

doi:10.16060/j.cnki.issn2095-8072.2018.06.005

中国股票市场的趋势预测 ——基于多维交互验证法^{*}

林仁文¹ 杨 熠²

(1. 上海对外经贸大学, 上海 201620; 2. 华东政法大学, 上海 201620)

摘 要: 股票市场是现代金融的核心组成部分, 准确有效地预测其运行规律, 对于投资者和管理者来说都极为重要。尽管大量研究表明, 股票市场是一个高度复杂的非线性动态系统, 其变化规律具有一定的趋势性, 但如何预测与确认其趋势一直都是困扰投资者的难题。本文提出了多维交互验证法, 尝试从驱动面、心理面、行为面等多个维度出发, 来预测股票市场的趋势。本文还针对上证综合指数建立多维决策树模型, 对该方法进行检验。实证分析的结果表明: 多维交互验证法在预测股票市场的月度趋势方面, 具有较好的有效性和可行性。

关键词: 多维交互验证法; 趋势预测; 决策树模型

中图分类号: F830.91

文献标识码: A

文章编号: 2095—8072(2018)06—0040—10

一、引言

2018年3月以来特朗普政府针对进口钢铁和铝产品征收高额关税的影响席卷全球金融市场。随之而来的中美贸易战, 造成了中美股市的巨大波动。如何及时预测股票趋势并制定对策, 对于防范金融系统波动有重要意义, 并已成为当前金融研究的热点。

从文献上看, 股票市场是现代金融的核心组成部分, 大量学者致力于股票趋势的预测, 并形成了一系列纷繁复杂的预测方法。其中, 格雷厄姆和大卫·多德是该研究领域的先驱。他们在1934年提出了基本分析方法(fundamental analysis)预测股票价格, 该方法使用简单, 但精确度不高。随后, 人们开始采用技术分析方法, 以提高股票趋势预测的精确度, 代表性文章包括: Alexander(1961)、Kho(1996)、Lento和Gradojevic(2011)、Atsalakis等(2016)和Liu等(2016)等。此外, 预测股市趋势的方法还包括行为金融预测方法和时间序列分析预测方法, 如Bolton等(2006)、Matsumura等(2013)和Shi等(2013)等。另一方面, 随着计算机技术的发展, 各类非线性的分析预测方法(如神经网络、模糊逻辑和混沌预测方法等)也被广泛应用于股票市场, 如: White(1988)利用神经网络模型预测IBM股票的每日收益率; Tsaih等(1998)

^{*} 基金项目: 本文受到国家社科基金重大项目(15ZDA058)、国家自然科学基金资助项目(11501355和71571116)、中国博士后科学基金第59批面上资助项目(167567)、教育部人文社会科学研究项目规划基金项目(15YJA790039)、上海市教育委员会科研创新项目(15ZZ090)、上海市哲学社会科学规划课题青年项目(2016EJL002)和华东政法大学科学研究项目(16HZK014)资助。

运用神经网络模型对S&P500指数波动进行了预测检验；Bildirici和Ersin(2009)等运用人工神经网络扩展的GARCH模型族对股票交易进行预测。同时，还有学者结合多种方法进行估算，以期得到更理想的预测结果，比如Kuo等(2001)将遗传算法和模糊神经网络相结合，开发出混合模型GFNN；Chakravarty和Dash(2012)提出模糊神经网络模型，用以分析预测S&P500指数和孟买股票指数。

国内也有大量研究预测股票走势，其发展路径与国外文献基本一致。其中一类文献是基于传统时间序列模型及其扩展的研究，如韩豫峰和汪雄剑等(2014)基于技术分析工具，研究了中国股市趋势中的惯性效应；谭政勋和张欠(2016)基于ARFIMA模型，考察了上证指数趋势变化中的长期记忆性等。另一类文献在预测过程中引入了各类非线性的预测方法，包括非线性动力系统、支持向量机和神经网络系统等，如刘文财和刘豹等(2002)运用混沌动力系统理论，预测了上证综合指数、深证成分指数；龙建成和李小平(2005)基于神经网络，对股票市场趋势进行了趋势预测；秦宇(2008)基于经验模态分解方法，研究了上海股票市场价格的趋势特征；隋学深和齐中英(2008)、陈海英(2013)应用支持向量机对股票市场趋势进行预测；邱望仁 and 刘晓东(2013)基于FCM的广义模糊时间序列模型，研究了沪指的趋势特征；张贵生和张信东(2016)基于近邻互信息的SVM-GARCH模型预测股票价格趋势。

然而，关于股市趋势预测的已有研究，仍存在一定改进空间。其中，大量研究往往局限于股市某一方面的影响因素进行趋势预测，导致其预测结果具有一定片面性。尽管也有研究结合了多个影响因素进行趋势预测，比如近年来的机器学习方法，但由于缺乏系统标准来筛选所采用的影响因素，其建模效果并不理想。

因此，本文提出了多维交互验证法，来预测股市的趋势。具体而言，本文结合驱动因素、心理因素和行为因素这三个维度的影响因素，通过它们之间的交互验证预测股市走势，从而提高趋势预测的可靠性和准确性。此外，还针对上证综指建立了相应的决策树模型，以验证本文方法的可靠性。实证结果表明：多维交互验证方法在预测股市趋势方面具有一定的有效性和可行性，能够对投资者的投资策略产生一定的指导作用。

二、股市趋势的影响因素分析

（一）驱动维度分析

驱动因素是市场趋势形成的根本原因。影响股市趋势的驱动因素有很多，譬如经济周期、市场流动性、通胀、行业与上市公司的业绩、股市政策等等。此外，政治因素以及自然灾害等因素，对于股市的影响也不可忽略。

本文从经济周期和市场流动性两个影响因素出发，分析它们对股市趋势的影响。这是由于相比其他驱动因素，经济周期和市场流动性对于预测市场表现最为直接有效：通胀的作用一定程度上体现在经济周期中；股市政策基本是逆周期的，其制定与经济周

期所处的阶段、通胀水平以及流动性状况都有密切联系；行业与公司业绩对市场的影响，主要是通过影响局部板块进而影响市场的整体走势来完成，一般来说投资者难以很好地把握它们对市场整体趋势的影响程度；突发性事件由于其本身的不可重复性以及数据量小等局限性，很难被精确地度量，因此对实践的指导意义也较有限。

首先，本文采用“美林投资时钟”的方法，分析驱动因素对于股市趋势的影响。

“美林投资时钟”是一种将经济周期与资产和行业轮动联系起来的基本方法，它将以股票为主的金融市场分为4个阶段(见表1)，来帮助交易者识别经济的转折点，从而预测大盘趋势。因此，通过基于美林投资时钟的经济周期分析，可以较为准确地预测股市的顶、底，为市场趋势的判断打下基础。

表 1 以股票为主的金融市场的经济周期：基于“美林投资时钟”的经济周期分析

1. **衰退阶段：**经济增长率和通胀率都在下降，流动性已基本见底，但由于实体经济缺乏好的投资机会，因此充裕的流动性很可能流入股票市场，从而表现为股市往往在经济衰退的中后期见底。
2. **复苏阶段：**经济进一步增长，同时没有加息的预期，所以股市在经济复苏阶段一般处于上涨态势。
3. **繁荣阶段：**经济加速增长，收益率曲线向上移动且逐步变得平坦，此时的债券价格开始下降，使得债券不再是好的投资标的。然而，工业企业的利润步入最后冲刺阶段，这使得绝大多数股市在该阶段构筑顶部，而原材料企业的利润仍处于上升阶段，因此对于能源企业和原材料企业占比较高的股市而言，顶部的构筑可能要等到滞胀阶段。经济加速增长，收益率曲线向上移动且逐步变得平坦。此时，债券价格开始下降，使得债券不再成为好的投资标的。工业企业的利润步入最后冲刺阶段，原材料企业的利润仍旧处于上升阶段。对于绝大多数股市来说，这个阶段一般是构筑顶部的阶段，而对于能源企业和原材料企业占比较高的股市而言，顶部的构筑可能要等到滞胀阶段。
4. **滞胀阶段：**实际的经济增长率逐步降低，并最终低于潜在经济增长率，但通胀水平却在不断上升。该阶段可能存在先前超发货币引发的产业链延伸过长而导致的滞胀。并且在这个阶段中，股市继续下跌，大宗商品也将在该阶段的中后期构筑顶部。

其次，本文从流动性角度，对市场趋势做进一步分析。这是因为，股市趋势同时受基本面的存量推动和资金面的流量推动，所以仅通过经济周期进行趋势判断略显不足，有必要结合流动性再做分析。本文采用货币供应量，作为代表性的流动性变量，以考察流动性和股市趋势间的关系。原因在于，货币供应量能够对股票市场产生显著的影响：当货币供应量增加时，将通过财富效应促使投资者购买公司的股票，从而引发股价上涨；而股价上升，反过来又将提高投资者的投资和消费欲望，从而进一步催生更高的货币需求。

(二) 心理维度分析

对市场趋势进行精确预测，还离不开心理维度的分析。原因在于，尽管驱动因素是市场趋势形成的根本原因，但驱动因素需要经过心理因素的过滤才能作用于市场，并促使市场产生周期性的变化，以及震荡和单边趋势交替出现。

市场情绪的三阶段模型是最简单的心理分析模型，如果投资者能够把握市场主流意见处于何种阶段，就能比较准确的判断预测市场的趋势。第一阶段(酝酿阶段)：股价处于不规则的震荡状态，没有明显的趋势。在该阶段，市场缺乏主流观点(主流观点渗透率一般低于50%)，投资者没有投资欲望，市场参与度很低，从而使得市场成交量惨淡。第二阶段(发展阶段)：股价开始了单边趋势，机构投资者一般在该阶段的早期就已介入，并一直持仓到该阶段的末期。在该阶段，随着驱动因素的变化，市场开始形成对其变化的主流观点，并且市场主流观点的渗透率逐步增加、直至超过50%的比例。第三阶段(反转阶段)：股价开始反转，进入下跌趋势。散户交易者在该阶段的末期才开始进入市场，但此时已没有更多的买家进场，缺乏进一步推动行情的资金，从而导致股价反转下跌。

总之，通过观察投资者本身的情绪状态，就可以基本判断市场所处的情绪阶段，进而判断未来的股票价格趋势。而投资者的情绪状态，可以通过如下心理维度指标进行分析：涨跌比率、市场关注度、封面指数、开户数和顶部底部心理特征等。

(三) 行为维度分析

行为维度的分析，即根据市场行为、运用技术分析手段来预测推断股价未来走势。股票市场的行为可以被简单总结为“价量时空”——价格、成交量、价量的变化、完成价量变化所经历的时间和空间。当趋势开始酝酿形成时，会通过某种市场行为的特征加以展现。因此，通过分析这些市场行为，投资者就能够对趋势形成进行确认。

目前，有效的行为维度分析方法包括：3N法则和P-V直角坐标系价量分析方法。其中，3N法则是一种比较高效的趋势识别方法，它是由空间上的N%法则、形态上的N字法则和时间上的N期法则三者结合构成。而P-V直角坐标系价量分析方法，是通过模仿经济学分析供求关系，将价量关系投射到直角坐标系中进行分析，可以简化推理过程。

三、多维交互验证分析

本文采用多维交互验证的方法，进行股市的趋势预测。在这里，多维交互验证包括维度内的交互验证和维度之间的交互验证。

(一) 维度内的交互验证

按照三个维度对市场影响的顺序，维度内交互验证的具体步骤是：

首先，进行驱动维度的分析。利用美林时钟划分经济周期，然后根据经济周期所处的阶段，判断市场是否处于趋势所能产生的时间窗口。其中，上升趋势产生于衰退阶段中后期，即通胀率高点到增长率低点的阶段；下跌趋势产生于繁荣时期，即通胀率低点到增长率高点的阶段。接下来，将该结果与流动性做印证，从而完成驱动维度内的交互验证。

其次, 进行心理维度的分析。考察开户数、涨跌家数、封面指数、情绪指数等心理指标, 通过它们之间的相互印证, 再依据帝娜心理分析框架来考察市场情绪处于什么状况, 从而确定市场趋势所处的阶段。

最后, 进行行为维度的分析。先通过3N法则, 从空间、形态和时间上判断股市趋势是否成立, 然后用P-V直角坐标系从供求的角度进行分析, 以验证上述驱动维度和心理维度的判断, 从而完成整个交互验证的过程。

(二) 维度之间的交互印证

本文结合驱动、心理和行为三个维度的分析结果, 进行相互印证。理想状态下, 各维度对市场趋势的预测结果是一致的; 如果预测结果不一致, 则需要重新验证。

各维度间的交互印证过程为: 如果某维度的市场趋势预测结果, 与上一维度的预测相一致, 那么可以继续下一维度的分析; 反之, 则需要重新进行本维度的分析, 或回溯到上一维度进行再次验证。如果各维度的预测结果始终不一致, 就需要查看各维度中大部分影响因素的预测结果是否一致: (1) 如果各维度中大部分影响因素的预测结果一致, 则维持大部分因素的判断, 并找出不一致的原因并加以合理解释, 否则就需要

适当调低预测结果准确率的期望值; (2) 如果各维度中大部分因素预测结果的分歧较大, 则有可能是市场趋势改变的征兆。图1是多维交互验证的流程图, 展示了从驱动维度、到心理维度、到行为维度的整个检验的递进过程。

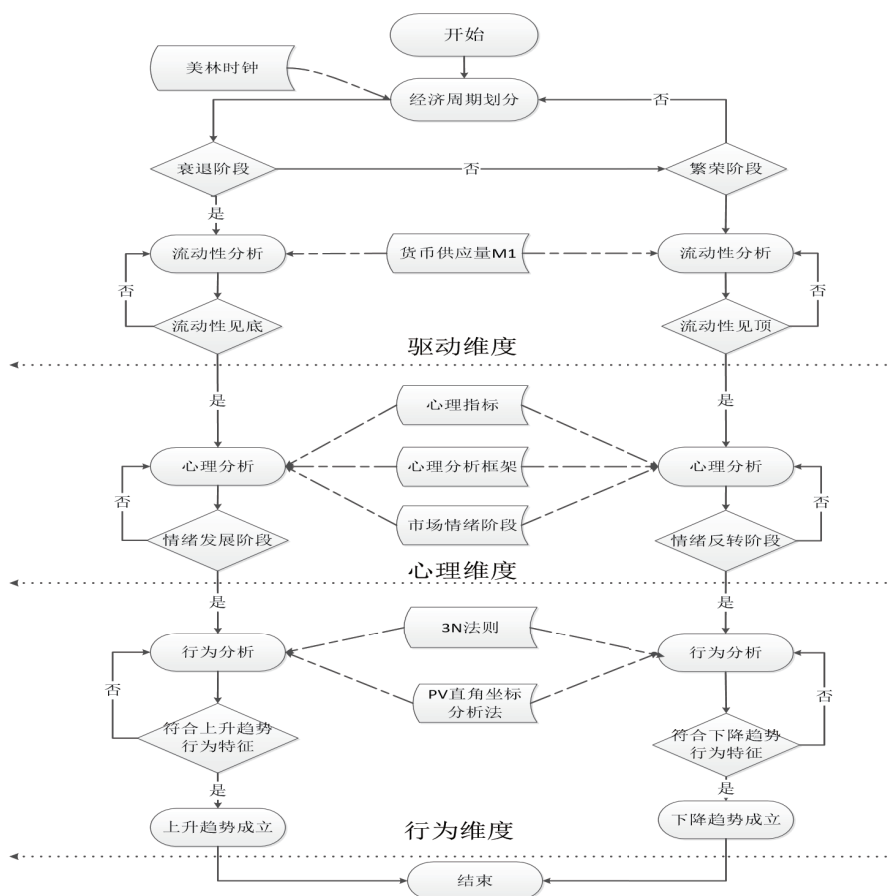


图1 多维交互验证流程图

四、决策树模型

决策树模型源于人工智能领域的机器学习技术，可以应用于数据分类和预测。作为经典的研究方法，它采用逼近离散函数值的形式，自顶向下递归构造而成。

早在20世纪60年代，决策树方法就已产生并广泛应用，而且发展出了多种算法。其中，Quinlan(1986)提出了ID3算法，可以有效减少树的深度。随后，在ID3算法基础上，决策树算法被进一步拓展，包括：C4.5、CART、CHAID、QUEST等。其中，基于信息论方法(Information Theory)构建的C4.5算法(Quinlan, 1993)在预测变量的缺值处理、剪枝技术以及派生规则等方面都有了很大改进，使之计算复杂度低、处理速度快、适用范围较广(Berretti, Thampi and Srivastava, 2015)。

本文采用C4.5算法，建立决策树模型，将信息增益率(Information Gain Ratio)定义为：

$$GainRatio(A) = \frac{Gain(A)}{SplitInfo(A)}, \quad (1)$$

$$\text{其中, } Gain(A) = Info(D) - Info_A(D), \quad (2)$$

$$Info(D) = -\sum_{i=1}^m p_i \log_2(p_i) \text{ 和 } Info_A(D) = \sum_{j=1}^v \frac{|D_j|}{|D|} Info(D_j), \quad (3)$$

$$SplitInfo_A(D) = -\sum_{j=1}^v \frac{|D_j|}{|D|} \log_2 \left(\frac{|D_j|}{|D|} \right), \quad (4)$$

其中， p_i 是子集中不同性(二元分类，即正样例和负样例)样例的比例， j 是属性 A 中的索引。 D 是集合样本， D_j 是 D 集中属性 A 上值等于 j 的样本集合。分裂信息 $SplitInfo(A)$ (Split Information)刻画了按照属性 A 分裂样本集合 D 的广度和均匀性。

五、模型应用

本文选取的样本为上证综指1994~2013年的月度收盘价，数据来自证券之星网站(www.stockstar.com)。本文采用月度数据的原因在于：趋势分析中，周期太短将难以形成明显的趋势，而周期太长则无法保证数据的充足性。下面，本文将驱动维度、心理维度和行为维度相结合，用这几个维度的月度历史数据作为输入并进行训练，以生成决策树模型，从而预测月度的股市趋势。

(一) 多维交互验证的分析结果

在上证股指的交易历史中，牛熊市交替出现。本文以2008年10月28日至2009年8月4日这段时间内的上升趋势为例，来实证分析多维交互验证的过程。^①

首先，在驱动维度方面，从经济周期和流动性的两个角度分析。经济周期的指标，选取月度工业增加值和月度CPI，而没有选用GDP和总体价格水平指标。其原因

^① 其中，2008年10月28日，上证指数最低杀跌到1664点；而2009年8月4日，上证指数最高反弹到3478点。

在于,虽然后者能够更好反映经济周期的走势,但缺乏月度数据,不利于及时判断股指趋势。分析结果显示,2008年金融危机导致世界范围内的股指剧跌,尽管中国央行推出多项利好政策(譬如持续降息、汇金回购、拓宽资金进入股市的渠道等),但仍无法阻止股市持续下挫。随后,中国经济在通胀率高点后自2008年二季度进入衰退期。根据美林投资时钟原理可以推算出衰退中后期孕育着股市的上升趋势。另一方面,还从流动性角度考察股市是否见底。我们选取M1作为流动性指标,这是由于中国的M1比M2波动更为剧烈,能够更好地反映流动性因素中的政策作用。结果显示,M1同比增速从2008年9月就降至10%以下;而国家推出4万亿投资计划后,M1同比增速迅速从10%以下开始回升,并逐步确认了流动性的底部。至此,驱动维度验证了趋势产生的条件已经具备。

其次,在心理维度方面。本文采用投资者新增开户数作为情绪指标。选取该指标的原因是:首先,在国内外研究市场情绪的文献中,投资者新增开户数指标都不可或缺;其次,投资者新增开户数能够直观反映投资者对证券市场的参与度及需求度,开户数越多、投资者情绪越高涨,因此本文采用开户数的同比增长率作为心理因素的代表性变量,并将其与上证指数的走势做对照(见图2)。从图2可以看出,2008年当股指跌至1664点,A股开户数的增长率也急速下降(股票基金的平均仓位已经降至其下限60%)。而随着4万亿投资计划的推出,开户数明显增多,基金仓位也开始回升。该结果表明,心理层面开始对驱动因素的变化做出积极反应,市场进入情绪发展阶段。同时,心理维度的预测,也验证了驱动维度的预测结果。

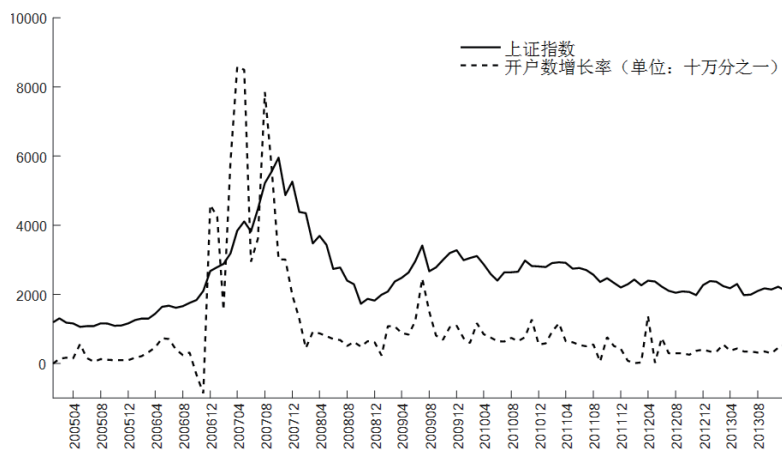


图2 2005年4月到2013年8月开户数增长率与上证指数的大盘走势关系

最后,在行为维度方面。本文选取股票价格和数量作为指标。这两个指标量化了股票的交易行为,是行为分析中的核心指标,同时也方便我们进行股票供给和需求分析。图3的股价和成交量直角坐标价涨量增图表明,多方势力开始加

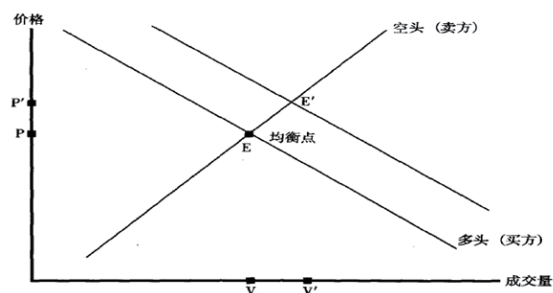


图3 股价和成交量直角坐标价涨量增图

强，而空方势力基本保持不变，从而将推动股价由P涨到P'、成交量由V增加到V'。根据哈尔兹法则和N期法则，股市经过N字发散收敛再发散将带来第二轮上涨，为后市的股指再一轮上涨趋势提供可能性。事实上自1664点开始，市场走出明显的复合N字底部，并出现价量齐增的现象。这样，行为维度的分析进一步确认了驱动和心理维度对股市上涨趋势的判断。至此，交互验证预测上涨趋势的过程全部完成。

上述分析表明，本文通过驱动、心理和行为的多维度交互验证的方法来判断股市趋势是可行的。

(二) 决策树模型的分析结果

本节详细说明决策树模型的分析结果，并讨论该模型对上证股指趋势预测的准确程度。决策树模型中的因变量为上证股指的下月趋势，赋值1或-1。其中，1代表上涨趋势；-1代表下跌趋势。模型中的自变量，来自驱动、心理和行为维度的代表性变量，自变量的具体定义如表2所示。

表2 源数据属性说明

维度	自变量	说明
行为	最近一个月涨幅(上证)	本月之前一个月的涨幅
	最近一个季度涨幅(上证)	本月之前三个月内的涨幅
	最近一年涨幅(上证)	本月之前一年内的涨幅
	月度成交手数(上证)	
驱动	月度M1同比增长率	
	月度M1环比增长率	
	月度工业增加值	由于GDP没有月度数据，故用月度工业增加值代替，该数据有极少部分缺失，本文用平均值填充法补齐数据
	月度CPI	
心理	月度涨跌家数比率(上证)	该比率等于月度上涨家数与下跌家数之比，由于其它情绪指标数据不足，故选择该指标作为心理维度

接下来，本文考察决策树模型对趋势预测的正确性。具体方法是，使用较早时间段的数据进行训练并在较晚时间段的数据上进行测试。1994到2013年间的月度数据中，选取后n年的数据为测试集，前20-n年的数据为训练集。本文将对n取不同值(n=0、1、3、5)，以对决策树模型进行预测。局限于20年的样本量，本文将决策树的深度设定为6，随机种子设定为2。决策树中属性的选择结果表明：在驱动因素中，流动性因素的影响最大，其次是CPI，然后才是工业增加值，这反映了短期内市场对经济增长的反映并不特别强烈，而对流动性和通胀率的变化更加敏感；在行为维度中，成交手数和最近一个季度涨幅占据了最为重要的位置，这说明在短期趋势预测中，行为因素的影响力较大。由于决策树模型同时包含了

表3 决策树分析结果

n	K值	正确预 测率	混淆矩阵			比值比
			真实值	预测值		
				-1	1	
0	3	75.8%	-1	103	12	14.74
			1	46	79	
1	1	66.7%	-1	7	0	8.75
			1	4	1	
3	3	61.1%	-1	10	11	4.00
			1	3	12	
5	2	63.3%	-1	18	11	2.98
			1	11	20	

注：当混淆矩阵中的出现0时，在计算比值比时将以0.2计算，这样将避免出现极值(0或无穷大)。

三个维度的因素,这进一步论证了多维交互验证预测趋势的方法在实证上的可行性。

表3是基于决策树模型的分析结果。其中,“正确预测率”一行说明的是决策树模型预测股市趋势的准确率。可以看到,预测的准确率均超过60%,而如果采用所有数据同时做训练集和预测集,那么预测的准确度将高达75.8%。这意味着,决策树模型对市场未来趋势的可预测性。“比值比”(odds ratio)一行说明的是决策树模型预测股市趋势的有效性,所有的比值比都大于1,这说明了用决策树对股市未来预测的效果较好。

六、结论

本文提出了一种新的股票市场预测方法,即多维交互验证预测法,来预测股市的趋势。该方法将影响股票市场趋势的驱动维度、心理维度、行为维度相结合,并通过各个维度内及维度之间的交互验证法来增强趋势预测的正确性。此外,本文还针对上证综合指数、建立多维决策树模型,对多维交互验证预测法进行检验。实证分析的结果表明,多维交互验证法在预测股票市场月趋势预测时,具有较好的有效性和可行性。

然而,本文提出的多维交互验证预测法,仍存在进一步挖掘改进的空间。接下来的研究工作包括:更深入地研究本文所涉及的三种维度的相互作用关系,从而提高交互验证的效果;同时,尝试将预测范围拓展到更短周期的趋势预测上,并逐渐细化到行业或个股,以对投资者的实际操作提供更好的指导。

参考文献

- [1] 陈海英. 基于支持向量机的上证指数预测和分析[J]. 计算机仿真, 2013(1): 297-300.
- [2] 韩豫峰, 汪雄剑, 周国富等. 中国股票市场是否存在趋势?[J]. 金融研究, 2014(3): 152-163.
- [3] 刘文财, 刘豹, 张维. 中国股票市场混沌动力学预测模型[J]. 系统工程理论方法应用, 2002(1): 12-14.
- [4] 龙建成, 李小平. 基于神经网络的股票市场趋势预测[J]. 西安电子科技大学学报, 2005(3): 460-463.
- [5] 秦宇. 应用经验模态分解的上海股票市场价格趋势分解及周期性分析[J]. 中国管理科学, 2008(S1): 219-225.
- [6] 邱望仁, 刘晓东. 基于FCM的广义模糊时间序列模型[J]. 模糊系统与数学, 2013(6): 111-117.
- [7] 隋学深, 齐中英. 基于多尺度特征和支持向量机的股市趋势预测[J]. 哈尔滨工业大学学报(社会科学版), 2008(4): 77-82.
- [8] 谭政勋, 张欠. 中国股票市场的长期记忆性与趋势预测研究[J]. 统计研究, 2016(10): 57-66.
- [9] 张贵生, 张信东. 基于近邻互信息的SVM-GARCH股票价格预测模型研究[J]. 中国管理科学, 2016(9): 11-20.
- [10] Alexander, S.S., "Price Movements in Speculative Markets: Trends or Random Walks", *Industrial Management Review of Mit*, 1961, 2(2): 7.
- [11] Atsalakis, G. S., E. E. Protopapadakis and K. P. Valavanis, "Stock Trend Forecasting in Turbulent Market Periods Using Neuro-Fuzzy Systems", *Operational Research*, 2016, 16(2): 245-269.
- [12] Berretti, S., S. M. Thampi and P. R. Srivastava, *Intelligent Systems Technologies and Applications*, Springer International Publishing, 2015.
- [13] Bildirici, M. and Ö. Ö. Ersin, "Improving Forecasts of GARCH Family Models with the Artificial Neural Networks: An Application to the Daily Returns in Istanbul Stock Exchange", *Expert Systems with*

- Applications*, 2009, 36(4): 7355–7362.
- [14] Bolton, P., J. Scheinkman and W. Xiong, “Executive Compensation and Short-Termist Behaviour in Speculative Markets”, *Review of Economic Studies*, 2006, 73(3): 577–610.
- [15] Chakravarty, S. and P. K. Dash, “A PSO Based Integrated Functional Link Net and Interval Type-2 Fuzzy Logic System for Predicting Stock Market Indices”, *Applied Soft Computing*, 2012, 12(2): 931–941.
- [16] Kho, B., “Time-Varying Risk Premia, Volatility, and Technical Trading Rule Profits: Evidence from Foreign Currency Futures Markets”, *Journal of Financial Economics*, 1996, 41(2): 249–290.
- [17] Kuo, R. J., C. H. Chen and Y. C. Hwang, “An Intelligent Stock Trading Decision Support System through Integration of Genetic Algorithm Based Fuzzy Neural Network and Artificial Neural Network”, *Fuzzy Sets and Systems*, 2001, 118(1): 21–45.
- [18] Lento, C. and N. Gradojevic, “The Profitability of Technical Trading Rules: A Combined Signal Approach”, *Journal of Applied Business Research (JABR)*, 2011, 23(1).
- [19] Liu, C., J. Wang and D. Xiao, et al, “Forecasting S&P 500 Stock Index Using Statistical Learning Models”, *Open Journal of Statistics*, 2016, 6(6): 1067.
- [20] Matsumura, K. and M. Kawamoto, “Derivative Trade Optimizing Model Utilizing GP Based on Behavioral Finance Theory”, *Electronics and Communications in Japan*, 2013, 96(4): 15–28.
- [21] Quinlan, J. R., “Induction of Decision Trees”, *Machine Learning*, 1986, 1(1): 81–106.
- [22] Quinlan, J. R., *C4.5: Programs for Machine Learning*, San Francisco Morgan Kaufmann Publishers, 1993.
- [23] Shi, W. and P. Shang, “Cross-Sample Entropy Statistic as a Measure of Synchronism and Cross-Correlation of Stock Markets”, *Nonlinear Dynamics*, 2013, 71(3): 539–554.
- [24] Tsaih, R., Y. Hsu and C. C. Lai, “Forecasting S&P 500 Stock Index Futures with a Hybrid AI System”, *Decision Support Systems*, 1998, 23(2): 161–174.
- [25] White, H., “Economic Prediction Using Neural Networks: The Case of IBM Daily Stock Returns”, *IEEE*, 1988: 451–458.

【作者简介】林仁文：上海对外经贸大学金融管理学院讲师，经济学博士。研究方向：宏观金融，数量经济。

杨 熠：华东政法大学商学院讲师，经济学博士。研究方向：货币政策，资产定价。

Stock Market Trend Prediction: Based on Multi-Dimensional Cross-Validation Method

LIN Ren-wen¹ & YANG Yi²

(1. Shanghai University of International Business and Economics, Shanghai 201620, China; 2. East China University of Political Science and Law, Shanghai 201620, China)

Abstract: The stock market is the key component to the modern finance. Forecasting its trend accurately and effectively, is of primary significance. Many researchers have already found that the stock market is a highly complicated nonlinear dynamic system, of which certain trend does exist. Nonetheless, how to predict the trend and make use of it remains an unsolved problem. The authors come up with a new method, multi-dimension cross-validation approach, to predict the stock market's trend, from the aspects of driving factors, psychological factors and behavioral factors. In the end, we build a statistical decision tree model to predict the Shanghai Composite Index's trend to testify our new method. The result shows that, our model performs very well with historical data, and it is feasible and reliable to employ our approach to forecast the market's trend.

Keywords: multi-dimension cross-validation approach; trend prediction; decision tree

(责任编辑：马莹)