

# 中国上市公司“高送转”动因研究

## ——基于BP神经网络模型方法分析

董克明 赵盛山

**摘要:**上市公司实施“高送转”股利分配方案是中国股票市场较为特殊的现象,先前的研究主要通过Logistic或Probit等线性模型来分析送转动因。本文基于2014—2016年沪深两市所有公司的数据,引入人工神经网络方法,建立了“高送转”预测模型。实证结果表明,该模型的预测准确率能够达到92%左右,较大的优于Logistic模型,且模型很好的支持了“股本扩张论”。此外,基于对模型的分析,本文发现上市公司的高送转行为与公司本身成长阶段有关,并提出了一种与上市公司成长阶段相关的高送转动因的理论解释。

**关键词:**高送转;神经网络模型;送转动因

**JEL分类号:**G23,G11,M31

## 一、引言

在股票市场上,支付现金股利和股票股利是上市公司进行股利分配的主要支付方式,上市公司采取送股或以公积金转增股本的形式发放股票股利的行为称为上市公司的送转行为,送转股比例较大的送转行为被称为“高送转<sup>①</sup>”。虽然“高送转”并没有改变公司的投资基本面,但是各国的实证研究表明实施“高送转”方案的股票能在短期或中长期获得不同程度的超额收益。美国市场数据表明,公司股票拆分公告前后,市场会给予公司股票以超额的溢价,在公告实施前市场给予的溢价比拆股方案实施后更高(Eng等人,2014)。Babitha Rohit等(2016)基于印度孟买证券市场的数据,检验并证明了印度半强有效市场给予发布拆股或配股的公司股票以非理性的溢价的统计显著性。陈珠明和史余森(2010)通过中国证券市场的数据,证明了除权除息日前后“高送转”股票会出现统计显著的超额收益。由此可见,分析和预测上市公司的“高送转”行为对于送转动因理论的完善与发展具有非常重要的作用,有助于为投资者和企业管理者提供合理的决策参考,对于进一步完善和发展中国上市公司的股利分配政策、提升中国资本市场的国际化水平具有重要的意义。

## 二、文献综述

### (一)对于“高送转”动因的解释理论

**作者简介** 董克明:浙江财经大学信息学院金融数据分析研究室讲师,博士;

赵盛山:浙江财经大学经济学院国际经济与贸易系。

①中国市场的送转股与国外市场的股票拆分为行类似,除特别说明外,本文不加以区分

根据 Modigliani 和 Miller(1961)的理论,在完美资本市场(Perfect capital market)条件下,公司的资本结构不会对公司的价值产生任何影响。然而 Black(1976)发现投资者对于股票股利存在着非理性偏好,他把这种现象称为“股利之谜”。由于股利之谜在证券市场上客观存在,因此引发了一系列对股利分配动因的研究。现有理论主要可概括为以下几种。

### 1.信号传递论

Bhattacharya(1979)最先提出了“信号传递理论”,该理论的主要观点是,由于上市公司管理者和投资者存在信息不对称,管理者希望通过实施股票股利形式分红或者股票拆分政策的方式,向市场释放出对公司未来发展积极乐观的信号。Grinblatt 等(1984)基于 1967-1976 年在华尔街日报上公布股票拆分方案公司股票的研究发现,发布送转公告的股票能给市场传递积极的信号,因此这类股票在公告后能获得显著的正收益。Miller 等(1985)论证了在信息不对称情况下,管理层利用股票分拆来吸引投资者行为的存在。John 等(1985)论证了信号传递的存在,公司内部人士会根据公司实际未来现金流的情况来选择股利政策并向市场传递信号。Brennan 等(1988)提出股票拆分会稀释每股盈余,因此只有公司在对未来盈利状况有信心的前提下才可能实施股票分拆方案。Asquith 等(1989)基于美国证券市场的数据,论证了美国证券市场支持信号传递论。陈浪南等(2000)运用事件研究法和上海股票市场的数据对股利政策的市场反应进行了实证研究。其结果表明,现金股利不能成为有效的信号传递机制,而股票股利确实具有比较明显的信号传递作用。

### 2.最适价格论

Baker 等(1993)通过邮件调查得到了纽交所和美国证券交易所实施拆股方案的 136 个公司样本数据,通过数据分析得出了这些公司的拆股动机。研究表明,拆股的主要动机是通过股票拆分增加公司股票的流动性,从而使公司股价维持在一个较为合理的范围。“最适价格论”的主要观点是,上市公司管理者倾向于将股票价格维持在一个合理的区间,因为股票价格过高会限制资金量较小的投资者介入,通过以股票股利发放或者股票拆分能够降低股票的单位价格,降低投资者的资金准入门槛,提高公司股票的流动性;在以成交股票数量计算交易佣金的市场中,股票价格过低则会增加交易费用。并且对美国的管理者股票股利政策的调查表明,上市公司管理者在制定股利分配政策时确实会考虑将公司股票价格维持在一个合理的区间。

### 3.股利迎合论

基于行为金融学角度出发,Baker 和 Wurgler(2004)首次提出了迎合理论(Catering Theory),该理论从上市公司管理者对投资者对送转行为的投机性角度进行了解释。该理论认为,管理者在不同时期制定股利分配政策的过程中会迎合该时期投资者的偏好,从而实现其自身利益最大化(Baker 等人, 2004; Li 等人, 2006)。黄娟娟和沈艺峰(2007)对于迎合理论的研究表明公司股利分配政策与股权集中程度相关,上市公司的股利分配政策主要是迎合了大股东的需要而非中小股东的利益。相关研究也表明,我国个人投资者偏好不以现金形式发放股利的公司,并给予此类公司股票相对较高的溢价,上市公司管理者迎合了投资者此类非理性行为,追求短期内公司股票价格的最大化。

### 4.价格幻觉论

何涛和陈小悦(2003)的研究指出,“信号传递论”和“最适价格论”无法解释中国上市公司的送转行为,并提出了“价格幻觉论”。该理论认为,实施送转方案的股票在除权除息日后名义价格下降,这种由送转造成的股价下降会使得投资者产生名义价格幻觉,从而影响其投资判断。上市公司的送转动机是通过利用市场的价格幻觉现象,间接提高公司的市值。Weld 等(2009)从名义价格角度分析了投资者的非理性行为。国外相关研究也表明了投资者价格幻觉的存在(Baker 等人, 2009)。后续研究者分别通过实证研究的方式对上述理论进行了研究(魏刚, 1998; 吕长江等人, 1999; 熊德华等人, 2007)

### 5.股本扩张论

薛祖云和刘万丽(2009)在对中国上市公司送转行为的相关研究中提出了“股本扩张论”,该理论的主要观点是,由于中国的政策原因,再融资受到上市时间、股本规模等一系列因素的限制,上市后股本规模较小的公司存在着相对强烈的股本扩张动机。

综上,对于“高送转”动机的研究相对较少,不同理论对于送转动机的解释存在着争议。

#### (二)机器学习方法在预测问题上的应用

对于中国上市公司“高送转”的动因的研究,学者一般采用 Logistic 和 Probit 模型来寻找解释“高送转”动因的相关因素或有效的解释模型。Ezell 等(1975)基于经验研究,提出了有过历史送转行为的公司再次送转概率会相对较高的假设。基于此假设,他们认为,可用现有的数据去预测并检验公司是否有过送转历史,以此来检验与送转行为相关的解释变量。Ezell 等(1975)的研究提高了送转预测的准确率,也一定程度上支持了基于历史数据来预测公司未来的送转情况的可能性。Hwang 等(2005)通过引入 Probit 模型,用股价、账面价值、利润增长率等变量来预测送转情况,并证明了这一系列变量的统计显著性。通过此模型,能够预测出一年范围内 30% 的送转公司股票,他的研究虽然没有得出用来预测送转行为的最优解释变量,却进一步证明了用包含公司状态信息的数据来预测送转情况的可能性。Kevin 等(2008)基于 Hwang(2005)的思路,对预测送转的 Probit 模型进行了改进,进一步提高了预测的准确率。龚慧云(2010)将“股利迎合论”加入到 Logistic 模型中,对中国股票市场的送转行为进行了研究。熊义明等(2012)将“信号传递论”、“最适价格论”等不同的送转理论加入到 Logistic 模型中,基于中国市场 2006-2010 年的高送转样本数据对不同理论进行检验并通过检验后的模型进行了“高送转”的预测,提供了中国市场的经验证据。李心丹、俞红海等(2014)利用 Logistic 和 Probit 模型来研究中国股票市场的“高送转”现象。然而到目前为止,还没任何研究通过采用机器学习方法来对“高送转”进行拟合及预测。

机器学习方法对于无法通过理论分析得出规律的复杂问题有着极其优越的拟合效果,近年来,机器学习算法凭借其高效的拟合预测效果,逐渐在各个研究领域被广泛应用,部分学者通过引入人工神经网络、支持向量机等机器学习算法来尝试解决相关分类与预测的问题。杨一文等(2005)通过支持向量机对上证综指序列趋势进行了多步预测。高洪深等(1999)提出了基于人工神经网络的非线性回归预测模型的有效性。王宜怀等(2004)进一步证明了基于神经网络的非线性回归对于解决无法得出具体函数表达的或者本身就没有具体的函数表达的问题更具有有效性和优越性。刘湘云等(2016)通过小波神经网络对股指期货市场的风险预警进行了研究。刘笑天和赵胜民(2017)将 BP 神经网络与在线学习算法相结合对股票价格的时间序列进行预测。

本文基于 2014-2016 年的沪深两市所有公司的观测数据,利用 BP 人工神经网络算法,建立高送转预测模型,且对所建立的模型所蕴含的高送转动因进行了解释分析。本文的贡献与创新之处在于:(1)高送转模型为非线性模型,然而迄今所有研究都通过线性模型拟合数据建立预测模型,因此模型的预测准确率较低。本文率先引进人工神经网络算法,该算法可以拟合非线性数据,实验结果表明该算法能显著提高“高送转”的预测准确率。(2)本次研究将目前主要的高送转动机理论均纳入预测模型,最终建立模型支持了“股本扩张论”。(3)以往的研究均是通过 Logistic 等线性模型来研究各自变量独立对高送转的线性影响,本文利用人工神经网络具有层次结构这一特性,研究了多自变量对高送转行为的影响。最终基于对训练模型的分析,本文提出了一种新的高送转动机解释理论,补充完善了现有理论体系。

## 三、研究设计

### (一)样本的选择与数据来源

本文的样本为2014-2016年沪、深两市所有的上市公司股票,剔除极少部分数据不完整的样本数据后,最终样本数量为8464。上市公司在一季报和三季报中发布股利分配方案的极少,半年报次之,大部分上市公司集中在年报公布年度分配方案,因此我们选择上市公司的三季报数据来预测该年度的年报送转情况。模型的输入数据全部源自于三季报显示的数据,最终年报显示的送转方案为模型输出结果的判断依据,样本数据源自Wind资讯。

表1 2014-2016年度“高送转”情况统计

年份	股票总数	“高送转” 股票总数	“高送转” 比例	10送20 以上股票数	最高送转比例
2014	3510	447	12.74%	12	10送20
2015	2367	403	17.00%	11	10送20
2016	2587	424	16.40%	44	10送30
合计	8464	1274	15.05%	67	10送30

如表1所示,2014-2015年期间,“高送转”比例有所上升。2015年-2016年间,“高送转”比例虽然相对稳定,但选择股票股利分红的公司的送转股比例却有所上升,实施10送20及以上的大比例送转的公司数量越来越多,2016年甚至开始出现10送30这样超大比例送转股的公司。

## (二)变量定义与说明

本部分设计思路是首先根据已有文献的研究结果,将“高送转”动机的影响因素全部纳入模型,然后将样本数据分为独立的两组——训练数据和测试数据。然后用训练集数据通过人工神经网络算法对模型进行训练,训练完毕后,用测试集数据进行样本预测效果的测试,以此来判断模型的有效性和准确性。加入模型的解释变量来自于上述文献涉及的相关理论。

根据“信号传递论”的观点,上市公司通过实施“高送转”向市场释放公司未来前景乐观的信号。市场对“高送转”的预期在送转公告前就已经发生,在半年报和年报发布前,投资者就会对有送转预期的公司股票给予超额溢价,即通常所说的“埋伏年报高送转”或“埋伏半年报高送转”。预期通常基于公司的留存收益和公司的成长性。

留存收益是实施“高送转”的必要条件,留存收益越高并不意味着送转概率会越大,但送转比例的上限会更高,从市场角度来看,投资者反应也会越强烈。因此,我们将每股收益、每股资本公积作为留存收益的指标纳入模型。

公司的成长性是公司发展速度的反映,近几年的成长性指标一定程度上反映了公司的发展趋势。如果按照“信号传递论”的观点,实施“高送转”能够向市场传递积极的信号,那么,公告实施前公司的成长性指标也会对“高送转”产生影响。因此,我们将今年前三季度的净利润增速、去年以及前年的前三季度的净利润增速作为成长性指标纳入模型。

根据“最适价格论”的观点,公司股票的单位价格会对管理者制定股利分配方案的时候产生影响。公



司股票价格过高、流动性较差的时候会考虑实施送转方案。因此,我们将股价、实际流通股数量作为反映公司股票流动性的指标,加入预测模型中。

根据“价格幻觉论”的观点,名义价格幻觉使得投资者愿意为“高送转”的股票承担超额的非理性溢价。若送转比例越大,除权除息后名义价格下降的幅度就越大,因此送转的比例越大,则名义价格幻觉越明显。

由于公司的每股公积金情况是送转比例上限的决定性因素,因此解释变量中需要包含反映公积金情况的指标。由于中国上市公司的送转股大多以资本公积转增股本,因此本文选取了每股资本公积来反映公积金情况。

根据“股本扩张论”的观点,股本较小的上市公司有扩大其股本的意愿,因此公司的总股本情况会对管理者股利分配方案的制定产生一定的影响。由于制度的原因,次新股的股本数量相对较少,股本扩张欲望强烈,因此我们同时将股本数量、是否为次新股作为变量加入模型。

根据“股利迎合论”的观点,管理者股利方案的制定会迎合市场的需求。首先,管理者的迎合动机与公司的属性有关,国企的管理者迎合的动机低于私企的管理者。其次,私企管理者的控股比例会影响其控制能力进而影响其迎合能力。再者,公司前几年的送转情况是公司迎合意愿的反映。综合以上三个因素,我们将企业的公私属性、第一大股东的控股比例、半年报送转、去年送转情况、前年送转情况、纳入模型当中。

基于上述分析,本文模型中的变量具体如表2所示:

表2 变量名称及定义

变量	变量符号	变量名称	变量定义	对应理论
因变量	<i>H</i>	高送转	每10股转增比例达到5股及以上记为“1”,否则记为“0”。	
	<i>eps</i>	每股收益	前三季度每股收益	
	<i>gro</i>	当年业绩增速	前三季度的净利润增速(%)	信号
	<i>groLY</i>	去年业绩增速	去年前三季度的净利润增速(%)	传递论
自变量	<i>groLLY</i>	前年业绩增速	前年前三季度的净利润增速(%)	
	<i>price</i>	股价	季报收盘价	最适
	<i>circul</i>	流通股	季报流通股数量(亿股)	价格论
	<i>res</i>	公积金	每股资本公积	价格幻觉论
	<i>scale</i>	股本数量	季报显示的总股本(亿股)	股本
	<i>new</i>	次新股	次新股记为1,否则记为0	扩张论
	<i>pri</i>	公私属性	公司是否为私企,国企记为0,私企记为1	
	<i>ctrl</i>	控股权	季报第一大股东的控股比例	股利
	<i>send</i>	半年报送转	半年报送转情况(每股送转比例)	迎合论
	<i>sendLY</i>	去年送转情况	去年送转情况(每股送转比例)	
	<i>sendLLY</i>	前年送转情况	前年送转情况(每股送转比例)	

### (三)本文所涉及的模型

#### 1.计量模型(Logistic)

为了研究各因素对高送转的影响,我们将表2涉及的所有解释变量加入到Logistic模型之中,得到如下计量模型:

$$H_t = \frac{1}{1 + e^{-L_t}} \quad (1)$$

其中

$$L_t = c + \beta_1 \text{eps} + \beta_2 \text{gro} + \beta_3 \text{groLY} + \beta_4 \text{groLLY} + \beta_5 \text{price} + \beta_6 \text{circul} + \beta_7 \text{res} + \beta_8 \text{scale} + \beta_9 \text{new} + \beta_{10} \text{pri} + \beta_{11} \text{ctrl} + \beta_{12} \text{send} + \beta_{13} \text{sendLY} + \beta_{14} \text{sendLLY} \quad (2)$$

$H_t$  表示因变量高送转,每10股转增比例达到5股及以上记为1,否则记为0;  $c$  为常数项,其他自变量含义参考表2。

#### 2.BP神经网络模型

人工神经网络算法是当今机器学习领域中比较热门的一类算法,已有研究证明三层人工神经网络<sup>①</sup>可以拟合任何连续有界函数(Cybenko, 1989; Hornik 等人, 1989),因此人工神经网络已经广泛应用于人脸识别,自动驾驶、气象预测等领域。人工神经网络模型是由大量处理单元单向加权相连而成的网络,该处理单元称为人工神经元。如图1所示,人工神经网络主要由输入层、隐藏层以及输出层组成。其中输入层神经元将输入值传递到隐藏层。隐藏层神经元将输入值加权求和,并通过激励函数产生一个输出值发送至输出层。隐藏层神经元的计算公式如公式(3)所示:

$$H = f(\vec{i} \times \vec{w}_{ih}) \quad (3)$$

其中  $\vec{i}$  为输入向量,  $\vec{w}_{ih}$  为隐藏层权重参数向量,  $H$  为隐藏层某个神经元的输出。函数  $f$  称为激励函数,一般选取可微非线性函数作为激励函数。本文选取的激励函数如下所示:

$$f(x) = \frac{1}{1 + e^{-x}} \quad (4)$$

上述函数又被称为sigmoid函数,采用sigmoid函数作为激励函数的隐藏层神经元又被称为sigmoid单元。

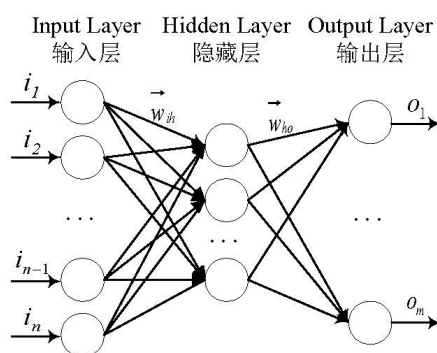


图1 人工神经网络结构

①符合该结论的神经网络需要满足以下条件:隐藏层采用sigmoid函数作为激励函数;输出层采用线性单元。

输出层的各神经元将隐藏层的计算结果加权求和并产生输出结果,所有输出神经元的输出结果即为神经网络模型的输出。输出层神经元计算公式如下:

$$o = \vec{h} \times \vec{w}_{ho} \quad (5)$$

其中  $\vec{h}$  为所有隐藏神经元的输出,  $\vec{w}_{ho}$  为输出层的权值参数。

模型输出与目标输出的差值称为误差,若误差不为0,则模型会根据误差调整其权值,使得误差向0逼近。根据误差调整权值的算法称为学习规则,人工神经网络可以根据学习规则的不同进行分类,其中BP神经网络是一类应用比较广泛的算法。在BP神经网络算法中,当误差非0时,算法会根据误差,依次修改输出层、隐藏层的连接权值,从而使得误差达到允许的范围之内,从而实现对数据的拟合。

假设现有  $K$  个训练数据组成的训练数据集,即  $\{(x_i, y_i)\}$ ,  $i = 1, \dots, K$ 。其中,  $x_i$  为输入值,  $y_i$  为目标输出。从节点  $i$  到节点  $j$  的输入表示为  $I_{ij}$ , 节点  $i$  到节点  $j$  的权值表示为  $w_{ij}$ 。BP神经网络学习过程具体如下:

第一步,初始化模型。人为设定隐藏层的节点数  $l$ ; 根据样本的输入输出维度确定输入层节点数  $n$  和输出层节点数  $m$ ; 初始化神经网络的参数,即对  $w_{ij}$  赋予接近于0的随机值

第二步,用训练数据  $(x_i, y_i)$  对模型进行训练,根据BP网络的学习规则进行参数调整,参数调整的具体步骤如下:

(1)将训练数据输入神经网络,通过公式(3)计算出隐藏层各个节点的输出  $H$ ; 然后通过公式(5)计算出模型输出  $o$ 。

(2)计算输出层节点的实际输出值  $o$  与目标输出值  $y$  之间的偏差,此偏差是当前模型的误差:

$$e_k = y_k - o_k, \text{ 其中 } k = 1, 2, \dots, m \quad (6)$$

(3)BP神经网络认为模型误差是各个节点产生的误差综合作用的结果,因此依次反向计算各个节点贡献的误差项。

1)对于网络的输出单元  $k$ , 计算它的误差项

$$\delta_k = o_k(1 - o_k)e_k \quad (7)$$

2)对于网络的隐藏单元  $h$ , 计算它的误差项

$$\delta_h = o_h(1 - o_h) \sum_{k=1}^m w_{kh} \delta_k \quad (8)$$

3)根据误差项,更新节点权值,其中  $\eta$  为学习速率,一般设置为接近于0的值

$$w_{ij} = w_{ij} + \Delta w_{ij}, \text{ 其中 } \Delta w_{ij} = \eta \delta_i I_{ij} \quad (9)$$

第三步,将训练数据依次放入模型,重复第二步训练方法,直到模型误差低于目标要求或者迭代次数达到一定设定值停止训练。至此神经网络训练完毕,可输入相应的  $x$  值利用该模型进行预测。

## 四、实证研究与结果

### (一)描述性统计与相关性分析

加入模型的各变量的描述性统计结果如表3所示,表4(见本文最后)给出了各变量之间的相关性分析。

表3 变量的描述性统计

变量	最大值	最小值	平均值	中位数	标准差
每股收益	9.36	-3.39	0.22	0.16	0.51
当年业绩增速(%)	61435.78	-141825	17.25	31.24	2634.52
去年业绩增速(%)	98058.14	-23885.1	-39.38	8.96	1577.04
前年业绩增速(%)	35041.74	-23885.1	-45.44	4.85	943.17
股价	190.31	1.91	16.59	12.96	12.91
流通股	2107.66	0.1	10.13	3.60	61.03
公积金	12.78	-0.96	1.70	1.34	1.51
股本	2943.88	0.41	14.05	4.73	93.74
次新股	1	0	0.06	0	0.242
公私属性	1	0	0.61	1	0.488
控股权	86.01	0.01	33.090	31.48	16.08
半年报送转	3	0	0.025	0	0.231
去年送转情况	3.1	0	0.219	0	0.489
前年送转情况	3	0	0.157	0	0.361

由相关性分析表可得,除股本与流通股存在高度相关外,其他各变量之间虽然存在一定的相关关系,但并非高度相关。如公积金和每股收益存在显著的正相关;股本和公积金存在显著的负相关。虽然股本与流通股存在高度相关,但这两个变量分别反映公司的两个不同指标,流通股数量直接反映公司股票的流动性,而总股本则会影响公司的股本扩张意愿。再者,对于人工神经网络模型来说,两者同时加入模型不会对模型整体构成太大的影响。

## (二)各模型最优变量与参数的选择

我们想通过引入人工神经网络算法来获得一个更加逼近于“高送转”决策的模型。首先,我们用随机抽样方法将数据集的80%作为训练数据集,10%做为验证数据集,10%作为测试数据集。然后,我们分别使用人工神经网络和Logistic模型对相同的训练数据集进行训练建模,最后利用测试数据集进行预测,通过对比两个模型的预测准确率来评价模型预测性能。

### 1. 计量模型(Logistic)

我们通过观测数据对模型各变量的显著性进行检验,以此来选取Logistic模型最合适的解释变量。表5给出了Logistic模型的回归结果的估计系数以及t统计量和对应P值。

回归结果显示,当年业绩增速( grow )、去年业绩增速( growLY )、前年业绩增速( growLLY )、流通股( circul )、次新股( new )、公私属性( pri )、控股权( ctrl )这几个变量统计上不显著。我们根据变量的显著性水平最终确定的Logistic模型为:

$$H_t = \frac{1}{1 + e^{-L_t}} \quad (4)$$

其中

$$L_t = c + \beta_1 \text{eps} + \beta_2 \text{gro} + \beta_3 \text{res} + \beta_4 \text{scale} + \beta_5 \text{send} + \beta_6 \text{sendLY} + \beta_7 \text{sendLLY} \quad (5)$$



表 4 变量之间的相关性

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14
V1	1													
V2	0.113**	1												
V3	0.047**	0.004	1											
V4	0.071**	0.006	0.002	1										
V5	0.455**	0.027	0.015	0.023*	1									
V6	0.019	0.002	0.005	0.005	-0.094*	1								
V7	0.191**	0.017	-0.026	-0.008	0.412**	-0.056*	1							
V8	0.045**	0.002	0.004	0.008	-0.085*	0.778**	-0.054*	1						
V9	0.223**	0.006	0.011	0.013	0.410*	-0.043*	0.187**	-0.013	1					
V10	0.017	0.044**	0.008	0.040*	0.236**	-0.101*	0.072**	-0.109*	0.149**	1				
V11	0.093*	-0.009	-0.008	0.001	0.088**	0.117**	0.071**	0.114**	0.158**	-0.162*	1			
V12	0.003	0.005	0.005	-0.001	0.048**	-0.015	0.059**	-0.008	0.017	0.074**	-0.012	1		
V13	0.176**	0.033*	0.011	0.027*	0.370**	-0.063*	0.233**	-0.047*	0.231**	0.255**	0.047**	-0.018	1	
V14	0.018	0.019	0.026	0.018	0.046**	-0.030*	-0.143*	-0.036	-0.044*	0.193**	0.019	0.06	0.01	1

V1:每股收益; V2:当年业绩增速; V3:去年业绩增速; V4:前年业绩增速; V5:股价; V6:流通股; V7:公积金; V8:股本; V9:次新股; V10:公私属性; V11:控股权; V12:半年报送转; V13:去年送转情况; V14:前年送转情况

\*\*, 在 0.01 级别(双尾), 相关性显著。

\*, 在 0.05 级别(双尾), 相关性显著。

表5 初始Logistic模型回归结果

变量名称	系数	t统计量	p值
每股收益	8.97	4.83	<0.001***
当年业绩增速	3.47	0.43	0.67
去年业绩增速	3.38	0.73	0.46
前年业绩增速	4.71	0.85	0.40
股价	-2.22	-2.20	0.03 **
流通股	0.24	0.12	0.91
公积金	3.28	5.89	<0.001 ***
股本	-239.70	-2.83	0.005 ***
次新股	-0.29	-1.19	0.21
公私属性	0.02	0.15	0.88
控股权	0.24	0.62	0.54
半年报送转	-2.16	-2.21	0.03**
去年送转情况	13.88	30.41	<0.001 ***
前年送转情况	3.23	6.75	<0.001 ***
截距(intercep)	-11.52	-1.84	0.07*

注:\*\*\*表示在1%的显著性水平上显著,\*\*表示在5%的显著性水平上显著,\*表示在10%的显著性水平上显著。

## 2.BP神经网络模型

构造结构合理的人工神经网络仍然处于研究阶段,但已有研究表明(Cybenko, 1989; Hornik 等人, 1989),若构建的神经网络隐藏层使用 sigmoid 单元、输出层使用线性单元,则三层的神经网络结构具有逼近任意有界连续函数的能力,因此本文选择三层的神经网络进行建模。人工神经网络算法需要对隐藏层节点数、迭代次数、学习速度三个参数进行设置。我们赋予参数一系列具体数值,接着通过训练集建模然后基于验证集算出所建模型的预测准确率。多次重复,算出预测准确率的平均值。最后选取具有最高预测准确率平均值的参数作为最终选择。

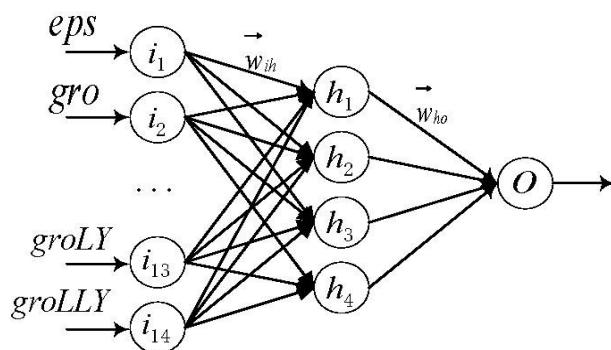


图2 本文采用的人工神经网络模型结构

基于人工神经网络的已有研究,隐藏层节点数会影响网络的整体结构,从而进一步影响拟合的精度,学习速度参数设置不合理会使得最终误差陷入局部最小值而非全局最小值,迭代次数设置不合理会使得模型陷入轻度拟合或过度拟合的状态。根据三个参数对模型有效性的影响程度,本文先确定隐藏层节点数来确定较为合理的网络结构,再确定最优的学习速度和迭代次数。下图为隐藏节点数在不同取值下1000次随机抽样预测的概率最大的前10%和20%股票的平均预测准确率。

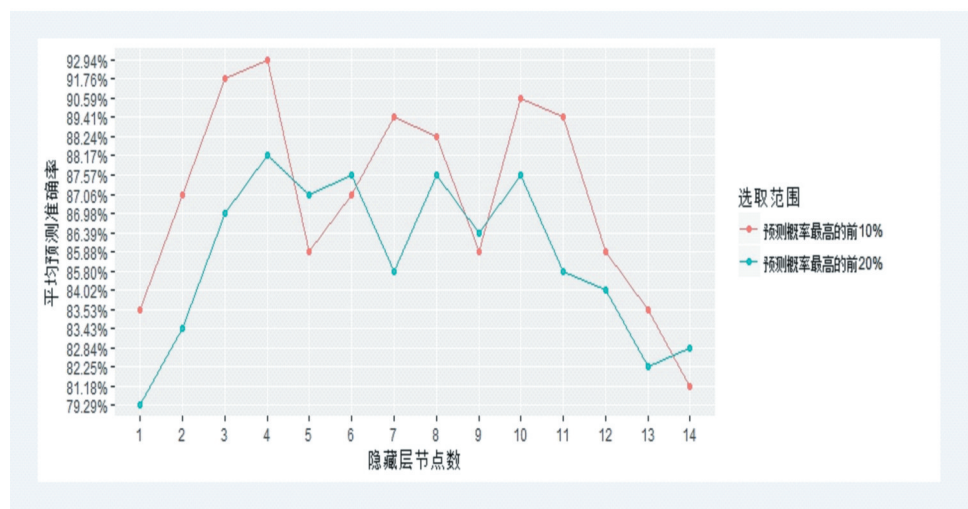


图3 不同隐藏层节点数的预测准确率

如图3所示,当隐藏层节点数设置为4时模型能达到较好的拟合状态,因此设置隐藏层节点为4。通过同样的方法,我们将模型的学习速率设置为0.0005,迭代次数设置为300次。

### (三)模型拟合与预测结果

实验采用训练集数据对两模型进行拟合,拟合完毕后用测试集数据进行有效性测试。由于模型的输出是高送转概率的拟合,我们对模型输出结果从大到小排序,然后选取模型输出最大的前10%和前20%的股票中高送转的概率作为模型的预测准确率。由于抽样的随机性会造成训练集数据改变,从而影响模型的拟合精度,我们通过进行多次抽样取平均预测准确率的形式来避免这一因素的影响。本次实验进行10000次的随机抽样,得到如下结果如表6所示。

表6 平均预测准确率

模型	预测结果	前10%最大输出	前20%最大输出
BP	正确	78	148
	错误	7	21
	总数	85	169
	准确率	91.8%	87.6%

Logistic	正确	66	128
	错误	19	41
	总数	85	169
	准确率	77.6%	75.7%

从上述模型的有效性验证结果来看,对于“高送转”的预测效果,BP神经网络模型优于Logistic模型。针对前10%的股票来说,BP神经网络模型预测出来的送转概率靠前的股票的预测准确率能够达到92%左右,比Logistic模型提高了大约14%。针对前20%的股票来说,也能达到88%左右,比Logistic模型提高了大约12%。因此通过BP神经网络模型,能够预测出潜在的“高送转”股票,建立“高送转”投资组合,获得较稳定的超额收益。

## 五、基于BP神经网络权值的送转动因分析

绝大多数送转动因的研究以及“高送转”预测都是基于Probit或Logistic等计量回归模型,使用该模型的优点在于根据回归结果能够很好的解释自变量对应变量的影响程度,缺点在于该模型是广义线性模型,对于本质是复杂的、非线性的实际问题不能达到较高的拟合精度。而人工神经网络模型是非线性模型,优越的拟合效果是其最大的优势,在“高送转”的预测上,实证结果已经表明该模型较Logistic模型有更好的预测效果。

表7 各层之间的权值分布表

输入变量	iàh	h1	h2	h3	h4	hào	o
每股收益	i1	-2.41	-1.13	-3.88	-2.13	h1	<u>-4.97</u>
当年业绩增速	i2	2.18	0.93	4.12	2.97	h2	<u>5.03</u>
去年业绩增速	i3	-1.86	-1.45	-0.62	1.74	h3	0.89
前年业绩增速	i4	-1.77	-1.8	0.97	3.11	h4	-0.95
股价	i5	-0.73	-2.26	8.74	-0.25		
流通股	i6	2.32	<u>6.50</u>	-1.96	2.85		
公积金	i7	<u>-12.72</u>	<u>-13.95</u>	13.58	0.79		
股本	i8	<u>9.51</u>	2.11	-5.62	4.31		
次新股	i9	1.3	1.37	-1.69	0.23		
公私属性	i10	0.31	0.74	-2.11	3.57		
控股权	i11	-0.41	-0.50	0.51	-1.15		
半年报送转	i12	1.33	1.53	-1.98	-3.52		
去年送转情况	i13	<u>10.11</u>	<u>9.77</u>	-2.52	-1.99		
前年送转情况	i14	-1.85	-1.99	2.51	0.84		



本部分内容,我们在拟合精度更高、预测效果更好的人工神经网络模型基础上,通过分析人工神经元节点间的权值,来进一步分析各个解释变量与送转概率之间的作用机制。

各层之间的权值包括:输入层(i)与隐藏层(h)之间的权值,隐藏层(h)与输出层(o)之间的权值参数,训练完的模型的权值具体情况如表7所示。

输入变量对输出结果的影响程度被称为输入变量的敏感性,我们将基于BP神经网络神经元节点之间的权值参数定性分析各变量对输出结果的敏感性,从模型角度来研究高送转动因。

图4展示了训练完的神经网络模型中对输出结果有决定性影响的权值情况,其中红色表示权值为正,蓝色表示权值为负。从隐藏层与输出层之间的权值来看,其中隐藏层节点 $h_1$ 与 $h_2$ 对输出结果具有较大的影响, $h_1$ 与输出结果呈负相关, $h_2$ 与输出结果呈正相关。因此这两个节点是企业是否会进行高送转的主要因素。我们对这两个节点的输入层权值进行分析:1) $i_8$ 、 $i_{13}$ 对 $h_1$ 会产生较大的正面影响, $i_7$ 对 $h_1$ 产生较大的负面影响,又因为输出结果与 $h_1$ 负相关,因此 $i_{13}$ 、 $i_8$ 会对输出结果产生较大负面影响,即股本规模越大,去年送转比例越高则企业进行送转的概率会越低; $i_7$ 与输出结果正相关,即每股公积金越高,则企业送转概率越高。2) $i_{13}$ 、 $i_6$ 对 $h_2$ 的有较大的正面影响, $i_7$ 对 $h_2$ 有较大的负面影响,又因为 $h_2$ 与输出结果正相关,因此 $i_{13}$ 、 $i_6$ 对输出结果产生正面影响, $i_7$ 与输出结果负相关。即去年送转比例越高则企业送转概率越大,流通股数量越多则送转的概率越大,每股公积金越多反而会使得送转概率下降。

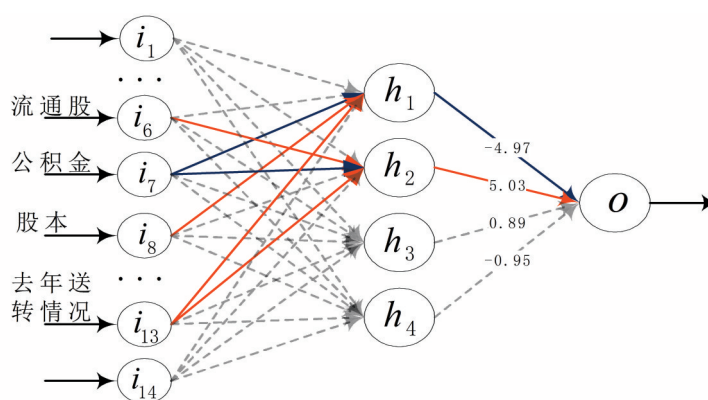


图4 神经网络二层权值结构图

结论1)与2)是两个看似十分矛盾的结论,但却能解释实际送转概率都相对较高的两类公司的送转动机:第一类是股本较小,每股公积金较高的,去年未进行送转的公司,这类公司具有强烈扩张股本的意图,符合这一类特点的公司以新上市的次新股为主。这类公司因为上市不久,受融资政策的限制,股本往往较小,每股公积金较高,具备送转动机,这一定程度上支持了“股本扩张论”。第二类公司通常具有送转经历,因此这类公司的公积金较低,流通股数量较大。这类公司的管理层具有较强的高送转动机,每当公积金、流通股数量达到送转的基本条件,管理层就会立即进行高送转。这类公司往往已经有一定的上市时间,管理层基于历史经验倾向于通过连续实施高送转达到做大公司市值的目的。这类送转行为一定程度上符合“股利迎合论”与“信号传递论”,公司管理层一方面为了迎合投资者对于“高送转”股票的非理性偏好从而实施高送转,并最终得到了市场积极回应,扩大了公司市值。同时,对于投资者来说,公司市值成功扩大又进一步证实了当初高送转所释放出的积极乐观信号。

## 六、结论与研究展望

高送转是中国股票市场一种较为特殊的现象。本文基于2014-2016年中国股票市场的数据,引入BP神经网络模型对高送转现象进行建模,并利用建立的模型对高送转进行预测。此外,本文还利用建立的模型对高送转的动因进行分析。结论如下:

(1)高送转方案是公司管理层基于多方面因素综合考虑后的决策结果,先前针对高送转模型的研究都是通过Logistic或Probit等线性计量模型进行。本文的实证结果表明,相比于上述计量模型,基于BP神经网络算法建立的模型能更好的拟合高送转现象。基于BP神经网络建立的模型对高送转的预测正确率比Logistic模型高14个百分点,能够达到92%左右的正确率。

(2)本文所建立的模型显示,高送转的公司可以分为两类。第一类公司的送转股本数量越小、每股公积金越高,且去年未进行送转则进行高送转的概率越大。这类公司反映了新上市的次新股公司,由于新上市不久,股本往往较小且未进行高送转。这类公司若公积金变高,则为了扩张股本,就会就行高送转。这个结论支持了“股本扩张论”。第二类公司的管理层为了做大公司市值,具有较强的连续送转的动机。只要公司公积金、流通股数量达到送转要求,管理层就会考虑实施高送转。因此这类公司往往是公积金较小,流通股数量较多,且在去年已经实施了高送转。本文根据建立的模型的启发,首先对市场中高送转现象进行了分类,然后对不同类别的高送转的动因分别进行解释,从而提出了一种与公司成长阶段相关联的高送转动因解释,实现了不同动因解释理论的统一。

(3)传统高送转动因解释理论分别从不同角度对高送转现象进行解释,由于高送转动因的复杂性,每个理论均不能很好的解释高送转现象。本文提供了一种研究框架,该框架基于数据挖掘技术对多种理论进行统筹研究,从而得出一个系统性的结论。当高送转的新动因解释理论出现时,该研究框架可将新理论纳入到原有理论体系中:构建一个解释变量,该变量能够充分反应新理论提出的送转动因;将此变量加入到现有的变量体系中,然后通过数据挖掘算法获得高准确率率的训练模型;通过分析模型权值,探究送转动因,从整体上对送转进行解释分析。

## 参考文献

- [1]陈浪南和姚正春,2000,《中国股利政策信号传递作用的实证研究》,《金融研究》第10期,69-77。
- [2]陈珠明和史余森,2010,《高送转股票财富效应的实证研究》,《系统工程》第10期,8-14。
- [3]龚慧云,2010,《基于股利迎合理论的我国上市公司送股、转股行为研究》,《上海金融》第9期,67-72。
- [4]高洪深和陶有德,1999,《基于人工神经网络的非线性回归预测模型的研究》,《北方工业大学学报》第3期,68-73。
- [5]黄娟娟和沈艺峰,2007,《上市公司的股利政策究竟迎合了谁的需要》,《会计研究》第8期,36-43。
- [6]何涛和陈小悦,2003,《中国上市公司送转股、转增行为动机初探》,《金融研究》第9期,244-256。
- [7]李心丹等,2014,《中国股票市场“高送转”现象研究》,《管理世界》第11期,133-145。
- [8]吕长江和王克敏,1999,《上市公司股利政策的实证分析》,《经济研究》第12期,31-39。
- [9]刘湘云、吴文洋、朱泓翰,2016,《基于小波神经网络的股指期货市场极端风险预警研究》,《投资研究》第12期,31-47。
- [10]刘笑天和赵胜民,2017,《基于神经网络的在线学习策略》,《投资研究》第6期,33-48。
- [11]沈海平,2011,《我国上市公司高送转公告效应的实证研究》,《区域金融研究》第4期,47-51。
- [12]魏刚,1998,《我国上市公司股利分配政策的实证研究》,《经济研究》第6期,30-36。
- [13]熊德华和刘力,2007,《股利支付决策与迎合理论——基于中国上市公司的实证研究》,《经济科学》第5期,89-99。

- [14]熊义明等, 2012,《中国上市公司送转行为动因研究——基于高送转样本的检验》,《经济与管理研究》第5期, 81-88。
- [15]王宜怀和王林, 2004,《基于人工神经网络的非线性回归》,《计算机工程与应用》,第12期, 79-82。
- [16]徐慧玲和吕硕夫, 2012,《中国上市公司“高送转”股利分配政策分析》,《经济问题》第11期, 84-88。
- [17]薛祖云和刘万丽, 2009,《中国上市公司送转股行为动因的实证研究》,《厦门大学学报》第5期, 114-121。
- [18]杨一文和杨朝军, 2005,《基于支持向量机的时间序列预测》,《系统工程理论方法应用》第4期, 176-181。
- [19]Asquith, P., P. Healy and K. Palepu, 1989, “Earnings and Stock Splits”, *The Accounting Review*, 64, pp. 387-403.
- [20]Baker, M. and Wurgler, J, 2004, “A Catering Theory of Dividends”, *Journal of Finance*, 59, pp. 1125-1165.
- [21]Baker, M., and Wurgler, J, 2004, “Appearing and Disappearing Dividends: The Link to Catering Incentives”, *Journal of Financial Economics*, 73, pp. 271-288.
- [22]Bhattacharya, S, 1979, “Imperfect Information, Dividend Policy, and the Bird in the Hand Fallacy”, *Bell Journal of Economics*, 10, pp. 259-270.
- [23]Black, F, 1976, “The Dividend Puzzle”, *The Journal of Portfolio Management* Winter, pp. 634-639.
- [24]Baker, Kent and Gary, 1993, “Further evidence on managerial motives for stock splits”, *Quarterly Journal of Business and Economics*, 32(3), pp. 20-31.
- [25]Babitha Rohit, Pinto Prakash and Shakila B. Drishtikon, 2016, “Impact of Stock Splits and Rights Issue Announcements on Market Price: Evidence from India”, *A Management Journal*, pp. 1-16.
- [26]Cybenko, G., 1989, “Approximation by Superpositions of a Sigmoidal Function”, *Mathematics of Control Signals and Systems*, 2, pp. 303-314.
- [27]Eng Li, Li Ha, Joohyung and Nabar Sandeep, 2014, “The impact of regulation FD on the information environment: evidence from the stock market response to stock split announcements”, *Review of Quantitative Finance and Accounting*; PP. 829-853.
- [28]Ezell J and Rubiales C, 1975, “An empirical analysis of the determinants of stock splits”, *Finance Review*, 20, pp. 21-30.
- [29]Grinblatt M, Masulis R and Titman S, 1984, “The Valuation Effects of Stock Splits and Stock Dividends”, *Journal of Financial Economics*, 13, pp. 461-490.
- [30]Hwang S, Keswani A and Shackleton M, 2005, “Stock Splits: what Does the Market Tell Us Ex Ante?”, *Working Paper*.
- [31]Hornik k., 1989, “Multilayer Feedforward Networks Are Universal Approximators”, *Neural Networks*, 2, pp. 359-366.
- [32]John K and Williams J, 1985, “Dividends, Dilution, and Taxes: A Signaling Equilibrium”, *Journal of Finance*, 40, pp. 1053-1070.
- [33]Krieger Kevin and Peterson David R., 2009, “Predicting stock splits with the help of firm-specific experiences”, *Journal of Economics and Finance*, pp. 410-421.
- [34]Miller M. and Rock K, 1985, “Dividend Policy under Asymmetric Information”, *Journal of Finance*, 40, pp. 1031-1051.
- [35]Modigliani F. and Miller M., 1961, “Dividend Policy, Growth and the Valuation of Shares”, *Journal of Business*, 34, pp. 411-433.
- [36]Spudeck Raymond E and Moyer Charles, 1986, “Reverse Splits and Shareholder Wealth: The Impact of Commissions: I”, *Introduction. Financial Management*, pp. 52-56.
- [37]Weld, W. C., R. Michaely, and R. H. Thaler and S. Benartzi, 2009, “The Nominal Price Puzzle”, *Journal of Economic Perspective*, 23, pp. 121-142.

**Abstract:** The high stock dividend is a very interesting issue in the Chinese stock market. Previous researchers mostly use Logistic or Probit method to model the high stock dividend behavior. This paper applies the artificial neural network algorithm to predict the high stock dividend decision, and the prediction accuracy reaches around 92%, which is highly higher than the Logistic method. Moreover, based on the analysis of the model, we find that the motives of high stock dividends relate to the development stage of the firm.

**Key Word:** High Stock Dividend; Neural Network; Dividend Policy