在回复女王疑问的信中也表达了类似观点。^①

本文从经济金融预测对象的特点,该领域已有的人工智能尝试,以 及与传统预测方法对比的优势等方面阐述人工智能是金融经济预测的 突破性工具。同时讨论人工智能提升社会科学领域预测效果的未来发 展方向。

一、经济金融等社会复杂系统的特点

(一)社会复杂系统的特点

在经济金融等社会复杂系统中,有两个关键要素:博弈和反馈。 前者是个体和个体之间的相互作用(往往目标不完全一致),后者是 系统与外部环境和资源之间的相互作用(正或负反馈)。这些复杂模 式通常很难预测,原因可以归结为几方面:

- 1. 高维性:构成现实生活的系统往往被大量可以独立变化的变量决定。比如经济金融的核心载体——市场由无数的交易者组成,不同交易者可能有不同的交易目的和习惯。如果套用物理的描述方法来预测,就是极高维度空间的运动问题。
- 2. 非线性:高维度系统的维度之间具有复杂的相互作用。集体行为的非线性(总体不等于个体之和),导致我们不能把系统分解为单一维度然后做加法的分析方法(Analysis)进行研究。非线性的相互作用,进而形成具备无数层级的复杂组织。笔者认为,更好的研究方法应该是以组织为对象的综合分析(Synthesis)。
 - 3. 反馈性: 最典型的反馈是记忆效应, 使得系统产生路径依赖,

① 他们在信中写道"很多原因导致了没有人预测出危机的发生时间、程度及严重性,没能避免危机的发生,但最主要的原因是未能从国内外诸多智者身上集思广益,从整体上来理解系统的风险"。

此刻现实与历史深刻关联。反身性是一种由预测产生的特殊反馈:预测股市的价格时,会引起交易策略的变化从而影响预测。社会集体的同向预测更是产生泡沫和狂热的根源。

4. 随机性:复杂系统往往包含有规律以外的随机噪声,导致很难区分发现的模式是噪声导致还是由于系统成员之间的相互作用。

(二)经济金融理论的发展

随着认识的深入,经济金融学术界也逐渐从传统的古典均衡理论 向博弈论,以及非传统非理性假设的行为经济、金融学过渡。举例来说,在资产定价领域,最初很多学者提出了各种基于"理性预期"的理论 来解释金融中风险资产的"异常"回报现象,但是这些理论依然受到了实证研究的挑战。Daniel 和 Titman (2012)认为,这些基于"理性预期"的风险定价模型之所以能够与实证数据相吻合,主要是由于其极低的统计功效(low statistical power),当控制了市值和估值变量以后,传统的风险定价模型无法通过统计检验。

另一些学者则提出了与传统的"理性预期"截然不同的理论模型。他们从行为金融学的角度出发,认为市场的参与者本身是非理性的。2017年诺贝尔经济学奖得主 Thaler(1985)提出过一个经典的心理账户模型。当投资者买人一只股票时,会在心中为此开设一个单独的账户,对股票的收益进行单独核算。这样投资者往往会偏好于卖出上涨的股票,保留亏损的股票,而不是基于理性做出决策。另外,套利的有限性使得风险资产的价格存在着被低估或者高估的可能性(mispricing)。Barberis 和 Thaler(2003)指出,风险资产的"异常"回报现象源于参与者的非理性行为。一方面,由于认知的有限性,非理性的投资者广泛存在于资本市场当中;另一方面,即便市场上存在理性的投资者(通常为专业的投资机构),他们也无法通过无风险套利促使资产被正确的定价。

传统的有效市场理论认为,当市场上的金融资产被低估(高估)时, 理性的投资人将会发现这一机会,立即进行无风险套利,进行做多(空), 从而使得股价迅速回归其内在价值。然而,在现实世界中,"无风险" 套利是存在高昂的成本的。可想而知,基于此类理论衍生的各种传统 多因子量化模型在预测能力方面也缺乏可靠性。而先抛开繁杂理论假 设的约束,以预测效果为导向,以数据模型结果为依据的人工智能模 型则为经济金融的预测开辟了一条崭新的道路。

二、人工智能在经济金融领域的运用

如上文所述,传统的理论模型和计量经济学等实证方法对于推进对经济金融现象的理解起了巨大作用,但是它们有自身缺陷,并且预测效果较弱。比如在经济决策方面,我们有效用 - 偏好理论、激励 - 强化理论、博弈论、行为经济学等诸多分散的理论,各自从不同角度进行分析。各个理论都有一定的解释作用,但是预测的效果都不太好(Fletcher, 2011a)。这种各自为战,"公说公有理、婆说婆有理"的现象反映了经济金融关系的复杂性,也使得个别经济学家甚至消极地认为"经济学不是拿来作预测的"。本节从分析解释和预测两个方面介绍人工智能在经济金融领域的卓有成效的运用。

(一)分析和解释

一些经济金融学家也尝试运用以机器学习为代表的人工智能方法,来分析和理解相关现象。Campbell, Lo 和 Mackinlay 在他们的经典教科书《金融市场计量经济学》(The Econometrics of Financial Markets, 1996)中,除了总结已有的各种计量方法及实证结果外,还介绍了非参数估计包括人工神经网络(Artificial Neural Network,即 ANN)。比如Hutchinson,Lo 和 Poggio(1994)用 ANN 来评估 Black-Scholes 理论期权定价模型的研究,并且表明这类带参数的学习网络在其他参数模型失效时,可以起到很好的作用。

当时由于条件所限,ANN 这些方法在他们的书中只是作为对传统金融计量方法的一个补充,仅仅在全书的最后一章加以介绍。过去

二十多年,随着梯度下降学习(Gradient Descent. Qian, 1999)和反向传播(Backward Propagation. Rumelhart, Hinton 和 Williams, 1986)算法在机器学习理论和实际中的发展,以及以摩尔法则为基础的单机计算能力成千上万倍的增加和 GPU、并行计算等工具的广泛应用,人工智能取得了飞速发展。

2017 年初,哈佛大学教授 Sendhil Mullainathan 在全世界最顶级的金融学术会议美国金融学年会(American Finance Association,即AFA)上专门作了机器学习在经济金融预测上运用的专题报告,展现了机器学习相对于传统的经济计量方法的优势。Chalfin(2016)等改变了传统经济模型生产函数中的劳动者具有同质生产效率的假设,通过机器学习来预测劳动者的不同生产效率。他的方法运用到警察和教师的雇用上,极大地提高了劳动政策的效率和社会福利。

Kleinberg (2015)和 Glaeser (2016)等学者在经济研究中的重要领域资源分配上,运用监督学习取得了明显的成果。他们优化了健康政策领域的资源分配问题,比如健康监测点的位置选择和监测能力分配。Naik (2016)等人运用街景图结合计算机视觉算法来度量十九个美国城市中人们的安全感受。他们发现平均的安全感受度与人口密度和家庭收入有强烈的正相关关系,并且该感受度的变化与收入的不平等性也强烈正相关。

(二)预测

有效预测比仅仅分析解释难度更大,在经济金融领域尤其如此。 深度学习、卷积神经网络、循环神经网络等人工智能模型在经济金融 预测上也被有效运用。

与传统的算法不同,深度神经网络预先并不设定任何特定的模型和规则,而是通过训练数据提供足够多的案例来"学习"完成相应的任务。比如,传统的多因子选股模型通常是依据简单的线性回归挑选出对超额收益有显著影响的因子,进而根据因子排序筛选,构建多空组合。深度学习模型则与之不同,事先并不主观的选定特定的因子进

行排序组合。深度学习模型先根据历史收益率数据,标记出一定时间 周期内相对强势(弱势)的股票。接下来,选择一系列可能对股票收 益造成影响的因子(几十到上百个不等),作为输入数据构成输入层 的若干个节点。输出层则为股票的标记信息,可采用虚拟变量表示。 若股票为强势股,则可将其标记为1,反之,则标记为0。构建好模型 的输入层和输出层以后,下一步是设计隐含层的数量、隐层的节点个数、 模型采用的激活函数、惩罚系数等一系列参数,之后根据模型训练的 结果对样本外数据进行预测打分,选取得分最高的若干股票构建组合, 进行策略回测。

在卷积神经网络模型中,每个卷积神经元仅处理其对应的局部接受域,提取该局部的特征。但是整个网络在对输入数据进行卷积处理后,会继续通过池化层对数据进行进一步的降维处理。具体地说,池化层将卷积层的多维输入数据进一步压缩为一个值进行输出。Ding(2015)等人运用卷积神经网络开发了一个基于新闻事件驱动的股票预测模型。他们首先从新闻文本出提取机构化的事件向量,再通过卷积神经网络网络对股票的短期和长期走势进行预测。研究结果表面,与传统的深度神经网络(DNN)相比,采用卷积神经网络模型的预测准确率提升了6个百分点。

与传统的时间序列方法类似,Hochreite 和 Schmidhuber(1997)提出了长短期记忆模型(long short-term memory,即 LSTM),通过模型的搭建,数据中的"短期记忆"(short-term memory)也能够在模型中保持较长的一段时间。由于以 LSTM 为代表的循环神经网络(Recurrent Neural Network,即 RNN)在处理序列数据上的优异表现,RNN 被广泛应用于自然语言处理领域。相似地,金融市场的量价信息本身就属于时间序列数据,因此,非常适合运用循环神经网络模型。

Nelson(2017)等人采用历史 K 线图数据(开盘价、收盘价、最高价、最低价)和相应的技术指标(均线、MACD、KRJ等),实证对比了传统机器学习模型和长短期记忆模型对股票收益率预测的能力。他们

发现,LSTM 模型的预测准确率为 55.9%,显著高于多层感知器和随机森林等机器学习模型。此外,基于循环神经网络在文本处理领域的优势,传统的基于文本挖掘和情感分析的股票价格预测系统未来也将更多地运用深度学习的算法不断改进。

这些使用结果表明,人工智能能够帮助经济金融学术界摆脱"经济学不是拿来作预测的"的认识误区,以及经济学家在社会大众心目中的"事后诸葛亮"尴尬形象。

三、人工智能与传统预测方法的比较

在上面机器学习等人工智能研究现状的描述中,同时提到了一些 二者在经济金融建模上的区别。它们之间的更多区别在于以下各方面。

(一)模型假设和检验

人工智能更多运用训练测试验证(train-test-validate)来避免过度 拟合,采用交叉验证(cross validation)、自举(bootstrapping)等数 据方法和非线性模型。因此对于大数据(Big data)包括文本分析等也 很有效。传统计量方法试图找出数据间的因果关系,着眼于解释变量 的参数和统计推断(Statistical Inference),运用的技巧包括自然实验 (natural experiments)、工具变量(instrumental variables)、双重差法 (difference-in-difference)、反事实设定(counterfactual thinking)、 实验经济学等。而人工智能不作因果假设,只作相关性假设,并且更 加强调模型的整体预测性。二者共用了简单回归、逻辑回归、主成分 分析等方法。

经济金融系统很难通过控制一些变量来研究另外一些变量间的特定关系。也很难通过通常的所谓拆开分析(analysis)来进行研究,因为各个变量之间的关系很难被真正完全分开。虽然运用实验经济学,可以再造实验和反复验证,用实验数据代替历史数据用来克服以往经验检验的不可重复性,但是它的主要运用局限于微观经济方面。对于

宏观经济现象是没有条件控制一些主要宏观变量不变,同时让另外一些宏观变量变化来构造的可重复实验的。

(二)数据使用

传统统计和计量方法需要更多的对数据的基本假设,比如一阶二阶的平稳性。因此传统的金融研究的对象,主要是一阶差分后的收益率和基于二阶差分构造的波动率,但是在取得收益率(尤其是对数收益率)和隐含波动率的过程中,往往损失了很多有用的信息。由此建立的模型很难回到需要预测的初始变量比如价格上,这是导致预测效果不好的一个重要原因。

深度神经网络在学习过程中可以自动提取特征(类似于传统分析方法中的因子)。比如在人脸识别过程中,深度卷积网络逐步从浅层的原始像素,组合为简单的几何图形,再到深层中形成眼耳口鼻等器官,最后确定恰当位置形成人脸。不需要像传统人脸识别方法中人为定义形状和器官,以及相互位置等判别因素。

人工智能方法能够对数据进行更全面的使用,而不是局限于既有的观念和模型确定的因素。我们可以直接对观察到的原始数据,比如说价格进行建模和预测。这是传统金融研究难以涉及的方面。很多金融投资公司也逐渐把人工智能方法作为量化投资的一个重要扩展应用到分析框架中,并取得了不错的预测和投资结果。

(三)扩展潜力

人工智能最核心的是学习功能。这更为其不断进化和海纳百川提供了坚实的模型和框架。另外,人脑和传统方法对于多维数据的处理,尤其是如何得出综合性的结论有很大的局限性。人工智能中的集成学习更是可以整合人工智能和传统方法的不同模型,综合得出更好结论的有力方法(Kuncheva 和 Whitaker, 2003; Brown, Wyatt, Harris 和 Yao, 2005; Polikar, 2006; Rokach, 2010; Hamed 和 Fazli, 2016)。集成学习中的各种具体方法可以弥补各种传统理论和研究方法盲人摸象、"公说公有理、婆说婆有理"的局限性,把各种分散的优势综合概括起来,

形成统一的有效结论。

经济金融等社会现象中的信息不对称(Information Asymmetry)以及非完美信息(Imperfect Information),用其他方法是很难作出好的模型进行研究和预测,尤其在多方参与者的多期决策过程中。《科学》(Science)杂志 2017 年 2 月的"预测及其限制"特刊中专门指出预测人类行为成为下一个前沿的研究课题(Science,2017)。与此相关,关于人类行为的博弈论(Game Theory),在经济金融研究中取得了非常显著的理论性成果(Shapiro,1989;Myerson,1991)。

然而,这方面的理论结果和实际情况还是有差异。即使在最简单的试验中,人们也只是在百分之六七十的情况下,按照博弈论的理论预测进行操作。一个很重要的原因就是,人们在博弈过程中,对其他参与方的最终目的不是很明确或者怀疑其他参与方是否还有其他目的。而结合博弈论的强化学习方法在,德州扑克游戏中已经能够完全战胜人类顶尖的职业牌手(Brown 和 Sandholm, 2017; Moravčík等,2017)。其中运用了基于人类常识(Heuristics)的启发式的神经网络学习,并取得了很好的简化运算和综合预测的效果。

四、未来方向与思考

在相关学科的预测方面,未来主要可能的发展方向包括双向迁移 学习(不只是人工智能方法和模型间的迁移,更包括将人类已有的分析结果比如说分析师的研报与人工智能之间的双向迁移和互动),深 度集成学习等方法进一步用以解决多维及综合分析(Synthesis)等,以 及自动机器学习(Auto Machine Learning,即 AutoML)用以促进非人工 智能专业领域的人才来更好结合人工智能。2017 年,Google 提出的自 动学习框架训练出了比专业数据分析师得到的更好的图像识别模型。

许多学者和业界人士认为非监督学习是未来人工智能发展的方向。 在涉及社会科学的应用中,对监督学习和非监督学习二者之间的比较 会更加复杂,因为人工智能的结果和应用最终往往受到别人和社会的评判。2017年"一石激起千层浪"的同性恋识别算法和结果即为典型例证。那么在这种情况下,即使由非监督学习算法产生的结果,还能被简单地认为是传统意义上的非监督吗?那么另一方面,我们又是否需要把人和社会的评判,提前加入到学习的过程中作为预期呢?人工智能对博弈的处理已经加入一些类似预期。是否有更好的融入方法呢?这些方面也是重要的研究方向。

带给人类的改变是最需要探讨的。比如现在很多人工智能学家往往把人工智能和人类来进行简单的比较来得出一些结论。而笔者认为,在未来人工智能会造成人的分化,导致人类这个概念没有以前那么强烈。或者说很难再把人类作为一个整体和人工智能来比较。人工智能在专业技能上(比如现在的外语翻译、图像识别、自动驾驶等方面)极可能超过大多数人。由此带来的经济和社会结构的变化,值得更多探讨。

参考文献:

- 1. Barberis N., Thaler R., "A survey of behavioral finance", *Handbook of the Economics of Finance* 1 (2003), pp. 1053-1128.
- Brown N., Sandholm T., "Safe and Nested Endgame Solving for Imperfect-Information Games", Proceedings of the AAAI workshop on Computer Poker and Imperfect Information Games, 2017.
- 3. Brown G., Wyatt J., Harris R., Yao X., "Diversity creation methods: a survey and categorisation", *Information Fusion*, 2005, 6 (1), pp.5–20.
- Campbell, J. Y., Lo, A. W., Mackinlay A. C., "The Econometrics of Financial Markets", *Princeton University Press*, 1996.
- Chalfin A., Danieli O., Hillis A., Jelveh Z., Luca M., Ludwig J., Mullainathan S.,
 "Productivity and Selection of Human Capital with Machine Learning", *American*

- Economic Review, 2016, Papers and Proceedings 106, no. 5, pp. 124-127.
- 6. Daniel K., Titman S., "Testing factor-model explanations of market anomalies", *Critical Finance Review*, 2012, 1 (1), pp. 103–139.
- Ding X., Zhang Y., Liu T., et al., "Deep learning for event-driven stock prediction", 2015 International Joint Conferences on Artificial Intelligence, 2015, pp. 2327–2333.
- 8. Fletcher, T., "Machine Learning for Financial Market Prediction", *University College London*, 2011, pp. 17–33.
- Glaeser E. L., Hillis A., Kominers S. D., Luca M., "Predictive Cities Crowdsourcing City Government: Using Tournaments to Improve Inspection Accuracy", *American Economic Review*, 2016, 106, pp.114–118.
- 10. Hochreiter S., Schmidhuber J., "Long short-term memory", *Neural computation*, 1997, 9 (8), pp. 1735–1780.
- Hutchinson J., Lo A., Poggio T., "A Nonparametric Approach to the Pricing and Hedging of Derivative Securities Via Learning Networks", *Journal of Finance*, 1994, 49, pp. 851–889.
- 12. Jasny B. R., Stone R., "Prediction and its limits", *Science*, 2017, vol. 355, issue 6324, pp. 468–469.
- 13. Kleinberg J., Ludwig J., Mullainathan S., Obermeyer Z., "Prediction Policy Problems", *American Economic Review*, 2015, 105, pp.491–495.
- Kuncheva L., Whitaker C., "Measures of diversity in classifier ensembles", *Machine Learning*, 2003, 51, pp. 181–207.
- Moravčík M., Schmid M., Burch N., Lisý V., Morrill D., Bard N., Davis T., Waugh K., Johanson M., Bowling M., "Deep Stack: Expert-level artificial intelligence in headsup no-limit poker", *Science*, 2017, 356 (6337), p.508.
- 16. Myerson R. B., "Game Theory: Analysis of Conflict", *Harvard University Press*, 1991, p. 1. Chapter-preview links, pp. vii–xi.
- 17. Naik N., Raskar R., Hidalgo, C. A., "Cities Are Physical Too: Using Computer Vision

- to Measure the Quality and Impact of Urban Appearance", *American Economic Review*, 2016, 106, pp. 128–132.
- Nelson D.M.Q., Pereira A.C.M., de Oliveira R.A., "Stock market's price movement prediction with LSTM neural networks", Neural Networks (IJCNN), 2017 International Joint Conference on IEEE, 2017, pp. 1419–1426.
- 19. Thaler R., "Mental accounting and consumer choice", *Marketing science*, 1985, 4 (3), pp. 199–214.
- 20. Polikar R., "Ensemble based systems in decision making", *IEEE Circuits and Systems Magazine*, 2006, 6 (3), pp. 21–45.
- 21. Qian N., "On the momentum term in gradient descent learning algorithms", *Neural Networks*, 1999, 12 (1), pp.145–151.
- 22. Rokach L., "Ensemble-based classifiers", *Artificial Intelligence Review*, 2010, 33 (1–2), pp.1–39.
- 23. Rumelhart D.E., Hinton G. E., Williams R. J., "Learning representations by back-propagating errors", *Nature*, 1986, 323 (6088), pp.533–536.
- 24. Shapiro C., "The Theory of Business Strategy", *RAND Journal of Economics*, 1989, 20 (1), pp. 125–137, JSTOR 2555656.