AdaBoost-LSTM Ensemble Learning for Financial Time Series Forecasting

Shaolong Sun[1,2], Yunjie Wei[1,3], Shouyang Wang[1,2,3]

[1] Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China
[2] School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100190, China
[3] Center for Forecasting Science, Chinese Academy of Sciences, Beijing 100190, China
weiyunjie@amss.ac.cn

Abstract. A hybrid ensemble learning approach is proposed to forecast financial time series combining AdaBoost algorithm and Long Short-Term Memory (LSTM) network. Firstly, by using AdaBoost algorithm the database is trained to get the training samples. Secondly, the LSTM is utilized to forecast each training sample separately. Thirdly, AdaBoost algorithm is used to integrate the forecasting results of all the LSTM predictors to generate the ensemble results. Two major daily exchange rate datasets and two stock market index datasets are selected for model evaluation and comparison. The empirical results demonstrate that the proposed AdaBoost-LSTM ensemble learning approach outperforms some other single forecasting models and ensemble learning approaches. This suggests that the AdaBoost-LSTM ensemble learning approach is a highly promising approach for financial time series data forecasting, especially for the time series data with nonlinearity and irregularity, such as exchange rates and stock indexes.

Keywords: Financial time series forecasting, long short-term memory network, AdaBoost algorithm, ensemble learning.

1 Introduction

Financial markets are affected by many factors, such as economic conditions, political events, traders' expectations and so on. Hence, financial time series forecasting is usually regarded as one of the most challenging tasks due to the nonlinearity and irregularity. How to forecast financial time series accurately is still an open question with respect to the economic and social organization of modern society. Many common econometric and statistical models have been applied to forecast financial time series, such as autoregressive integrated moving average (ARIMA) model [1], vector auto-regression (VAR) model [2] and error correction model (ECM) [3]. However, traditional models fail to capture the nonlinearity and complexity of financial time series which lead to poor forecasting accuracy. Hence, exploring more effective forecasting methods, which possess enough learning capacity, is really necessary for fore-

ICCS Camera Ready Version 2018
To cite this paper please use the final published version:
DOI: 10.1007/978-3-319-93713-7_55

casting financial time series. Thus, nonlinear and more complex artificial intelligence methods are introduced to forecast financial time series, such as artificial neural networks (ANNs) [4-5], support vector regression (SVR) [6] and deep learning method [7].

The forecasting accuracy of those nonlinear artificial intelligence methods are usually better than the common econometric and statistical models, while they also suffer from many problems, such as parameter optimization and overfitting. Hence, many hybrid forecasting approaches are proposed to get better forecasting performance [8-13]. So far, the decomposition ensemble learning approach has been widely used to forecast time series in many fields, such as financial time series forecasting [14-15], crude oil price forecasting [16], nuclear energy consumption forecasting [17], PM2.5 concentration forecasting [18], etc. According to the existing literatures, ANNs are the most common used methods both in single model forecasting and hybrid model forecasting, which demonstrates that ANNs are really suitable for time series forecasting. If the advantages of different ANNs methods are combined, a better forecasting performance can be obtained. Long short-term memory (LSTM) neural network is a kind of deep neural networks, while it possesses similar properties of recurrent neural network (RNN). Therefore, LSTM is a better choice for financial time series forecasting. In addition, the above ensemble learning approach usually chooses AdaBoost to integrate different LSTM forecasters.

In this study, an AdaBoost-based LSTM ensemble learning approach is firstly proposed to forecast financial time series, combining AdaBoost ensemble algorithm and LSTM neural network. LSTM is considered as weak forecasters and AdaBoost is utilized as an ensemble tool. The rest of this paper is organized as follows: the proposed method is briefly introduced in Section 2. Section 3 gives the empirical results and Section 4 provides the conclusions.

2 AdaBoost-LSTM ensemble learning approach

Suppose there is a time series, we would like to make the m-step ahead forecasting. It is noticing that the iterative forecasting strategy is implemented in this paper, which can be expressed as:

$$\hat{x}_{t+m} = f(x_t, x_{t-1}, \dots, x_{t-(p-1)}) \tag{1}$$

where $\hat{x}$ is the forecasting value, x_t is the actual value in period t, and p denotes the lag orders.

In this study, the AdaBoost algorithm is introduced to integrate a set of LSTM predictors. An AdaBoost-LSTM ensemble learning approach is proposed for financial time series forecasting, and the flowchart is illustrated in **Fig. 1**. The proposed AdaBoost-LSTM ensemble learning approach consists of three main steps:

1) The sampling weights $\{D_n^t\}$ of the training samples $\{x_t\}_{t=1}^{T}$ are calculated as follows:

$$D_n^t = \frac{1}{N}, n = (1, 2, \dots, N; t = 1, 2, \dots, T) \tag{2}$$

where N is the number of LSTM predictors, T is the number of training samples.

ICCS Camera Ready Version 2018
To cite this paper please use the final published version:
DOI: 10.1007/978-3-319-93713-7_55

2) The LSTM predictor F_n is trained by the training samples which are sampled according to the sampling weights D_n^t.
3) The foresting error $\{e_n^t\}$ and ensemble weights $\{W_n\}$ of the LSTM predictor F_n are calculated as follows:

$$e_n^t = \frac{|x_i - \hat{x}_i|}{x_i}, (n = 1, 2, \dots, N; t = 1, 2, \dots, T) \tag{3}$$

$$w_n = \frac{1}{2}\ln\left(\frac{1-\sum_{t=1}^{T} e_n^t}{\sum_{t=1}^{T} e_n^t}\right) \tag{4}$$

4) Update the sampling weights $\{D_{n+1}^t\}$ of the training samples $\{x_t\}_{t=1}^{T}$ as follows:

$$D_{n+1}^t = \frac{D_n^t \beta_n^t}{\sum_{t=1}^{T} D_n^t \beta_n^t} \tag{5}$$

where $\beta_n^t = exp(e_n^t)$ is the update rate of training sample x_t .
5) Repeat the step 2-4 until all the LSTM predictors are obtained.
6) The final forecasting result is generated by integrating the forecasting results of all the LSTM predictors with ensemble weights.

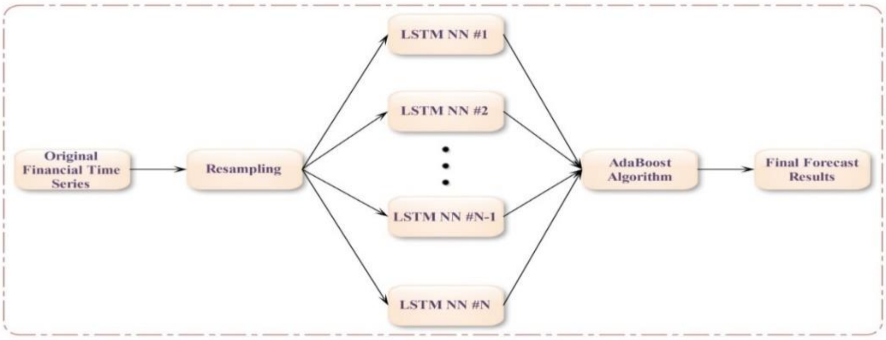

Fig. 1. The flowchart of the AdaBoost-LSTM ensemble learning approach.

3 Empirical study

3.1 Data Description and Evaluation Criteria

The data in this research comprises of two typical stock indexes (S&P 500 index and Shanghai composite index (SHCI)) and two main exchange rates (the euro against the US dollars (EUR/USD) and the US dollars against the China yuan (USD/CNY)). The historical data are daily data, collected from the Wind Database (http://www.wind.com.cn/). The datasets are then divided into in-sample subsets and out-of-sample subsets, as illustrated in **Table 1**.

Table 1. In-sample and out-of-sample dataset of the stock indexes and exchange rates.

Time Series	Sample type	From	To	Sample size
S&P 500	in-sample	January 2, 2015	December 30, 2016	504
	out-of-sample	January 3, 2017	May 31, 2017	103

ICCS Camera Ready Version 2018
To cite this paper please use the final published version:
DOI: 10.1007/978-3-319-93713-7_55

SHCI	in-sample	January 5, 2015	December 30, 2016	488
	out-of-sample	January 3, 2017	May 31, 2017	97
EUR/USD	in-sample	January 1, 2015	December 30, 2016	527
	out-of-sample	January 2, 2017	May 31, 2017	108
USD/CNY	in-sample	January 5, 2015	December 30, 2016	488
	out-of-sample	January 3, 2017	May 31, 2017	97

Table 2 shows the descriptive statistics of those research data. The difference of statistics between the four series can be obviously seen from **Table 2**.

Table 2. The descriptive statistics of the stock indexes and exchange rates.

Time series	Minimum	Maximum	Mean	Std.*	Skewness	Kurtosis
S&P 500	1828.08	2415.82	2123.51	127.69	0.50	2.90
SHCI	2655.66	5166.35	3332.45	488.95	1.71	5.69
EURUSD	1.04	1.21	1.10	0.03	0.09	2.98
USDCNY	6.19	6.96	6.54	0.25	0.15	1.70

Note: Std.* refers to the standard deviation.

In order to evaluate the forecasting performance of the proposed AdaBoost-LSTM ensemble learning approach, mean absolute percentage error (MAPE) and directional symmetry (DS) are employed to evaluate the level forecasting accuracy and directional forecasting accuracy, respectively. MAPE and DS are defined as follows:

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{Y_t-\hat{Y}_t}{Y_t}\right| * 100\% \tag{6}$$

$$DS = \frac{1}{n-1}\sum_{i=2}^{n} d_i \times 100\% \text{ , } d_i = \begin{cases} 1, & (y_i - y_{i-1})(\hat{y}_i - y_{i-1}) \geq 0 \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

where $\hat{y}_i$ is the forecasting value, y_i is the actual value, and n is the number of observation samples.

3.2 Forecasting performance comparison

The forecasting performances of the proposed AdaBoost-LSTM ensemble learning approach and benchmarks are discussed in this section. **Tables 3-6** show the comparison results of MAPE and DS evaluation criteria, which show that the out-of-sample forecasting performance of the proposed approach is better than that of the benchmarks for all of the four financial time series and demonstrates that the proposed approach is an effective tool for financial time series forecasting.

As shown in **Tables 3-6**, the proposed approach significantly outperforms all of the benchmark models by means of level forecasting accuracy and directional forecasting accuracy for the stock indexes and exchange rates. Overall, the ensemble learning approaches outperform the single models, while individual LSTM, ELM, SVR and MLP models consistently outperform ARIMA models in terms of MAPE and DS. Moreover, the proposed AdaBoost-LSTM ensemble learning approach produces 19.44-22.33% better directional forecasts than ARIMA models, reaching up to an accuracy rate of 76.68% in out-of-sample directional forecasting for the EUR/USD.

ICCS Camera Ready Version 2018
To cite this paper please use the final published version:
DOI: 10.1007/978-3-319-93713-7_55

Table 3. Forecasting performance of different models for stock indexes

	Models	S&P 500		SHCI	
		MAPE (%)	DS (%)	MAPE (%)	DS (%)
Single forecasts	ARIMA	4.973	52.43	5.162	51.55
	MLPNN	3.114	63.11	2.661	55.67
	SVR	2.025	66.02	2.126	60.82
	ELM	1.974	66.02	1.024	59.79
	LSTM	1.045	66.99	0.925	62.89
Ensemble forecasts	AdaBoost-MLP	1.023	70.87	0.918	67.01
	AdaBoost-SVR	0.841	72.82	1.106	71.13
	AdaBoost-ELM	0.782	71.84	0.692	70.10
	AdaBoost-LSTM	0.413	74.76	0.312	73.20

Table 4. Forecasting performance of different models for exchange rates series.

	Models	EURUSD		USDCNY	
		MAPE (%)	DS (%)	MAPE (%)	DS (%)
Single forecasts	ARIMA	4.169	57.41	3.973	55.67
	MLPNN	1.973	67.59	2.034	60.82
	SVR	1.164	70.37	1.615	70.10
	ELM	1.035	68.52	0.993	67.01
	LSTM	0.917	69.44	1.024	69.07
Ensemble forecasts	AdaBoost-MLP	0.643	75.00	0.781	71.13
	AdaBoost-SVR	0.534	73.15	0.497	72.16
	AdaBoost-ELM	0.346	74.07	0.268	74.23
	AdaBoost-LSTM	0.172	76.85	0.113	76.29

Table 5. MAPE comparison with different ensemble forecasting approaches.

	Ensemble models	Number of forecasters				
		K=10	K=20	K=30	K=40	K=50
S&P 500	AdaBoost-MLP	1.023	0.993	1.126	1.205	1.021
	AdaBoost-SVR	0.841	0.917	0.864	0.968	0.845
	AdaBoost-ELM	0.782	0.754	0.793	0.801	0.785
	AdaBoost-LSTM	0.413	0.397	0.402	0.419	0.385
SHCI	AdaBoost-MLP	0.918	1.216	1.039	1.114	1.063
	AdaBoost-SVR	1.106	0.987	1.025	1.203	1.287
	AdaBoost-ELM	0.692	0.682	0.705	0.712	0.695
	AdaBoost-LSTM	0.312	0.295	0.323	0.298	0.347
EUR/USD	AdaBoost-MLP	0.643	0.711	0.669	0.683	0.702
	AdaBoost-SVR	0.534	0.602	0.585	0.596	0.562
	AdaBoost-ELM	0.346	0.369	0.401	0.327	0.364
	AdaBoost-LSTM	0.172	0.119	0.187	0.254	0.306
USD/CNY	AdaBoost-MLP	0.781	0.816	0.798	0.833	0.778
	AdaBoost-SVR	0.497	0.506	0.485	0.523	0.519
	AdaBoost-ELM	0.268	0.314	0.296	0.337	0.274
	AdaBoost-LSTM	0.113	0.107	0.235	0.196	0.273

ICCS Camera Ready Version 2018
To cite this paper please use the final published version:
DOI: 10.1007/978-3-319-93713-7_55

Table 6. DS comparison with different ensemble forecasting approaches.

	Ensemble models	Number of forecasters				
		K=10	K=20	K=30	K=40	K=50
S&P 500	AdaBoost-MLP	70.87	69.90	71.84	70.87	69.90
	AdaBoost-SVR	72.82	70.87	70.87	72.82	73.79
	AdaBoost-ELM	71.84	72.82	72.82	73.79	71.84
	AdaBoost-LSTM	74.76	74.76	73.79	74.76	75.73
SHCI	AdaBoost-MLP	67.01	65.98	67.01	64.95	68.04
	AdaBoost-SVR	71.13	69.07	70.10	71.13	69.07
	AdaBoost-ELM	70.10	71.13	69.07	70.10	71.13
	AdaBoost-LSTM	73.20	74.23	72.16	75.26	74.23
EUR/USD	AdaBoost-MLP	75.00	72.22	74.07	73.15	74.07
	AdaBoost-SVR	73.15	74.07	75.00	73.15	72.22
	AdaBoost-ELM	74.07	73.15	73.15	74.07	76.85
	AdaBoost-LSTM	76.85	75.93	76.85	75.93	77.78
USD/CNY	AdaBoost-MLP	71.13	70.10	72.16	70.10	74.23
	AdaBoost-SVR	72.16	75.26	73.20	74.23	74.23
	AdaBoost-ELM	74.23	73.20	74.23	72.16	72.16
	AdaBoost-LSTM	76.29	77.32	75.26	75.26	76.29

In summary, some interesting findings can be summarized: (1) the proposed AdaBoost-LSTM outperforms all of the benchmark models in different forecasting horizons, which implies that the AdaBoost-LSTM ensemble learning approach is a powerful learning approach for financial time series forecasting in both level accuracy and directional accuracy; (2) it clearly shows that the hybrid ensemble approach with AdaBoost is much better than the one without ensemble by means of level accuracy and directional accuracy, which reveals that AdaBoost is a more effective ensemble algorithm; (3) the forecasting performance of hybrid ensemble learning approaches are significantly better than single models. The possible reason is that the ensemble can dramatically improve the forecasting performance of single models.

4 Conclusions

This paper proposes an AdaBoost-LSTM ensemble learning approach which employs AdaBoost algorithm to integrate the forecasting results of LSTM forecasts. Then, the proposed AdaBoost-LSTM ensemble learning approach is applied to forecast financial time series, including stock indexes and exchange rates. For model evaluation and model comparison, four typical financial time series data are collected to test the model performance. The empirical results show that the proposed AdaBoost-LSTM ensemble learning approach can significantly improve forecasting performance and outperform some other single forecasting models and some other ensemble learning approaches in terms of both level forecasting accuracy and directional forecasting accuracy, which demonstrates that the proposed approach is really a promising approach for financial time series forecasting. What's more, the proposed approach can also be employed to solve other complex time series forecasting problems, such as crude oil price forecasting, wind speed forecasting, traffic flow forecasting, etc.

ICCS Camera Ready Version 2018
To cite this paper please use the final published version:
DOI: 10.1007/978-3-319-93713-7_55

References

1. Chortareas G, Jiang Y, Nankervis JC: Forecasting exchange rate volatility using high-frequency data: Is the euro different? International Journal of Forecasting 27(4):1089-1107 (2011).
2. Carriero A, Kapetanios G, Marcellino M: Forecasting exchange rates with a large Bayesian VAR. International Journal of Forecasting 25(2):400-417(2009).
3. Moosa IA, Vaz JJ: Cointegration, error correction and exchange rate forecasting. Journal of International Financial Markets, Institutions and Money 44:21-34 (2016).
4. Galeshchuk S: Neural networks performance in exchange rate prediction. Neurocomputing 172:446-452 (2016).
5. Zhang G, Hu MY (1998) Neural network forecasting of the British pound/US dollar exchange rate. Omega 26(4):495-506.
6. Huang S, Chuang P, Wu C, Lai H: Chaos-based support vector regressions for exchange rate forecasting. Expert Syst Appl 37(12):8590-8598 (2010).
7. Shen F, Chao J, Zhao J: Forecasting exchange rate using deep belief networks and conjugate gradient method. Neurocomputing 167:243-253 (2015).
8. Chen A, Leung MT: Regression neural network for error correction in foreign exchange forecasting and trading. Comput Oper Res 31(7):1049-1068 (2004).
9. Nag AK, Mitra A: Forecasting daily foreign exchange rates using genetically optimized neural networks. Journal of Forecasting 21(7):501-511 (2002).
10. Sermpinis G, Stasinakis C, Theofilatos K, Karathanasopoulos A: Modeling, forecasting and trading the EUR exchange rates with hybrid rolling genetic algorithms—Support vector regression forecast combinations. Eur J Oper Res 247(3):831-846(2015).
11. Sermpinis G, Theofilatos K, Karathanasopoulos A, Georgopoulos EF, Dunis C: Forecasting foreign exchange rates with adaptive neural networks using radial-basis functions and particle swarm optimization. Eur J Oper Res 225(3):528-540 (2013).
12. Yu L, Wang S, Lai KK: A novel nonlinear ensemble forecasting model incorporating GLAR and ANN for foreign exchange rates. Comput Oper Res 32(10):2523-2541 (2005).
13. Yu L, Wang S, Lai KK: Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm. Energy Economics 30(5):2623-2635 (2008).
14. Plakandaras V, Papadimitriou T, Gogas P: Forecasting daily and monthly exchange rates with machine learning techniques. Journal of Forecasting 34(7):560-573 (2015).
15. Yu L, Wang S, Lai KK: A neural-network-based nonlinear metamodeling approach to financial time series forecasting. Appl Soft Comput 9(2):563-574 (2009).
16. Yu L, Wang Z, Tang L: A decomposition-ensemble model with data-characteristic-driven reconstruction for crude oil price forecasting. Appl Energ 156:251-267 (2015).
17. Tang L, Yu L, Wang S, Li J, Wang S: A novel hybrid ensemble learning paradigm for nuclear energy consumption forecasting. Appl Energ 93:432-443 (2012).
18. Niu M, Wang Y, Sun S, Li Y: A novel hybrid decomposition-and-ensemble model based on CEEMD and GWO for short-term PM2.5 concentration forecasting. Atmos Environ 134:168-180 (2016).

ICCS Camera Ready Version 2018
To cite this paper please use the final published version:
DOI: 10.1007/978-3-319-93713-7_55