

基于神经网络集成学习算法的金融时间序列预测^①



徐晓芳, 管 瑞

(南开大学 金融学院, 天津 300350)

通信作者: 徐晓芳, E-mail: xuxf_7@126.com

摘 要: 本文在传统神经网络 (NN)、循环神经网络 (RNN)、长短时记忆网络 (LSTM) 与门控循环单元 (GRU) 等神经网络时间预测模型基础上, 进一步构建集成学习 (EL) 时间序列预测模型, 研究神经网络类模型、集成学习模型和传统时间序列模型在股票指数预测上的表现. 本文以 16 只 A 股和国际股票市场指数为样本, 比较模型在不同预测期间和不同国家和地区股票市场上的表现. 本文主要结论如下: 第一, 神经网络类时间序列预测模型和神经网络集成学习时间序列预测模型在表现上显著稳健优于传统金融时间序列预测模型, 预测性能提高大约 35%; 第二, 神经网络类模型和神经网络集成学习模型在中国和美国股票市场上的表现优于其他发达国家和地区的股票市场.

关键词: 神经网络; 集成学习; 金融时间序列; 预测; 机器学习

引用格式: 徐晓芳, 管瑞. 基于神经网络集成学习算法的金融时间序列预测. 计算机系统应用, 2022, 31(6): 29-37. <http://www.c-s-a.org.cn/1003-3254/8551.html>

Financial Time Series Forecasting Based on Neural Network Ensemble Learning Algorithms

XU Xiao-Fang, GUAN Rui

(School of Finance, Nankai University, Tianjin 300350, China)

Abstract: On the basis of time series forecast models of neural networks (NNs) such as the traditional NN, recurrent neural network (RNN), long short-term memory (LSTM), and gated recurrent unit (GRU), this study builds a time series forecast model of ensemble learning (EL) to study the performance of NN models, the EL model, and traditional time series models in stock index prediction. This study takes 16 Chinese and international stock market indexes as samples to compare the performance of the models in different forecast periods and stock markets in different countries and regions. The main conclusions of this study are as follows: First, the NN time series forecast model and the EL time series forecast model based on NNs are significantly more robust than the traditional financial time series forecast model, and the prediction performance is improved by about 35%. Second, the performance of NN models and EL models in Chinese and American stock markets is better than that of the rest of developed countries and regions.

Key words: neural network; ensemble learning; financial time series; forecast; machine learning

金融时间序列预测是金融领域中一个重要研究方向, 可以为二级股票市场投资提供市场走势、风险和入场出场时机的预测. 传统金融时间序列预测方法包括技术分析和计量经济学模型. 技术分析是利用技术指标或者结合多种技术指标对未来一段时间股票走势进行方向性预测, 计量经济学方法是利用多元线性回

归、ARIMA 模型等模型对历史股票数据进行最优化拟合, 然后对未来进行预测.

人工智能可以学习到时间序列中的非线性关系, 能够有效处理金融数据中低信噪比问题, 在金融时间序列预测上具有显著的优势, 尤其是以大数据为背景的神经网络类模型. 国外相关文献主要以宏观经济数

① 收稿时间: 2021-08-22; 修改时间: 2021-09-29; 采用时间: 2021-11-07; csa 在线出版时间: 2022-05-26

据和股票价格预测为主, Ahmed 等利用传统机器学习模型对宏观经济中 M3 时间序列进行预测^[1]; Cao 等利用 EMD 分解和 LSTM 模型对美国股票价格时间序列进行预测, 获得了较高的预测精度^[2]; Siami-Namini 等比较了 LSTM 模型和 ARIMA 模型在美国重要股票指数上的预测水平差异^[3]; Dingli 等研究了 CNN 模型在美国股票价格序列上的预测能力^[4]; Kumar 等测试了 LSTM 对纳斯达克股票价格的预测能力^[5]; Shen 等比较了 GRU 和 SVM 在美国和欧洲不同股票指数上的预测能力^[6]. 国内相关文献主要以单只股票指数预测为主, 谷丽琼等利用结合注意机制的 GRU 模型对科大讯飞股票价格进行预测, 预测效果优于其他模型^[7]; 乔若羽比较研究神经网络在上证指数上的预测能力, 研究了模型优化方向^[8]; 李坤等利用支持向量机对 A 股大盘指数和个股进行预测^[9]; 王文波等结合 EMD 分解和神经网络对上证指数和深证成指进行预测^[10].

本文在传统神经网络 (NN)、RNN、LSTM 和 GRU 等模型基础上, 构建集成学习模型, 研究神经网络类模型和集成学习模型在金融时间序列预测上的表现. 本文选择 ARIMA 模型和线性回归模型为基准模型, 选择 16 只 A 股市场和全球重要发达国家和地区股票市场指数为样本, 研究模型在长期、中期和短期 3 种不同预测期间上的预测能力, 比较模型在不同国家和地区的适用情况.

1 模型构建

1.1 传统神经网络模型

神经网络模型由全连接层组成, 按照全连接层作用分类, 可以分为输入层、隐藏层和输出层, 如图 1(a). 神经网络中的节点是非线性转换单元, 负责对输入进行线性组合并且利用激活函数进行非线性输出, 如图 1(b).

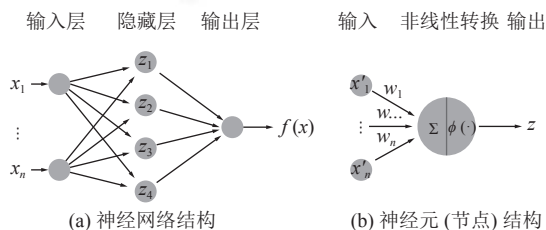


图 1 传统神经网络模型

激活函数是神经网络可以进行非线性拟合的核心, 常见激活函数有 S 型函数 (Sigmoid)、双曲正切函数

(tanh) 和修正线性单元 (ReLU), 如式 (1).

$$\begin{cases} \text{Sigmoid}(x) = \frac{1}{1 + e^{-x}} \\ \text{tanh}(x) = \frac{2}{1 + e^{-2x}} - 1 \\ \text{ReLU}(x) = \max(0, x) \end{cases} \quad (1)$$

1.2 循环神经网络

传统神经网络的输入是同质的, 无法区分输入序列的先后关系, RNN 将输入看成一个新的维度, 允许模型按照时间顺序输入特征, 如图 2 所示.

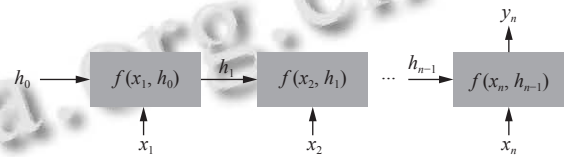


图 2 RNN 模型

RNN 模型输入向量为 (x_1, x_2, \dots, x_n) , 输出为 y_n , 记忆向量为 h_t , RNN 前向运算过程如式 (2) 和式 (3), 其中 W_{hh}, W_{xh}, W_{hy} 是权重共享的参数矩阵, 是 RNN 最后的输出^[11].

$$h_t = \tanh(W_{hh}h_{t-1} + W_{xh}x_t) \quad (2)$$

$$y_t = W_{hy}h_t + b \quad (3)$$

RNN 虽然可以利用状态向量拥有记忆能力, 但是记忆能力有限, 只能记忆当前时刻最近的状态, 无法处理长期依赖问题. LSTM 模型解决了长期依赖问题, 拥有长期记忆能力, 同时也解决了 RNN 梯度爆炸和梯度离散的问题.

LSTM 由遗忘门 f_t , 输入门 i_t , 输出门 o_t 组成, 输入当前时刻输入 x_t , 上一时刻单元状态 c_{t-1} 和隐藏状态 h_{t-1} , 输出 c_t, h_t 和 y_t . LSTM 前向运算如图 3 所示, 其中, * 表示矩阵对应元素乘法, + 表示矩阵加法^[12-14]. LSTM 数学运算过程如式 (4).

$$\begin{cases} f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \\ i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \\ o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \\ z_t = \tanh(W_z[h_{t-1}, x_t] + b_z) \\ c_t = f_t * c_{t-1} + i_t * z_t \\ h_t = o_t * \tanh(c_t) \end{cases} \quad (4)$$

其中, σ 是 Sigmoid 激活函数, 输出在 0 到 1 之间, \tanh 是激活函数, 输出在 -1 到 1 之间, * 表示矩阵中对应元素相乘. LSTM 依然采用权值共享机制, 不同时刻的参数矩阵相同.

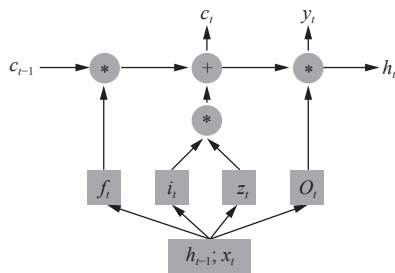


图3 LSTM模型

GRU 是 LSTM 的简化版本, 性能与 LSTM 类似, 但是大大减少了计算复杂度, 提高了计算效率^[15]. GRU 减少了门控的数量并且去掉了状态向量 c_t , 只保留重置门 r_t 和更新门 z_t , 如图 4 所示.

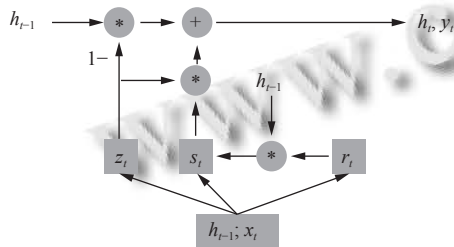


图4 GRU模型

GRU 输入为 h_{t-1} 和 x_t , 输出为 h_t 和 y_t , h_t 作为下一个时刻的输入, y_t 作为最终输出结果, GRU 的数据运算过程如式 (5).

$$\begin{cases} z_t = \sigma(W_z[h_{t-1}, x_t] + b_z) \\ r_t = \sigma(W_r[h_{t-1}, x_t] + b_r) \\ s_t = \tanh(W_s[h_{t-1} * r_t, x_t] + b_s) \\ h_t = (1 - z_t) * h_{t-1} + z_t * s_t \end{cases} \quad (5)$$

其中, W_z, W_r, W_s 是参数矩阵, b_z, b_r, b_s 是偏置向量, z_t, r_t 分别代表更新门和重置门, s_t 是中间变量, h_t 是 GRU 的输出, 最后时刻的 h_t 是 GRU 最终输出 y_t , * 表示矩阵中对应元素相乘.

1.3 集成学习

股票指数在不同的资本市场和不同的时期会呈现出不同的趋势和特点. 机器学习模型有其各自的优缺点和最优的适用场景, 单一机器学习模型很难适用于股票市场中的所有情况. 集成学习 (ensemble learning) 可以结合多个基学习器, 获得一个表现稳定且优异的强学习器, 能够同时减少预测偏差和方差^[16].

本文选择传统神经网络 (NN)、RNN、LSTM 和 GRU 4 个模型作为基学习器, 利用 bagging 集成学习构建强学习器, 如图 5 所示.

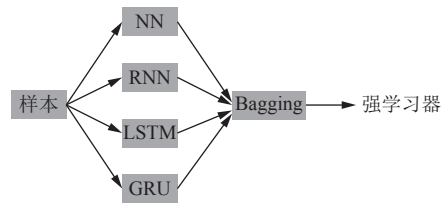


图5 EL模型

假设金融时间序列为 $\{x_t\}(t = 1, 2, 3, \dots)$, 第 $i (i = 1, 2, 3, 4)$ 个基学习器如式 (6) 所示:

$$x_t^i = f_i(x_{t-1}, x_{t-2}, \dots, x_{t-n}) \quad (6)$$

其中, 输入为 $x_{t-1}, x_{t-2}, \dots, x_{t-n}$, 输出 x_t^i 为第 i 个基学习器在 t 时刻的预测值. 由于股票指数时间序列具有渐进独立性, 所以本文选择 n 等于 30, 即前 30 个交易日为自变量, 第 31 个交易日为因变量, 构建神经网络集成学习模型.

集成学习在 t 时刻的最终预测值 \hat{x}_t 等于 4 个基学习器的预测值的简单算术平均数, 如式 (7):

$$\hat{x}_t = \frac{1}{4} \sum_{i=1}^4 x_t^i \quad (7)$$

1.4 传统金融时间序列模型

传统金融时间序列模型主要包括 n 阶自回归模型 ($AR(n)$ 模型) 和差分自回归移动平均模型 ($ARIMA(p, d, q)$ 模型).

假设金融时间序列为 $\{x_t\} (t = 1, 2, 3, \dots)$, ε_t 为白噪声, $AR(n)$ 模型表达式如式 (8) 所示, 其本质是多元线性回归模型. 为了与神经网络模型保持一致, 本文选择 $AR(30)$ 模型作为基准模型.

$$x_t = a_1 x_{t-1} + a_2 x_{t-2} + \dots + a_n x_{t-n} + \varepsilon_t \quad (8)$$

$ARMA(p, q)$ 模型表达式如式 (9) 所示:

$$x_t = \sum_{i=1}^p a_i x_{t-i} + \sum_{j=1}^q b_j \varepsilon_{t-j} + \varepsilon_t \quad (9)$$

对于非平稳金融时间序列, 传统金融计量经济学引用 d 阶差分来平稳化金融时间序列, 并且使用经过差分后的金融时间序列进行 $ARMA(p, q)$ 建模, 即 $ARIMA(p, d, q)$ 模型^[17,18].

2 实证研究

2.1 数据选取

本文数据来源是 wind 数据库, 样本时间范围为

2005年1月1日到2020年12月31日. 本文选择7只我国A股市场指数和9只国际市场指数共计16只重要股票指数作为样本, 比较机器学习模型和传统时间序列预测模型在不同市场和拥有不同成分股的股票指数上的预测能力, 增加了结论的可靠性和稳健性.

我国A股市场指数分别为上证综指(000001), 上证50(000016), 上证180(000010), 沪深300(000300), 深证成指(399001), 创业板指(399006)和中小板指(399005)^[19]; 全球市场指数分别为道琼斯工业指数(DJI), 标普500(SPX), 纳斯达克指数(IXIC), 恒生指数(HSI), 法国CAC40(FCHI), 英国富士100(FTSE), 德国DAX(GDAXI), 日经225(N225)和韩国综合指数(KS11).

2.2 评价指标与DM检验

本文使用平均绝对误差(MAE)度量模型预测能力, 并且比较神经网络集成学习(EL)与传统金融时间序列预测模型直接预测误差MAE衡量EL性能提升水平. 本文进一步使用Diebold-Mariano检验来对模型预测能力进行统计显著性检验, 比较神经网络类模型和集成学习是否显著优于传统ARIMA和AR模型^[20,21].

MAE用于衡量预测值与真实值误差绝对值的均值, 可以比较模型预测能力的优劣, 本文为了便于展示, 将MAE结果扩大了100倍, 如式(10).

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \times 100 \quad (10)$$

EL性能提升水平是基于模型预测误差MAE计算的模型预测误差提升百分比, 如式(11), 其中MAE_{other}代表其他模型的预测误差MAE, MAE_{EL}代表神经网络集成学习模型预测误差MAE.

$$EL_{Growth} = \frac{MAE_{other} - MAE_{EL}}{MAE_{EL}} \times 100\% \quad (11)$$

Diebold和Mariano(2002)^[22]提出DM检验, 检验两个模型预测精度是否存在显著的差异. DM检验假设模型与模型在时间跨度的预测误差序列向量分别为: $\mathbf{e}_A = (a_1, a_2, \dots, a_T)$, $\mathbf{e}_B = (b_1, b_2, \dots, b_T)$, 二者的差值序列向量为: $\mathbf{d} = (d_1, d_2, \dots, d_T)$, 其中 $d_i = a_i - b_i$ ($i = 1, 2, \dots, T$). DM统计量表达式如式(12).

$$DM = \frac{\sqrt{T}\bar{d}}{\sigma(\mathbf{d})} \quad (12)$$

其中, $\sigma(\mathbf{d})$ 为差值序列向量 \mathbf{d} 的标准差, \bar{d} 为差值序列向量 \mathbf{d} 均值, DM统计量的分布服从标准正态分布.

2.3 模型参数

神经网络模型超参数包括隐藏层层数、神经元个数、激活函数和模型优化器等. 为了提高模型的收敛速度, 本文对数据进行最大最小值标准化, 对神经网络类模型采用随机梯度下降(SGD)优化器, 并且配合学习率递减策略^[23]. 4种机器学习器的具体超参数取值如表1所示. 本文基准模型选择ARIMA(p, d, q)模型和AR(n)模型, 两个基准模型的参数优化及选择如表2和表3所示.

表1 4种机器学习器超参数取值

超参数	NN	RNN	LSTM	GRU
隐藏层层数	1	1	1	1
神经元个数	10	100	100	100
激活函数	ReLU	tanh	tanh	tanh
损失函数	MAE	MAE	MAE	MAE
模型优化器	SGD	SGD	SGD	SGD
迭代次数	100	100	100	100
学习率	衰减	衰减	衰减	衰减
Batch size	64	64	64	64
Dropout	1.0	0.8	0.8	0.8

表2 ARIMA(p, d, q)模型超参数选择

超参数	参数取值
模型参数选取准则	AIC准则
参数 p	$2 \leq p \leq 5$
参数 d	$0 \leq d \leq 2$
参数 q	$2 \leq q \leq 5$
是否季节性调整	是
时间序列平稳性检验方法	KPSS

表3 AR(n)模型超参数选择

超参数	参数取值
滞后阶数 n	30
优化方法	OLS

2.4 实证结果

本文选择全球16只重要的股票价格指数时间序列为样本, 时间跨度为2005年1月1日到2020年12月31日, 共计3889个交易日, 由于我国A股市场某些指数编制起始时间不同, 交易日个数略少, 具体如表4.

表4 样本数据时间跨度

股票指数	样本初始时间	样本结束时间	交易日个数
中小板指	2005年6月7日	2020年12月31日	3633
创业板指	2010年6月1日	2020年12月31日	2575
沪深300	2005年4月8日	2020年12月31日	3828
其他	2005年1月1日	2020年12月31日	3889

本文将样本划分为不同长度的训练集和测试集,用以检验神经网络集成学习模型在不同预测期间长度上的表现,具体可分为长期预测(100个交易日预测),中期预测(50个交易日预测)和短期预测(30个交易日预测),即选择样本中最后100、50和30个交易日数据作为测试集,其他交易日数据作为训练集。

实证结果如表5-表7所示。表5-表7以MAE为评价指标,分别展示了神经网络类模型和神经网络集成学习模型在全球不同市场不同股票指数下,在不同预测期间的预测能力,比较了神经网络集成学习(EL)相比于传统金融时间序列预测模型ARIMA和AR模型的性能提升水平。

表5 长期模型预测误差(MAE)和性能提升

指数	模型预测误差						EL性能提升(%)		
	NN	RNN	LSTM	GRU	EL	ARIMA	AR	EL_AR	EL_ARIMA
000001	0.562	0.511	0.695	0.587	0.554	1.191	0.844	34.36	53.48
000016	0.779	0.704	0.930	0.771	0.748	1.797	1.158	35.41	58.38
000010	0.786	0.712	1.007	0.832	0.778	1.675	1.201	35.22	53.55
000300	0.904	0.813	1.145	0.982	0.911	2.527	1.295	29.65	63.95
399001	0.920	0.816	1.149	0.960	0.914	2.570	1.176	22.28	64.44
399006	1.238	1.193	1.538	1.302	1.265	5.161	1.651	23.38	75.49
399005	0.981	0.909	1.275	1.056	1.008	2.493	1.333	24.38	59.57
DJI	1.260	1.112	1.711	1.411	1.291	5.633	1.957	34.03	77.08
SPX	1.238	1.212	1.684	1.370	1.314	5.115	2.038	35.53	74.31
IXIC	1.386	1.605	1.935	1.519	1.515	11.608	2.757	45.05	86.95
HSI	1.005	0.903	1.225	1.068	1.023	3.972	1.272	19.58	74.24
FCHI	1.603	1.402	1.912	1.651	1.577	7.803	1.685	6.41	79.79
FTSE	1.504	1.301	1.806	1.510	1.473	5.576	1.576	6.54	73.58
GDAXI	1.472	1.282	1.905	1.603	1.489	4.717	1.771	15.92	68.43
N225	0.953	0.969	1.189	1.028	0.990	9.695	1.270	22.05	89.79
KS11	1.363	1.184	1.722	1.424	1.380	8.832	1.933	28.61	84.38

表6 中期模型预测误差(MAE)和性能提升

指数	模型预测误差						EL性能提升(%)		
	NN	RNN	LSTM	GRU	EL	ARIMA	AR	EL_AR	EL_ARIMA
000001	0.518	0.451	0.623	0.529	0.502	1.344	0.769	34.72	62.65
000016	0.764	0.642	0.920	0.738	0.714	2.332	1.130	36.81	69.38
000010	0.787	0.659	0.969	0.813	0.757	2.336	1.134	33.25	67.59
000300	0.852	0.740	1.077	0.876	0.837	3.439	1.189	29.60	75.66
399001	0.804	0.689	0.959	0.793	0.758	2.480	1.034	26.69	69.44
399006	1.083	0.954	1.295	1.084	1.058	2.153	1.430	26.01	50.86
399005	0.826	0.727	1.079	0.882	0.841	1.641	1.155	27.19	48.75
DJI	1.280	1.112	1.694	1.400	1.314	5.478	2.263	41.94	76.01
SPX	1.149	1.032	1.489	1.235	1.168	5.272	2.272	48.59	77.85
IXIC	1.106	1.217	1.375	1.172	1.130	3.417	2.610	56.70	66.93
HSI	0.938	0.805	1.054	0.944	0.902	6.300	1.117	19.25	85.68
FCHI	1.691	1.436	3.220	1.810	1.863	15.121	1.961	5.00	87.68
FTSE	1.552	1.276	1.963	1.602	1.557	12.379	1.589	2.01	87.42
GDAXI	1.560	1.320	2.153	1.747	1.624	7.217	2.122	23.47	77.50
N225	1.142	0.998	1.572	1.319	1.226	11.329	1.266	3.16	89.18
KS11	1.168	1.010	1.494	1.219	1.166	12.801	1.867	37.55	90.89

由表5-表7可知,NN、RNN、LSTM、GRU和集成学习模型的样本外预测能力显著优于传统ARIMA和线性回归模型,在不同股票指数和不同预测时间长度上的表现都优于传统模型,其平均预测性能提升大

约35%。

LSTM和GRU模型在英国富士100和日经225等极少数股票指数上表现不如线性回归模型,其他情况下都优于ARIMA和AR模型。其中原因是股票指数

与近期股市表现相关性最强,与长期表现相关性变弱,导致具有长期记忆能力的 LSTM 和 GRU 模型预测能力下降.集成学习作为强学习器,在预测稳定性和预测

精度上优于其他神经网络类模型,集成学习在不同期限上的样本外预测值和真实值的时间序列,如图 6-图 8 所示.

表 7 短期模型预测误差 (MAE) 和性能提升

指数	模型预测误差					EL性能提升 (%)			
	NN	RNN	LSTM	GRU	EL	ARIMA	AR	EL_AR	EL_ARIMA
000001	0.513	0.446	0.576	0.517	0.483	0.644	0.775	37.68	25.00
000016	0.764	0.635	0.860	0.708	0.695	1.841	1.114	37.61	62.25
000010	0.793	0.670	0.907	0.791	0.738	1.628	1.148	35.71	54.67
000300	0.864	0.723	0.973	0.835	0.794	1.553	1.242	36.07	48.87
399001	0.757	0.636	0.800	0.719	0.687	1.048	1.093	37.15	34.45
399006	0.949	0.835	1.000	0.888	0.869	3.208	1.336	34.96	72.91
399005	0.660	0.641	0.861	0.713	0.677	1.210	1.099	38.40	44.05
DJI	0.530	0.520	0.509	0.521	0.486	2.287	1.258	61.37	78.75
SPX	0.565	0.573	0.636	0.616	0.561	2.928	1.352	58.51	80.84
IXIC	0.557	0.665	0.699	0.643	0.561	3.157	1.548	63.76	82.23
HSI	0.835	0.730	0.829	0.790	0.775	0.994	1.031	24.83	22.03
FCHI	0.879	0.745	0.959	0.875	0.807	2.094	1.493	45.95	61.46
FTSE	1.139	0.897	1.240	1.060	1.056	3.446	1.188	11.11	69.36
GDAXI	0.994	0.803	0.980	0.930	0.903	3.157	1.476	38.82	71.40
N225	0.934	0.884	1.158	1.025	0.973	5.034	1.135	14.27	80.67
KS11	1.051	0.983	1.170	1.060	1.026	8.969	1.967	47.84	88.56

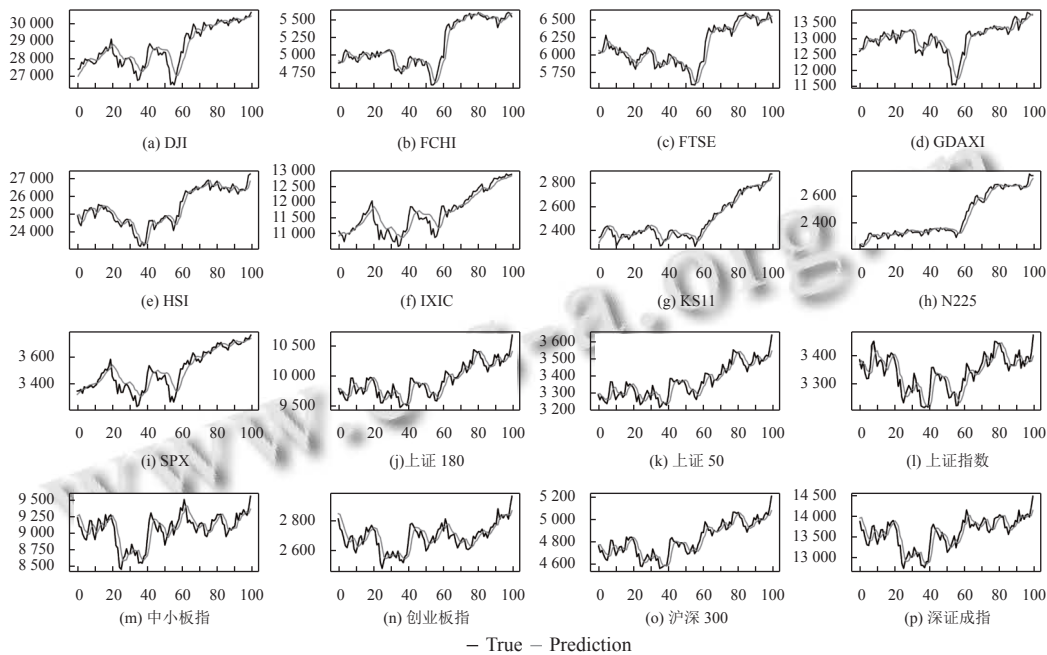


图 6 集成学习在长期预测 (100 天) 上的表现 (横坐标表示天数, 纵坐标表示股票指数)

神经网络类模型和集成学习在不同国家和地区的股票市场中的表现具有一定的差异.在中国股票市场和美国股票市场上,神经网络类模型和集成学习比 ARIMA 模型性能提升大约 45%,比 AR 模型性能提升

大约 35%;在其他发达国家股票市场上,神经网络类模型和集成学习比 ARIMA 模型依然有显著的优势,性能提升大约 70%,但是比 AR 模型性能提升只有 15%左右.

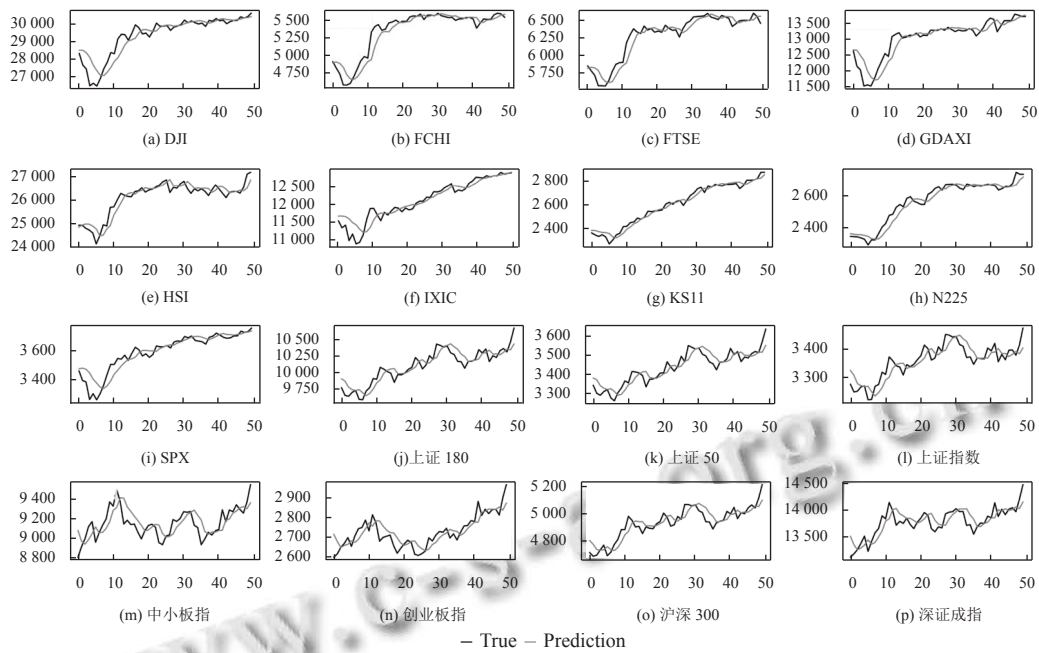


图7 集成学习在中期预测(50天)上的表现(横坐标表示天数,纵坐标表示股票指数)

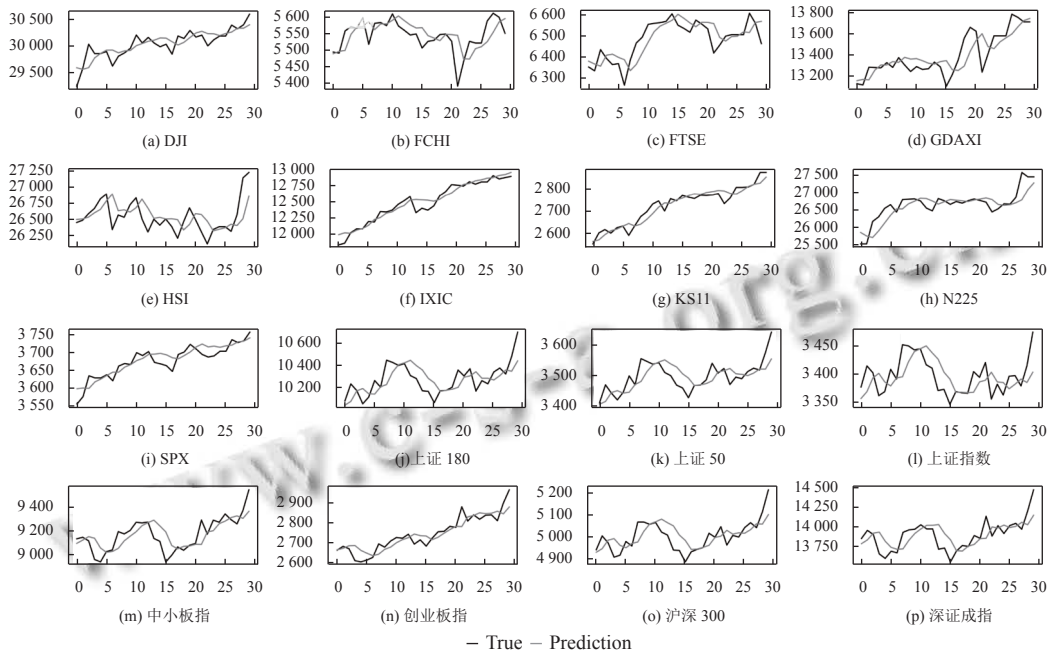


图8 集成学习在短期预测(30天)上的表现(横坐标表示天数,纵坐标表示股票指数)

本文利用 DM 检验,进一步检验神经网络类模型和集成学习在不同预测期间和不同国家股票市场上的表现的是否显著优于传统时间序列预测模型,结果如表 8 所示.表 8 检验了在不同预测期限上,神经网络类模型和集成学习分别在中国股票市场、美国股票市场

和其他发达国家股票市场上相比于传统时间序列模型是否存在显著的优势.NN、RNN、LSTM、GRU 和集成学习模型在中国市场和美国市场上的预测能力显著优于传统时间序列模型,但是在其他发达国家市场的表现不具有统计显著性.

表8 不同股票市场预测能力DM检验

方法	长期预测			中期预测			短期预测		
	China	US	Others	China	US	Others	China	US	Others
NN	4.36 (0.00)	5.18 (0.00)	2.49 (0.13)	3.42 (0.02)	4.09 (0.00)	1.91 (0.15)	3.21 (0.02)	3.58 (0.01)	1.69 (0.23)
RNN	4.73 (0.00)	4.79 (0.00)	3.23 (0.01)	3.74 (0.00)	3.95 (0.00)	2.47 (0.02)	3.54 (0.00)	3.36 (0.00)	2.33 (0.05)
LSTM	1.6 (0.30)	2.66 (0.03)	0.86 (0.28)	1.39 (0.28)	3.04 (0.02)	-1.28 (0.09)	1.85 (0.10)	3.49 (0.00)	1.30 (0.19)
GRU	3.84 (0.00)	4.71 (0.00)	1.39 (0.37)	2.95 (0.03)	4.01 (0.00)	0.48 (0.47)	2.89 (0.02)	3.44 (0.01)	1.86 (0.24)
EL	4.29 (0.00)	4.89 (0.00)	2.11 (0.24)	3.40 (0.01)	4.05 (0.00)	1.02 (0.45)	3.28 (0.01)	3.61 (0.00)	2.09 (0.14)

注: 表格中数字代表DM统计量, 括号内代表 p 值。

2.5 结果分析

ARIMA 和 AR 等传统金融时间序列预测模型只能捕获时间序列中的线性关系, 无法捕获时间序列中的非线性关系. 股票指数是包含了众多噪音在内的非线性时间序列, 线性模型只能预测股票指数的线性趋势部分, 无法预测股票指数非线性趋势部分.

神经网络类模型可以捕获时间序列中的非线性关系, 集成学习可以有效地降低单一神经网络模型的过拟合风险. 基于神经网络类模型的集成学习不仅仅可以有效地识别出股票指数中的非线性关系, 还可以避免模型拟合过多的噪音, 增强模型的泛化能力. 所以神经网络集成学习在金融时间序列预测上的表现远远优于传统金融时间序列预测模型.

3 结论

本文以 7 只 A 股市场指数和 9 只国际市场指数为样本, 研究了神经网络类模型和基于神经网络的集成学习在金融时间序列上的预测能力与传统时间序列模型的预测能力的差异, 提出基于神经网络类模型的集成学习时间序列预测模型, 提高了金融时间序列的样本外预测能力.

本文主要结论如下: (1) 神经网络类模型显著优于 ARIMA 模型和线性回归模型. LSTM 和 GRU 模型在英国富士 100 和日经 225 等极少数股票指数上表现不如线性回归模型, 在其他股票指数上的表现显著优于 ARIMA 和 AR 等传统时间序列模型. NN、RNN 和集成学习模型显著优于传统时间序列模型. (2) 基于神经网络类的集成学习模型在所有机器学习模型中表现最稳定, 在短期预测、中期预测和长期预测和全部股票

市场指数上的表现显著优于传统时间序列模型. (3) 神经网络类模型和基于神经网络的集成学习模型在中国股票市场和美国股票市场中的表现显著优于其他发达国家股票市场.

本文的研究拓展了金融时间序列预测的理论研究, 为股票投资提供风险参考和入场出场时机参考. 在宏观上, 本文研究可以为投资者提供股票走势的预测和风险预测; 在微观上, 本文研究可以减少股票市场不必要的波动, 促进中国股票市场合理定价, 促进中国股票市场繁荣稳定发展.

参考文献

- 1 Ahmed NK, Atiya AF, El Gayar N, *et al.* An empirical comparison of machine learning models for time series forecasting. *Econometric Reviews*, 2010, 29(5-6): 594–621. [doi: 10.1080/07474938.2010.481556]
- 2 Cao J, Li Z, Li J. Financial time series forecasting model based on CEEMDAN and LSTM. *Physica A: Statistical Mechanics and its Applications*, 2019, 519: 127–139. [doi: 10.1016/j.physa.2018.11.061]
- 3 Siami-Namini S, Tavakoli N, Namin AS. A comparison of ARIMA and LSTM in forecasting time series. 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA). Orlando: IEEE, 2018. 1394–1401.
- 4 Dingli A, Fournier KS. Financial time series forecasting—A deep learning approach. *International Journal of Machine Learning and Computing*, 2017, 7(5): 118–122. [doi: 10.18178/ijmlc.2017.7.5.632]
- 5 Kumar S, Ningombam D. Short-term forecasting of stock prices using long short term memory. 2018 International Conference on Information Technology (ICIT).

- Bhubaneswar: IEEE, 2018. 182–186.
- 6 Shen GZ, Tan QP, Zhang HY, *et al.* Deep learning with gated recurrent unit networks for financial sequence predictions. *Procedia Computer Science*, 2018, 131: 895–903. [doi: [10.1016/j.procs.2018.04.298](https://doi.org/10.1016/j.procs.2018.04.298)]
 - 7 谷丽琼, 吴运杰, 逢金辉. 基于 Attention 机制的 GRU 股票预测模型. *系统工程*, 2020, 38(5): 134–140.
 - 8 乔若羽. 基于神经网络的股票预测模型. *运筹与管理*, 2019, 28(10): 132–140.
 - 9 李坤, 谭梦羽. 基于小波支持向量机回归的股票预测. *统计与决策*, 2014, (6): 32–36.
 - 10 王文波, 费浦生, 羿旭明. 基于 EMD 与神经网络的中国股票市场预测. *系统工程理论与实践*, 2010, 30(6): 1027–1033. [doi: [10.12011/1000-6788\(2010\)6-1027](https://doi.org/10.12011/1000-6788(2010)6-1027)]
 - 11 Zhang JS, Xiao XC. Predicting chaotic time series using recurrent neural network. *Chinese Physics Letters*, 2000, 17(2): 88. [doi: [10.1088/0256-307X/17/2/004](https://doi.org/10.1088/0256-307X/17/2/004)]
 - 12 Hua YX, Zhao ZF, Li RP, *et al.* Deep learning with long short-term memory for time series prediction. *IEEE Communications Magazine*, 2019, 57(6): 114–119. [doi: [10.1109/MCOM.2019.1800155](https://doi.org/10.1109/MCOM.2019.1800155)]
 - 13 邓飞燕, 岑少琪, 钟凤琪, 等. 基于 LSTM 神经网络的短期价格趋势预测. *计算机系统应用*, 2021, 30(4): 187–192. [doi: [10.15888/j.cnki.csa.007855](https://doi.org/10.15888/j.cnki.csa.007855)]
 - 14 姜秋龙, 徐晓钟. 基于 FCGA 和改进 LSTM-BPNN 的燃气负荷预测. *计算机系统应用*, 2021, 30(4): 1–8. [doi: [10.15888/j.cnki.csa.007760](https://doi.org/10.15888/j.cnki.csa.007760)]
 - 15 Tseng FM, Yu HC, Tzeng GH. Combining neural network model with seasonal time series ARIMA model. *Technological Forecasting and Social Change*, 2002, 69(1): 71–87. [doi: [10.1016/S0040-1625\(00\)00113-X](https://doi.org/10.1016/S0040-1625(00)00113-X)]
 - 16 Chitra A, Uma S. An ensemble model of multiple classifiers for time series prediction. *International Journal of Computer Theory and Engineering*, 2010, 2(3): 454–458.
 - 17 Ho SL, Xie M, Goh TN. A comparative study of neural network and Box-Jenkins ARIMA modeling in time series prediction. *Computers & Industrial Engineering*, 2002, 42(2–4): 371–375.
 - 18 李勃旭, 南西康, 郑向东, 等. 基于 EMD-ARIMA 模型的地铁门传动系统早期故障预测. *计算机系统应用*, 2019, 28(9): 110–117. [doi: [10.15888/j.cnki.csa.007090](https://doi.org/10.15888/j.cnki.csa.007090)]
 - 19 张聪. Kelly 准则下的组合选择策略及其经验证据. *东北财经大学学报*, 2018, (6): 95–104.
 - 20 Diebold FX, Mariano RS. Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 1995, 13(3): 253–263.
 - 21 陈昊, 王玉荣. 一种负荷预测模型预测能力的评价标准. *电力需求侧管理*, 2010, 12(6): 24–26. [doi: [10.3969/j.issn.1009-1831.2010.06.008](https://doi.org/10.3969/j.issn.1009-1831.2010.06.008)]
 - 22 Diebold FX, Mariano RS. Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 2002, 20(1): 134–144.
 - 23 李海涛, 王腾, 王印庚. 基于 VGG-16 卷积神经网络的海水养殖病害诊断. *计算机系统应用*, 2020, 29(7): 222–227. [doi: [10.15888/j.cnki.csa.007483](https://doi.org/10.15888/j.cnki.csa.007483)]