

# 数字经济时代的智能投资决策

## ——基于多模型融合的股价预测研究

尹琳

南京信息工程大学数学与统计学院, 江苏 南京

收稿日期: 2025年12月19日; 录用日期: 2025年12月29日; 发布日期: 2026年1月23日

### 摘要

随着数字经济的迅猛发展, 数字化转型已经成为经济市场不可逆转的趋势, 数据是关键生产要素, 算法是使数据产生价值的核心生产力, 股票价格能够反映不同国际贸易情形下的不同市场表现。本文选取贵州茅台的股票价格数据以及公司财务报表中的17项关键特征数据, 包含了公司的盈利情况、偿债能力、运营效率和成长效率等多个方面, 全面反映公司的经营状况和未来发展潜力。利用LASSO回归模型对财务报表数据进行筛选和降维, 保留最具代表性的特征。构造基于LASSO和BP神经网络的符合模型, 结果表明, 该复合模型的均方根误差均低于单独的LASSO模型或者单独的BP神经网络模型。证明了在数字经济背景下, 基于LASSO和BP的宽深神经网络模型在数据预测方面的优越性, 以及将传统财务分析方法和现代机器学习技术相结合的有效性。

### 关键词

数字经济, LASSO回归模型, BP神经网络, 股票价格预测

# Intelligent Investment Decision-Making in the Digital Economy Era

## —Research on Stock Price Prediction Based on Multi-Model Fusion

Lin Yin

School of Mathematics and Statistics, Nanjing University of Information Science and Technology, Nanjing Jiangsu

Received: December 19, 2025; accepted: December 29, 2025; published: January 23, 2026

### Abstract

With the rapid development of the digital economy, digital transformation has become an irreversible

文章引用: 尹琳. 数字经济时代的智能投资决策[J]. 电子商务评论, 2026, 15(1): 69-80.

DOI: 10.12677/ecl.2026.151010

trend in the economic market. Data is a key production factor, and algorithms are the core productivity that enables data to generate value. Stock prices can reflect the different market performances under various international trade scenarios. This paper selects the stock price data of Kweichow Moutai and 17 key characteristic data from the company's financial statements, covering multiple aspects such as the company's profitability, debt-paying ability, operational efficiency, and growth efficiency, comprehensively reflecting the company's operating conditions and future development potential. The LASSO regression model is used to screen and reduce the dimension of the financial statement data, retaining the most representative features. A composite model based on LASSO and BP neural network is constructed. The results show that the root mean square error of this composite model is lower than that of the LASSO model or the BP neural network model alone. This proves the superiority of the wide and deep neural network model based on LASSO and BP in data prediction under the background of the digital economy, as well as the effectiveness of combining traditional financial analysis methods with modern machine learning techniques.

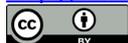
## Keywords

Digital Economy, LASSO Regression Model, BP Neural Network, Stock Price Prediction

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

国际和国内重视数字经济的发展,全球数字化转型的进程不断加速。在国家经济发展的宏大图景中,股票市场一直扮演着不可或缺的角色,它既是经济的晴雨表,又在一定程度上指引着国家经济前行的方向;股价反映企业特质信息,是衡量资本市场资源配置功能的核心指标。然而,近期我国股市的表现逐渐暴露出一些问题,由于股民缺乏科学系统的投资理论指导,投资者在行为上可能有冲动盲目的倾向性,股票市场中的价格波动愈发偏离其应有的合理区间。在这种情况下,为促进市场平稳运行,引导理性投资,需要构建科学系统的分析框架,将统计学的方法应用于股价分析中,对影响股票价格的因素进行建模预测,分析各方面因素对股票价格的作用,从而推导出数字经济时代的智能投资决策。

## 2. 文献综述

数字经济时代,股票市场的波动性和不确定性确实使得准确预测股票价格成为金融领域的热点研究问题,张虎等[1]发现数字化转型显著提升了制造业上市企业的股票流动性,也为制造业企业及政策制定者如何利用数字化转型提升资本市场表现提供了实践指导。陈奋强等[2]发现数字金融显著降低了股价同步性,同时有助于降低分析师预测偏差与预测分歧度。数字经济已成为驱动经济发展的新引擎,如何有效预测做出智能决策成为了一个问题。

在处理复杂的非线性关系和模式时,基于线性的传统模型的股票预测方法存在一定局限性。近年来,随着人工智能技术的不断发展,出现了机器学习和深度学习方向的各种新方法,2016年李红梅[3]对比了机器学习方法和传统统计建模方法的预测效果,发现对于多元时间序列数据而言,传统计量模型的效果不如神经网络模型,证明了神经网络模型预测的有效性。然而单一模型存在容易陷入局部最优的缺陷,刘佳祺、刘德红等[4]将遗传算法、PCA主成分分析法与BP神经网络算法相结合,构建了PCA-GA-BP整合模型预测股价变动情况,有效改善了BP神经网络运算速度慢和容易陷入局部极小值的缺点。周宁等[5]利用BP神经网络对中石化股票数据进行建模预测后,结合马尔可夫模型对误差序列进行处理,有效

提高了模型的准确度，但存在计算复杂的问题。苏兆辉等[6]基于改进经验模态分解与 A-LSTM 混合神经网络的方法预测股票价格具有优良的准确性和稳定性，但没有考虑各类因素的综合影响。然而过去的文献大多聚焦于单一的神经网络模型或者对基础模型的特定缺陷进行优化，没有考虑系统性融合多种具有互补优势的模型，本文聚焦 LASSO 回归和神经网络方法的融合，提升了预测的精确性。

### 3. 方法介绍与数据来源

#### 3.1. LASSO 回归方法介绍

当数据集的维度非常高，即协变量数量过多时，这些变量之间可能会出现显著的信息重合，这种现象被称为多重共线性，多重共线性会导致普通多元线性回归模型的估计参数不稳定，容易过高或过低地估计真实的参数值。为了解决这一问题，LASSO (Least Absolute Shrinkage and Selection Operator) 回归方法被广泛应用。LASSO 回归是一种线性回归模型，它通过向损失函数中添加 L1 正则化项来改进普通最小二乘估计。L1 正则化项是参数向量的绝对值之和，它会将模型的系数压缩到较小的值，以及将一些不重要的系数压缩至零，从而实现重要特征选择。这种方法不仅可以帮助我们避免过拟合，还可以简化模型，能够有效地处理协变量数量远大于样本数量的情况。在 LASSO 回归中，正则化参数  $\lambda$  用于控制正则化的强度，通过调整正则化参数  $\lambda$ ，LASSO 回归可以在模型拟合和特征选择之间找到平衡。在模型训练过程中，我们可以使用交叉验证等技术来选择最优的  $\lambda$  值，从而确保模型在训练集和测试集上都能表现出良好的性能。为了使用正则方法，需要引入惩罚函数：

$$\hat{\beta}_{LASSO} = \arg \min_{\beta} \left( \|Y - X\beta\|^2 + P_{\lambda}(|\beta|) \right)$$

其中， $P_{\lambda}(|\beta|) = \lambda \sum_{i=1}^p |\beta_j|$ ,  $n \geq 0$  为惩罚项， $\lambda$  为调节参数，当  $n=1$  时，该惩罚项为 LASSO 惩罚，回归系数估计值  $\hat{\beta}_{LASSO}$  要使下式达到最小的  $\beta_j$ ：

$$\hat{\beta}_{LASSO} = \arg \min_{\beta} \left( \|Y - X\beta\|^2 + \lambda \sum_{i=1}^p |\beta_j| \right)$$

当  $\lambda$  越小， $\lambda$  越大，引发系数趋向 0 的压缩程度越大，即压缩能力也越大，通过 LASSO 惩罚函数将不显著的解释变量的系数压缩至 0，从而剔除不显著变量。在股票预测中，特征选择是一个重要步骤。由于股票数据通常可能与大量的指标相关，同时其中许多变量可能与预测目标无关或相关性较低。LASSO 回归通过引入正则化项，能够在模型训练过程中自动选择对预测目标有显著影响的特征，从而有效避免模型过拟合，提高预测精度。

#### 3.2. BP 神经网络模型

20 世纪末 Rumelhart 等人提出名为反向传播的算法，即为 BP (Back Propagation) 神经网络，它是一种采用误差逆向传播算法进行训练的多层前馈网络，根据梯度下降算法进行权重和阈值的更新。该神经网络的信号是正向传播，而误差是反向传播的，BP 神经网络在训练过程中通过不断的前向传播和反向传播迭代优化网络参数，最终实现对输入数据的准确预测。首先通过前向传播计算出对于给定输入的预测输出，然后与期望输出进行对比，根据误差来反向调整网络中的权重，这个过程反复进行，直至网络能够在一定程度上准确地预测和分类数据。BP 神经网络包括输入层、输出层和隐藏层三个层对数据进行复杂的变换和整合，输出层将经过处理的数据转化为最终的预测结果或分类标签，层内内部神经元相互独立，各层之间相互通过每个层级的首尾处边端的神经元连通。我们需要选择非线性函数作为 BP 神经网络的激活函数，考虑数据取值范围和自变量数据类型之后，本文选择 Sigmoid 型函数，值域为 0 到 1，函数表达式为：

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

BP 神经网络[7]的工作过程分为两个阶段：前向传播和反向传播。在前向传播阶段，输入信号通过输入层传入网络，经过隐藏层的处理，最终到达输出层并产生输出结果。如果输出结果与期望输出之间存在误差，则进入反向传播阶段。在反向传播阶段，根据误差信号调整网络中的权重和偏差项，使网络的输出误差逐渐减小，图 1 为一个简单的三层 BP 神经网络结构图。

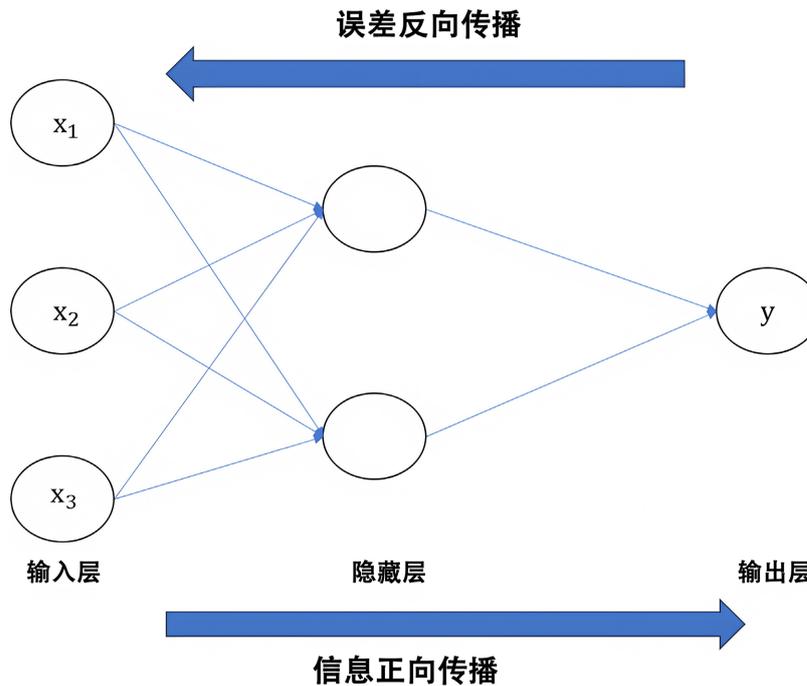


Figure 1. Three-layer BP neural network model structure diagram

图 1. 三层 BP 神经网络模型结构图

由于该神经网络的训练过程分为两部分，分别是数据前向传递和误差反向传递，据此给出它的训练步骤。设 BP 网络的输入数据为  $x = (x_1, x_2, x_3)$ ，输出为  $y$ ，输入层节、隐藏层以及输出层节点个数依次为 3、2 与 1，第一层权重矩阵为  $w^{(1)}$ ，其中元素  $w_{ij}^{(1)}$  表示第 1 层中第  $i$  节点对第二层中第  $j$  个节点的权值，偏置为  $b_1$ ；第二层权重矩阵为  $w^{(2)}$ ，其中元素  $w_{jk}^{(2)}$  与  $w_{ij}^{(1)}$  类似，偏置为  $b_2$ 。设此网络的激活函数为 Sigmoid，隐层输入值为  $a_1$ ，输出值为  $a_2$ 。

(1) 信号向前传播过程：

$$inp_{a_1} = x_1 w_{11}^{(1)} + x_2 w_{21}^{(1)} + x_3 w_{31}^{(1)} + b_1$$

$$outp_{a_1} = f(inp_{a_1}) = \frac{1}{1 + e^{-inp_{a_1}}}$$

$$inp_{a_2} = x_1 w_{12}^{(1)} + x_2 w_{22}^{(1)} + x_3 w_{32}^{(1)} + b_1$$

$$outp_{a_2} = f(inp_{a_2}) = \frac{1}{1 + e^{-inp_{a_2}}}$$

$$inp_{\hat{y}} = outp_{a_1} w_{11}^{(2)} + outp_{a_2} w_{21}^{(2)} + b_2$$

$$\hat{y} = f(\text{inp}_{\hat{y}}) = \frac{1}{1 + e^{-\text{inp}_{\hat{y}}}}$$

损失函数如下:

$$L_{(1)} = (y - \hat{y})^2$$

(2) 误差反向传播,  $\alpha$  为学习率, 使用下述更新方法更新权值  $w_{11}^{(2)}, w_{21}^{(2)}$ :

$$\Delta w_{11}^{(2)} = \frac{\partial L_{(1)}}{\partial w_{11}^{(2)}} = \frac{\partial L_{(1)}}{\partial \hat{y}} \times \frac{\partial \hat{y}}{\partial \text{inp}_{\hat{y}}} \times \frac{\partial \text{inp}_{\hat{y}}}{\partial w_{11}^{(2)}}$$

$$w_{21}^{(2)} = w_{11}^{(2)} - \alpha \Delta w_{11}^{(2)}$$

接下来, 按照下述更新方法更新权值  $w_{11}^{(1)}, w_{12}^{(1)}, w_{21}^{(1)}, w_{22}^{(1)}$ 。

$$\Delta w_{11}^{(1)} = \frac{\partial L_{(1)}}{\partial w_{11}^{(1)}} = \frac{\partial L_{(1)}}{\partial \hat{y}} \times \frac{\partial \hat{y}}{\partial \text{inp}_{\hat{y}}} \times \frac{\partial \text{inp}_{\hat{y}}}{\partial \text{out}_{a_1}} \times \frac{\partial \text{out}_{a_1}}{\partial \text{inp}_{a_1}} \times \frac{\partial \text{inp}_{a_1}}{\partial w_{11}^{(1)}}$$

$$w_{12}^{(1)} = w_{11}^{(1)} - \alpha \Delta w_{11}^{(1)}$$

由于 BP 神经网络强大的非线性映射能力和自学习能力, 它在股票预测领域得到了广泛应用[8]。通过对大量历史数据的学习, BP 神经网络能够捕捉股票价格变化的复杂规律和趋势, 进行准确的预测, 从而为投资者提供全面的决策支持。

## 4. 实例分析

### 4.1. 数据来源

为了对贵州茅台(600519)的股票价格进行预测分析, 本研究选取了从 2010 年 3 月 31 日至 2023 年 12 月 31 日这一时间段内, 公司各季度的财务报表数据作为研究基础, 共计 56 个报告期的数据。在数据收集过程中, 本文特别关注了能够反映公司财务状况和经营成果的主要财务指标, 选取了营业收入(万元)、营业成本(万元)、营业利润(万元)、利润总额(万元)、所得税费用(万元)、净利润(万元)和每股收益(元)这七项指标, 它们共同构成了分析公司盈利能力的重要框架。除了财务指标外, 同时还从资产负债表中提取了关键信息, 货币资金(万元)、存货(万元)、流动资产合计(万元)、固定资产净额(万元)、资产总计(万元)、流动负债合计(万元)、负债合计(万元)以及所有者权益(或股东权益)合计(万元)这八项特征, 为我们提供了公司资产结构、负债水平以及所有者权益状况的相关数据。此外, 特别选取了期初现金及现金等价物余额(万元)和期末现金及现金等价物余额(万元)这两项特征, 以便更全面地了解公司的现金流情况。综合以上 17 项主要特征, 我们构建了一个用于预测贵州茅台股票价格的截面数据特征集, 这些特征涵盖了公司的盈利能力、资产结构、负债水平、现金流情况等多个方面, 为股票价格预测提供了全面的数据支持。在股票价格数据方面, 选择了每个季度最后一个开盘日的收盘价格作为对应的股票价格。本文的股票价格来源于东方财富网, 季度财务报表数据来自新浪财经网。

如图 2 所示, 从 2011 年到 2023 年, 贵州茅台股票价格整体呈现指数型上升趋势, 但中间也存有波动。在 2011 年到 2014 年间, 欧债危机发酵、美联储流动性收紧引发全球风险偏好下行, 茅台业绩增长乏力且估值受压制, 股价相对较低, 大部分时间在 100 元以下。从 2015 年开始, 全球宽松周期重启, 数字经济时代开始成型, 股价逐渐上涨, 到 2021 年中期达到一个高点, 超过了 2500 元, 但此时股票价格被股民盲目的逐利心理操控, 具有很强的风险性和不确定性。2021 年下半年到 2023 年上半年, 股价出现了波动和下跌, 这是由于疫情期间, 人们的心态变得保守, 股票价格逐渐归于平缓。总体来说, 这只股

票在过去几年里表现出了一定的增长趋势，尽管中间有几次波动和回调，这反映了一只股票的稳定增长或周期性波动的典型行为。



**Figure 2.** Closing line chart of Guizhou Moutai stock  
**图 2.** 贵州茅台股票收盘折线图

## 4.2. LASSO 回归实例分析

本文首先利用 LASSO 回归模型对主要财务指标中选择出的 17 项特征进行筛选，挑选出对股票价格影响较大的特征用于后续的模型建立中。由于原始数据中各项自变量特征数值的平均值之间存在较大的差值，会使 LASSO 回归模型的结果造成较大偏差，严重影响回归结果的准确性，同时将会导致原本具有重要影响的特征在模型中的系数变得异常小。为了消除不同量纲和尺度对 LASSO 回归模型结果造成的影响[9]，应在建立模型前对各项自变量特征分别进行归一化处理。经过多次实验建模对比归一化处理后的回归结果和未归一化处理后的回归结果，证明归一化处理后的 LASSO 回归模型在该问题中具有更小的均方根误差，说明归一化处理确实有利于提高预测精度，具体的归一化方法如下：

$$\text{归一化后的数据} = \frac{\text{原特征数据} - \text{特征最小值}}{\text{特征最大值} - \text{特征最小值}}$$

通过 4.1 节的分析可知股票价格随着时间的推移呈现出指数型增长的趋势，而其他特征是随时间增加呈现出线性增长的趋势，因此应对股票价格归一化之后取对数后再进行回归。设股票价格为  $p$ ，其他各项影响因素分别为  $x_1, x_2, \dots, x_{17}$ ，则将  $p$  对数化后的 LASSO 回归模型如下所示：

$$\ln p = a_1 x_1 + a_2 x_2 + \dots + a_{17} x_{17} + \varepsilon$$

由 LASSO 回归的方法原理可知 LASSO 回归模型中存在一个超参数  $\lambda$ ，当  $\lambda$  的取值越大则 L2 范数的惩罚项对自变量回归系数的压缩作用越强，这会导致自变量回归系数被压缩为 0。使用 Python 中的交叉验证法进行求解，交叉验证法将数据集分成  $k$  个大小相等的互斥子集。每次  $k-1$  个子集的数据训练模型，剩下的 1 个子集用于测试。这个过程重复进行  $k$  次，每次选择不同的子集作为测试集，最后计算  $k$  次测试结果得出最优值作为模型的性能指标。本题最佳的  $\lambda$  值为 0.001047，将该值带入 LASSO 回归模型进行建模。

按照训练集占全集 80% 和测试集占全集 20% 的标准划分, 由表 1 结果可知返回的是 LASSO 回归模型的回归系数, 其中 0 表示经过 LASSO 回归后被压缩为 0 的指标, 正值表示正相关, 负值表示负相关, 可以看出十七个指标中有十个被压缩至 0, 由此说明这十个特征对股票季度价格没有显著影响, 在对价格有显著影响的几个特征中, 对价格有正向影响的特征有: 基本每股收益、货币成本(万元)、流动负债合计(万元), 流动资产合计(万元)、期初现金及现金等价物余额(万元)、期末现金及现金等价物余额(万元), 共计 7 项特征, 对价格有反向影响的特征有: 营业收入(万元), 共计一项特征。以上结果也反映了上述指标对股票价格变动的影响, 当基本每股收益、存货资金、固定资产和现金及现金等价物越多时, 说明该股票具有比较高的投资价值, 会造成股票价格的上涨。同时, 也有一些指标与现有的统计学经济学原理相违背, 按照传统观念, 当营业收入增加时, 销售额增加, 股票价格也应该随之上涨, 但结果却是负相关的关系, 反而流动负债资产却是正相关的关系, 这可能是由于公司正在进行扩张从而投入了大量资金导致营业收入的减少或者流动负债的增加, 这些现象也说明股票市场的预测存在难度, 我们需要使用顺应时代潮流的机器学习或深度学习方法进行挖掘。接下来使用测试集对训练后的 LASSO 回归模型进行预测性能测试, 使用测试集数据的预测值和实际值计算出评估模型性能的指标均方根误差 RMSE, 得到输出的结果 0.07387 左右, 其结果表明 LASSO 回归模型对股票价格预测的均方根误差值大概在 0.07387 左右, 因此还应该用其他方法建立更优模型从而得到更精准的预测结果。

**Table 1.** LASSO regression coefficient

**表 1.** LASSO 回归系数表

	LASSO 回归的各项特征	LASSO 回归系数
基本财务数据	营业收入(万元)	-0.228394
	营业成本(万元)	-0
	营业利润(万元)	0
	利润总额(万元)	0
	所得税费用(万元)	-0
	净利润(万元)	0
	基本每股收益(元)	0.111193
资产负债数据	存货资金(万元)	0.028663
	存货(万元)	0
	流动资产合计(万元)	0.069242
	固定资产净额(万元)	0.257351
	资产总计(万元)	0
	流动负债合计(万元)	0.05035
	负债合计(万元)	0
现金流数据	所有者权益(或股东权益)合计(万元)	0
	期初现金及现金等价物余额(万元)	0.142523
	期末现金及现金等价物余额(万元)	0.300973
	截距	-0.228394

### 4.3. BP 神经网络实例分析

本文收集了 2010 年 3 月 31 日到 2023 年 12 月 31 日共 56 个季度的股票价格数据, 并将每个季度的

前四个季度依次排列作为一组时间序列数据，如 2021 年第一季度作为目标值，并以 2020 年的四个季度作为特征，共生成 52 组数据。将这四个季度依次命名为  $j_1, j_2, j_3, j_4$ ，分别表示当期数据前的第  $i$  个季度数据。

在归一化后的数据集取出  $j_1, j_2, j_3, j_4$  共四列，以这几列数据作为神经网络的四个输入值，并像前文将数据集划分为训练集与测试集。我们利用 Keras 模块建立一个三层的 BP 神经网络模型，由于是选取四个季度的收盘价作为输入值，因此拥有四个神经节点，本文使用单层隐藏层的神经网络模型，经过两千次迭代之后，BP 神经网络的损失函数均方差收敛于 0.0067，本方法中损失函数减少速度很慢，使用训练后的 BP 神经网络模型对测试集数据进行预测，计算均方根误差 RMSE 如图 3，选择出最佳节点数，可得实验结果如下。通常情况下，经过训练后的神经网络模型在测试集的预测误差会比训练集的拟合误差更大一些，如果训练后模型的拟合表现远好于测试集的拟合表现，这可能是由于 BP 神经网络模型训练中出现了过拟合的情况。由表 2 可知，当建立的神经网络模型中隐藏层节点数为 8 个时，预测集和训练集的预测误差相对最小。

### 隐藏节点数的各均方根误差

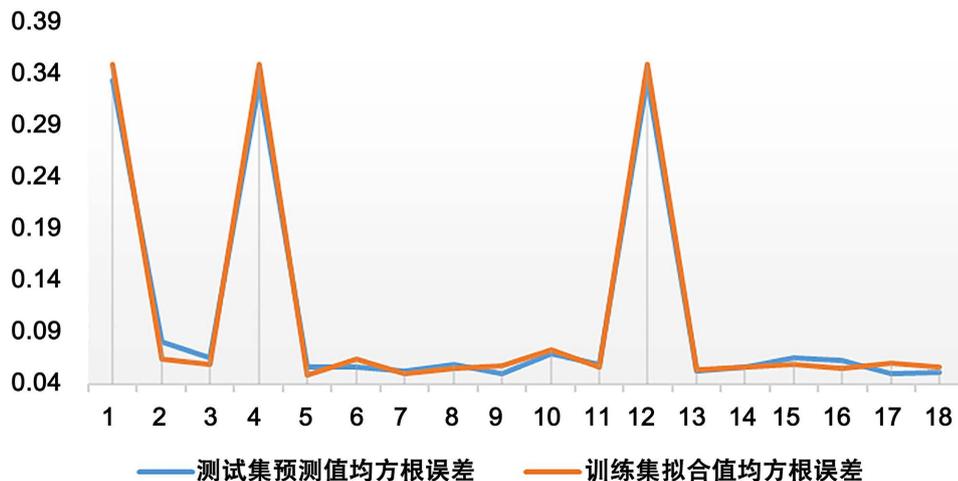


Figure 3. The root mean square error of the predicted values and the fitted values of each node in the hidden layer  
图 3. 隐藏层各节点预测值和拟合值的均方根误差

Table 2. The prediction errors and fitting errors under each hidden node  
表 2. 各隐藏节点下的预测误差和拟合误差

隐藏层节点数	测试集预测值均方根误差	训练集拟合值均方根误差
2	0.333545	0.348476
3	0.080690	0.064605
4	0.064982	0.058642
5	0.331311	0.348484
6	0.056953	0.049446
7	0.056943	0.064514
8	0.052096	0.050173
9	0.058961	0.055720

续表

10	0.050190	0.057205
11	0.069946	0.073424
12	0.058888	0.056854
13	0.334386	0.348504
14	0.052552	0.053616
15	0.056113	0.056848
16	0.064962	0.058722
17	0.063258	0.055273
18	0.050748	0.060627
19	0.051105	0.056882

#### 4.4. 基于 LASSO 和 BP 神经网络的宽深神经网络

##### 4.4.1. 建立宽深神经网络模型

利用 LASSO 结果对十七项特征筛选结果, 筛选出基本每股收益、货币成本(万元)、流动负债合计(万元), 流动资产合计(万元)、期初现金及现金等价物余额(万元)、期末现金及现金等价物余额(万元)和营业收入(万元)共计八项指标, 将这 8 项特征作为 BP 神经网络的特征输入神经网络后再与上文中前 4 期的时间序列数据在 BP 神经网络中的计算结果合并成新网络并加以训练, 称为基于 LASSO 和 BP 神经网络的宽深神经网络模型, 结构如图 4。

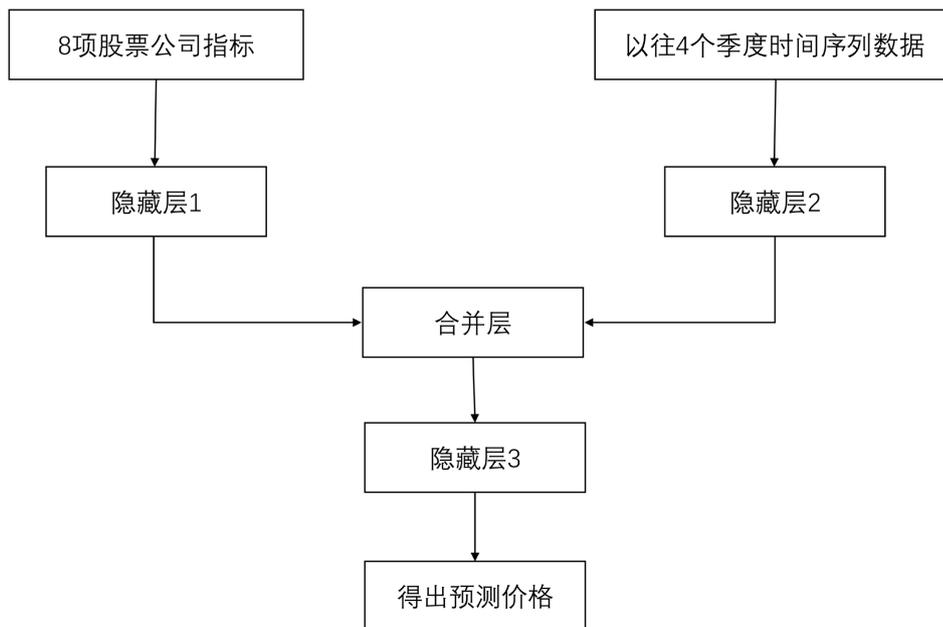


Figure 4. Wide and deep neural network architecture

图 4. 宽深神经网络结构图

根据上文第二节对以往 4 个季度时间序列的 BP 神经网络模型的测试, 隐藏层使用上文中得到的最优神经元个数, 因此在隐藏层 2 中设置 8 个隐藏神经元, 我们假定所建立的宽深神经网络模型中隐藏层

1 含 8 个隐藏神经元，隐藏层 3 中含 4 个隐藏神经元，我们利用前文已经进行归一化处理的数据进行模拟，同样按照上文中的方法划分测试集。首先建立输入层 A，用来接收经过 LASSO 回归筛选出来的表示公司盈利情况、经营情况和现金流的 8 项指标，设置为 8 个神经节点，建立输入层 B 接收公司股票价格的序列数据，设置为 4 个神经节点。建立隐藏层 1 处理 LASSO 回归筛选的八项指标的数据，建立隐藏层 2 对 4 期的时间序列数据进行处理，负区间的斜率设置为 0.2。将隐藏层 1 和隐藏层 2 合并，对该宽深神经网络迭代两千次之后输出预测结果。

从最后几次的迭代结果来看，该神经网络的损失函数收敛于 0.0013，迭代的时间和 BP 神经网络模型差不多，接下来计算该宽深神经网络的预测误差，测试集上的均方根误差为 0.0691，训练集上的均方根误差为 0.0618，好于 LASSO 模型的误差结果，但是没有 BP 神经网络的均方根误差小，我们需要对节点数进行调整。

#### 4.4.2. 寻找最佳节点数

通过上文假定隐藏层 2 的节点数为 8 个，利用网格搜索找到另外两个隐藏层的最佳节点，并进行最优测试，隐藏层神经元个数公式如下：

$$x = \sqrt{m+n} + y$$

$x$  表示隐藏层的节点数， $m$  表示输入的特征数， $n$  表示输出的特征数， $y$  是常数，由此可知神经元的个数大概为 [4, 14] 之间的整数数值，因此对神经网络进行 11\*11 共 121 网格搜索，为了简化计算并依据上一节的迭代结果，我们每次循环进行 200 次的迭代训练，并记录每一次训练结果的预测误差 MSE，得出输出结果如表 3，我们可以知道当节点数为 (9, 5)，即本节建立的宽深神经网络模型中隐藏层 1 的节点数为 9 个且隐藏层 3 的节点数为 5 个时，对测试集的误差最小，为 0.001419，远远小于别的模型的误差结果，说明基于 LASSO 和 BP 神经网络的宽深神经网络的预测结果更好。

**Table 3.** The error results of each node after grid search

**表 3.** 网格搜索后的各节点误差结果

节点数	MSE	节点数	MSE	节点数	MSE	节点数	MSE
(4, 4)	0.08073	(6, 12)	0.00812	(9, 9)	0.00409	(12, 6)	0.00445
(4, 5)	0.00545	(6, 13)	0.00848	(9, 10)	0.00667	(12, 7)	0.00692
(4, 6)	0.00255	(6, 14)	0.00379	(9, 11)	0.00345	(12, 8)	0.00758
(4, 7)	0.00499	(7, 4)	0.10068	(9, 12)	0.00535	(12, 9)	0.00358
(4, 8)	0.00305	(7, 5)	0.00912	(9, 13)	0.00167	(12, 10)	0.00459
(4, 9)	0.00277	(7, 6)	0.0113	(9, 14)	0.00506	(12, 11)	0.00601
(4, 10)	0.00476	(7, 7)	0.00596	(10, 4)	0.00308	(12, 12)	0.00521
(4, 11)	0.0069	(7, 8)	0.00547	(10, 5)	0.0989	(12, 13)	0.00323
(4, 12)	0.00694	(7, 9)	0.00442	(10, 6)	0.06768	(12, 14)	0.00291
(4, 13)	0.00521	(7, 10)	0.00529	(10, 7)	0.00379	(13, 4)	0.01171
(4, 14)	0.00302	(7, 11)	0.00773	(10, 8)	0.00327	(13, 5)	0.00807
(5, 4)	0.0051	(7, 12)	0.00788	(10, 9)	0.00388	(13, 6)	0.00304
(5, 5)	0.00315	(7, 13)	0.00635	(10, 10)	0.00614	(13, 7)	0.00389
(5, 6)	0.07357	(7, 14)	0.00545	(10, 11)	0.00329	(13, 8)	0.00703
(5, 7)	0.00368	(8, 4)	0.00563	(10, 12)	0.01567	(13, 9)	0.00642

续表

(5, 8)	0.00534	(8, 5)	0.0118	(10, 13)	0.0029	(13, 10)	0.00447
(5, 9)	0.00431	(8, 6)	0.00412	(10, 14)	0.00434	(13, 11)	0.00634
(5, 10)	0.00625	(8, 7)	0.02305	(11, 4)	0.0025	(13, 12)	0.00689
(5, 11)	0.00592	(8, 8)	0.00344	(11, 5)	0.00795	(13, 13)	0.00852
(5, 12)	0.0062	(8, 9)	0.00519	(11, 6)	0.00388	(13, 14)	0.00414
(5, 13)	0.00297	(8, 10)	0.00481	(11, 7)	0.00482	(14, 4)	0.00373
(5, 14)	0.00272	(8, 11)	0.00668	(11, 8)	0.00802	(14, 5)	0.03531
(6, 4)	0.00669	(8, 12)	0.00294	(11, 9)	0.00381	(14, 6)	0.00387
(6, 5)	0.01146	(8, 13)	0.00258	(11, 10)	0.01984	(14, 7)	0.00482
(6, 6)	0.00728	(8, 14)	0.00271	(11, 11)	0.0022	(14, 8)	0.0043
(6, 7)	0.00331	(9, 4)	0.00578	(11, 12)	0.00404	(14, 9)	0.00369
(6, 8)	0.00727	(9, 5)	0.00142	(11, 13)	0.02078	(14, 10)	0.00659
(6, 9)	0.00394	(9, 6)	0.00824	(11, 14)	0.0035	(14, 11)	0.00319
(6, 10)	0.00341	(9, 7)	0.00323	(12, 4)	0.09813	(14, 12)	0.00617
(6, 11)	0.00551	(9, 8)	0.00428	(12, 5)	0.00435	(14, 13)	0.00345
						(14, 14)	0.00425

宽神经网络模型在处理股票价格预测时，首先会对影响股票价格的关键数据进行细致的分类，去除影响较小的数据变量，再分别针对截面数据和时间序列数据两类数据进行迭代处理，确保它们各自的特性得到充分的捕捉，宽神经网络具有超越简单 BP 神经网络的优越性。接下来，模型会将截面数据和时间序列数据的训练输出合并，进一步通过迭代训练来优化预测效果，该处理方式旨在整合两种不同类型数据中的信息，以更全面地反映股票价格的变化规律，从而得出更精确的预测结果和投资方案。从理论上讲，模型仍有进一步优化的空间，可以通过引入更为精细的算法或策略来更好地捕捉不同类型数据中的不同信息[10]。

## 5. 总结与建议

经过对股票价格的深入分析，我们发现股票价格受到一系列主要财务指标的影响，经过 LASSO 回归模型的压缩和筛选，我们识别出对股票价格有显著影响的几个关键特征。其中，基本每股收益、货币成本(万元)、流动负债合计(万元)、流动资产合计(万元)、期初现金及现金等价物余额(万元)和期末现金及现金等价物余额(万元)这七项特征对股票价格具有正向影响，即这些指标的提升通常伴随着股票价格的上涨。而营业收入(万元)这一特征则对股票价格有反向影响，即营业收入的增加在某种程度上可能导致股票价格的下降。基于这些筛选出的特征，我们直接建立了 LASSO 回归模型，并在测试集上进行了预测。结果显示，该模型在测试集上的预测均方根误差为 0.07387，这表明我们的模型在预测股票价格方面具有较好的性能。

在使用 BP 神经网络模型对价格的时间序列数据进行建模预测时，经过不断测试，得出的最优隐藏层神经节点个数为 8 个，此时误差为 0.05 左右，以此建立最优模型，可见用 BP 神经网络模型进行预测效果优于 LASSO 回归模型的预测效果。将 LASSO 回归模型和 BP 神经网络结合后建立宽神经网络模型具有更强的预测效果，网格搜索后获得该模型的最佳参数，使用最佳参数训练后模型对测试集预测结果的均方根误差下降至 0.00142，这远远小于单独使用其他方式的预测误差，这足以说明结合后的宽深神经

网络模型对于股票数据的预测可以得到更好的结果，认为该多模型融合的股票预测方法具有较强的应用价值，为当下数字经济时代发展带来新的思考与策略。本文也存在一些不足，选择我国单只股票作为分析对象具有一定的局限性，同时股价波动也会受到国家政策的影响，但这一部分难以纳入考虑的范围[11]。综上所述，本文验证了多模型融合技术在股价预测上的显著优越性，更揭示了在数据成为关键生产要素的今天，谁能更有效地处理数据，采用精确度高的预测方法，谁就能在复杂市场中建立更有效的认知优势与决策优势。这不仅是投资领域的策略升级，更是各行各业在数字化浪潮中构建核心竞争力的缩影，我们需要顺应数字化时代发展，抓住各类机遇，挖掘各种智能投资决策方案。

## 致谢

感谢所有为本文提供帮助的人。

## 参考文献

- [1] 张虎, 向宇, 林宗海. 数字化转型对股票流动性的影响研究——来自我国制造业上市公司的证据[J]. 数量经济研究, 2025, 16(4): 1-20.
- [2] 陈奋强, 王焰辉. 数字金融何以赋能资本市场高质量发展——基于股价同步性视角[J]. 上海金融, 2024(10): 55-67.
- [3] 李红梅. 机器学习方法和统计建模方法的预测比较研究[D]: [硕士学位论文]. 昆明: 云南师范大学, 2016.
- [4] 刘佳祺, 刘德红, 林甜甜. 基于 BP 神经网络模型的股票价格研究[J]. 中国商论, 2018(8): 29-30.
- [5] 周宁. 基于神经网络和马尔可夫组合模型的股票价格预测[C]//中国自动化学会控制理论专业委员会. 第 36 届中国控制会议论文集. 2017: 545-550.
- [6] 苏兆辉, 尚领, 刘志中, 等. 一种基于改进经验模态分解与 A-LSTM 混合神经网络的股价预测方法[J]. 南京大学学报(自然科学), 2025, 61(4): 613-623.
- [7] 李彦冬, 郝宗波, 雷航. 卷积神经网络研究综述[J]. 计算机应用, 2016, 36(9): 2508-2515.
- [8] 陈毅. 基于多种神经网络的股价预测研究[D]: [硕士学位论文]. 贵阳: 贵州财经大学, 2022.
- [9] 宋罗瑶. 基于 Lasso 与集成学习的川渝地区二手车价格预测分析[D]: [硕士学位论文]. 重庆: 西南大学, 2024.
- [10] 耿晶晶, 刘玉敏, 李洋, 赵哲耘. 基于 CNN-LSTM 的股票指数预测模型[J]. 统计与决策, 2021, 37(5): 134-138.
- [11] 马美琛. 基于机器学习的股票趋势预测和投资策略研究[D]: [硕士学位论文]. 石家庄: 河北经贸大学, 2024.