UCLA Economics
Master of Quantitative Economics

Research on Quantitative Trading Strategy Based on Support Vector Machine and Neural Network Algorithm

--Take stock index futures as an example

JIE

UCLA Economics
Master of Quantitative Economics

CONTENTS

01

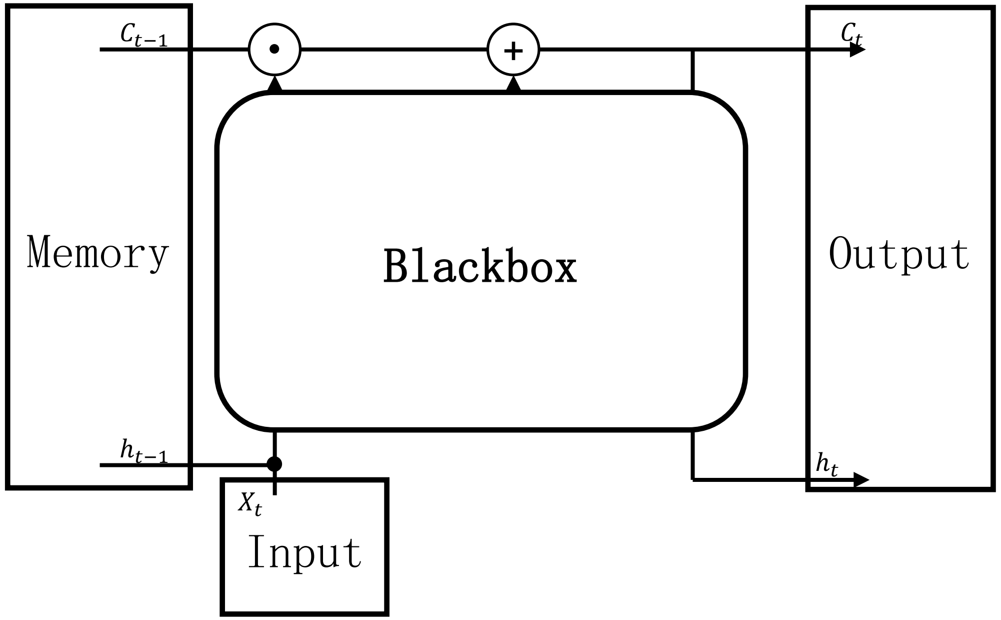
Introduction

Introduction

UCLA Economics
Master of Quantitative Economics

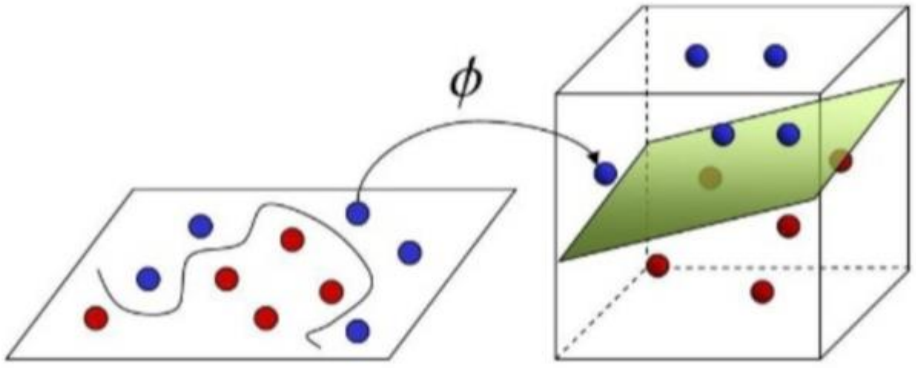
Long short-term memory model(LSTM)

Hochreiter(1997) proposed that linking memory and input in time dimension is suitable for predicting highly dynamic time series data.



Support Vector Machine(SVM)

Support Vector Machine (SVM) uses hyperplane to classify samples. The classifier is only determined by the support vector, which can effectively avoid over-fitting. With excellent nonlinear classification ability, it is widely used in forecasting the rise and fall of stock prices.

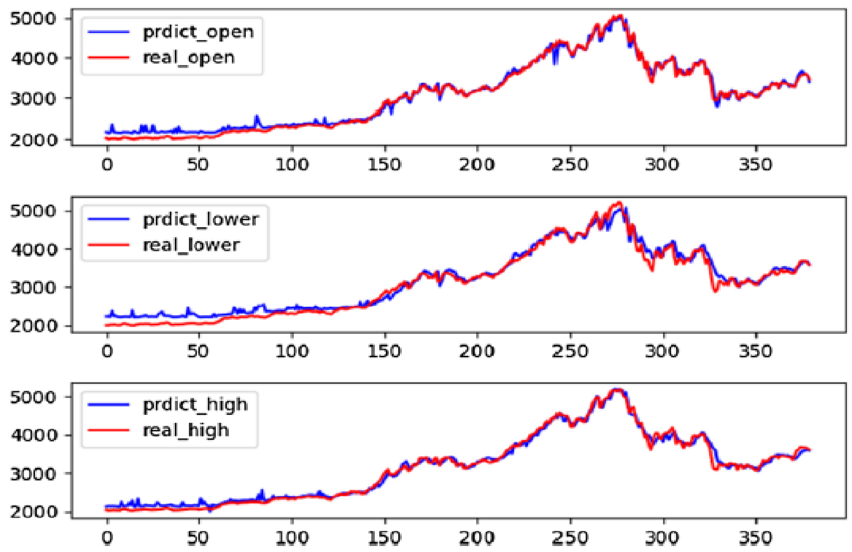


Introduction

UCLA Economics
Master of Quantitative Economics

Long short-term memory model(LSTM)

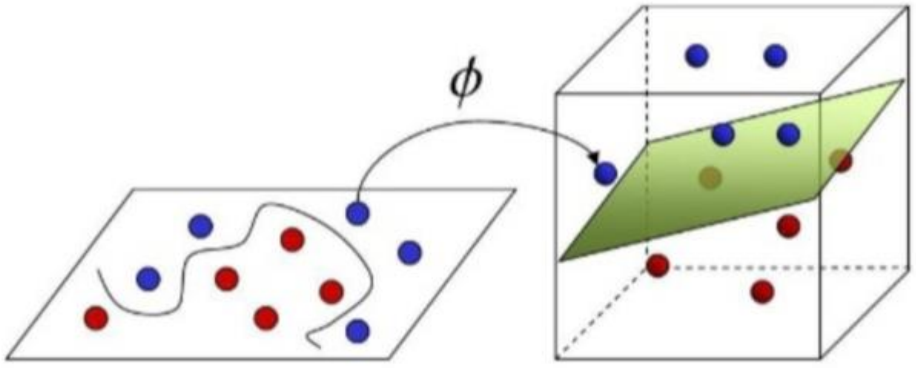
Long short-term memory model (LSTM) has been proved to be an effective method for forecasting time series data, and has been widely used in stock price forecasting.



(G. Ding and L. Qin., 2020)

Support Vector Machine(SVM)

Support Vector Machine (SVM) uses hyperplane to classify samples. The classifier is only determined by the support vector, which can effectively avoid over-fitting. With excellent nonlinear classification ability, it is widely used in forecasting the rise and fall of stock prices.



02

Sample Selection and Data Source

Sample selection and data source

Sample selection: **Shanghai and Shenzhen 300 Index Futures**

Time period: **04/16/2010-03/17/2022(2897 Trade days)**

Robustness test:

Sample selection: **CSI 500 Index Futures and SSE 50 Index Futures**

Time period: **04/16/2015-03/17/2022(1686 Trade days)**

Data source: **WIND**

UCLA Economics
Master of Quantitative Economics

03

Research Design

Index calculation

UCLA Economics
Master of Quantitative Economics

Original Data: **OPEN CLOSE HIGH LOW AMOUNT VOLUME**

Method: **TA-Lib (Technical Analysis Library)**

Summary: 21 technical factors in 6 categories with 42 output variables

Model Evaluation: Average Absolute Error(MAE);

Average Absolute Percentage Error(MAPE); Signal Test

Dependent Variable:

Factor Selection	LSTM	SVM
Return between today and yesterday	CLOSE of next trade day	CLOSE change of next trade day
$Y_i = \frac{CLOSE_i - CLOSE_{i-1}}{CLOSE_{i-1}}$	$Y_i = CLOSE_{i+1}$	$y_t = \begin{cases} 1, s.t. CLOSE_{t+1} > CLOSE_t \\ 0, s.t. CLOSE_{t+1} < CLOSE_t \end{cases}$

Factor Selection

UCLA Economics
Master of Quantitative Economics

Selection Basis: AIC; BIC; LASSO(Least absolute shrinkage and selection operator)

Result: The nine variables or technical indicators selected by LASSO model are as follows

DIF MACD CMO CMOMA ROC STOCH_K WAD MOM Delta1

LASSO

Result-1

Result-2

Prediction of the next day's closing price based on LSTM

Prediction of the next-day closing price of Shanghai and Shenzhen 300 index futures based on LSTM model

MSE	R-Squared	MAE	MAPE	SMAPE	Precision	Recall
12720.92	0.83	83.35	6.2	6.17	0.56	0.3

The next-day closing price of CSI 300 index futures predicted by LSTM model can basically correctly reflect the trend of the real closing price of CSI 300 index futures.

The error between the predicted result based on LSTM model and the real value is about 6.20%. Similarly, according to the calculation method given above, the signal accuracy of model prediction is 56%.

Result:

MSE	R-Squared	MAE	MAPE	SMAPE	Precision	Recall
0.41	-0.65	0.41	/	/	0.59	0.16

As shown in the table above, because the predicted value and the real value are both binary variables of (0,1), the evaluation indicators in statistics can't reflect the quality of the model, my work focuses on the signal evaluation indicators given by the model, namely accuracy and recall rate. Prediction results based on SVM model According to the calculation method given above, the accuracy of signal predicted by the model is 59%.

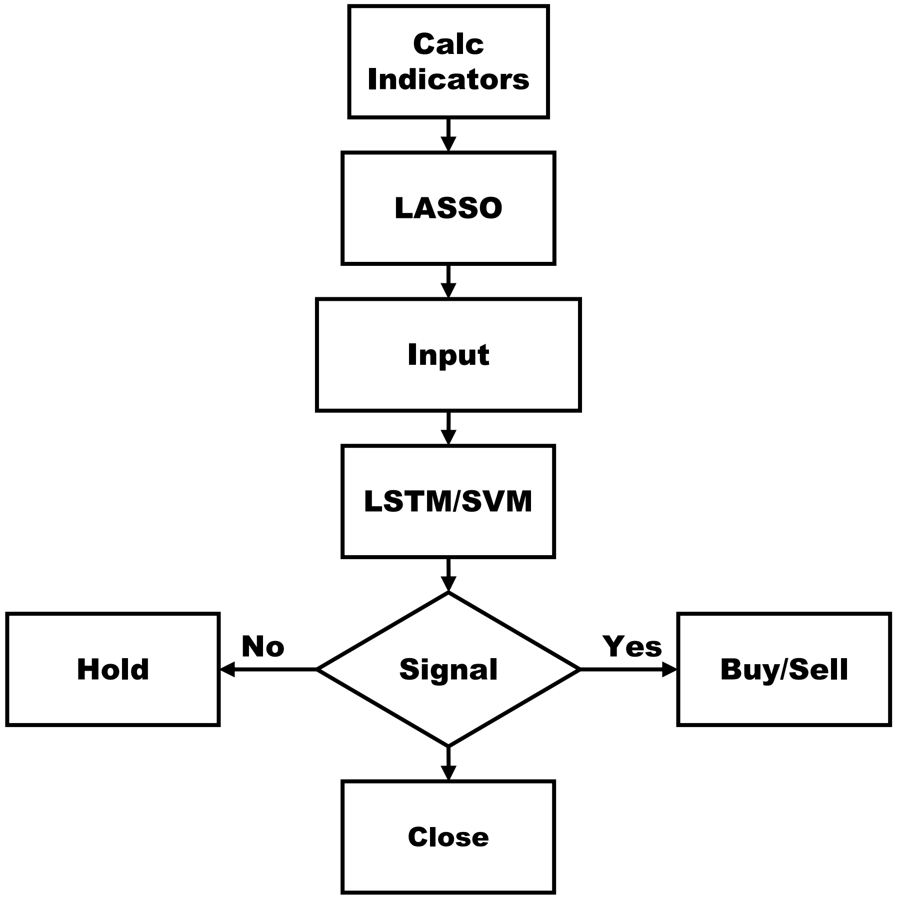
UCLA Economics
Master of Quantitative Economics

04

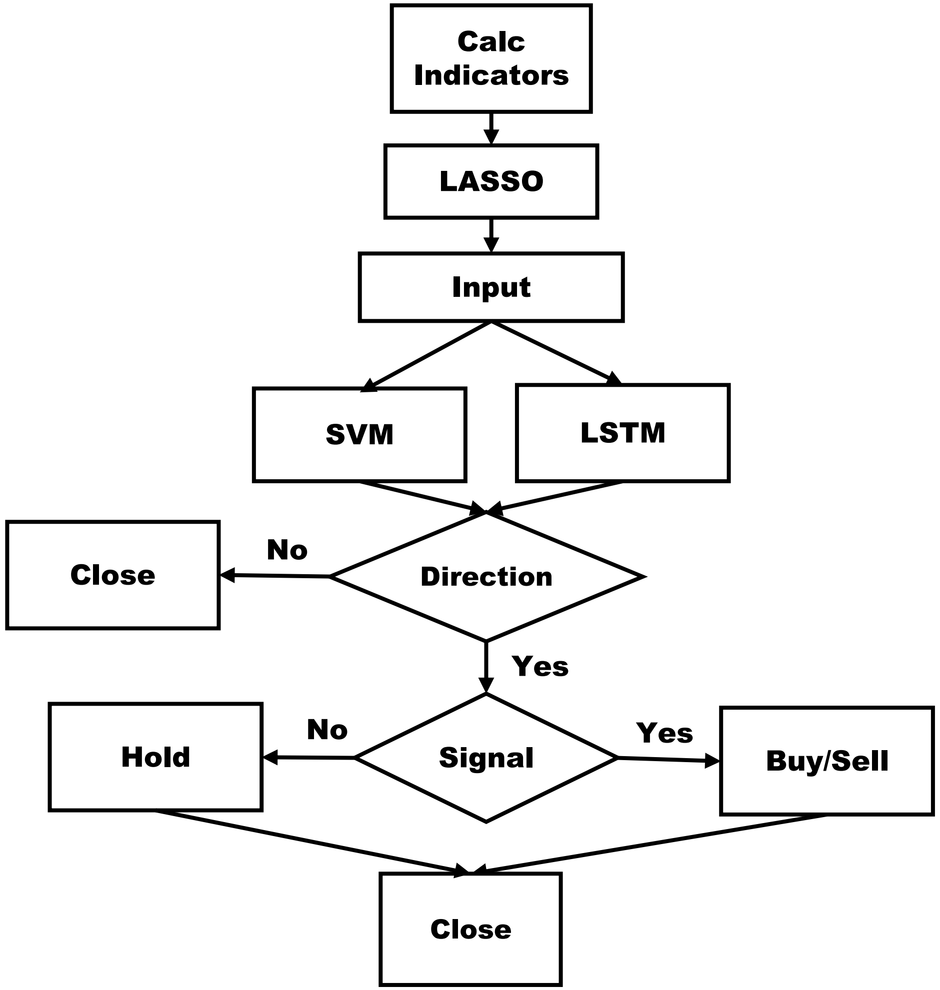
Strategy and Back-test

Strategy

LASSO+LSTM/SVM

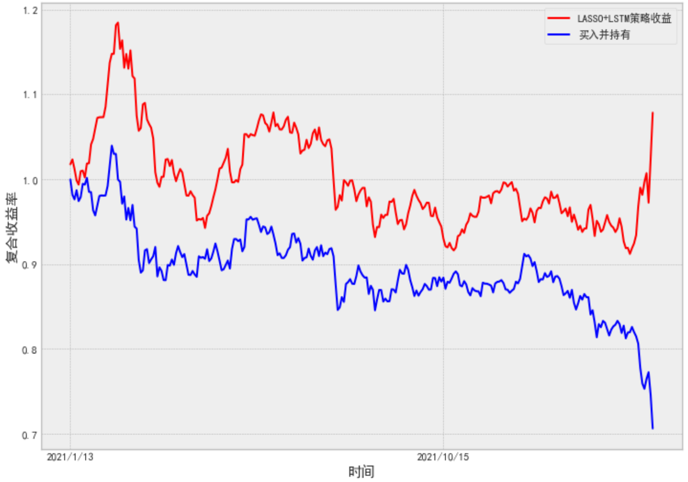


LASSO+LSTM+SVM

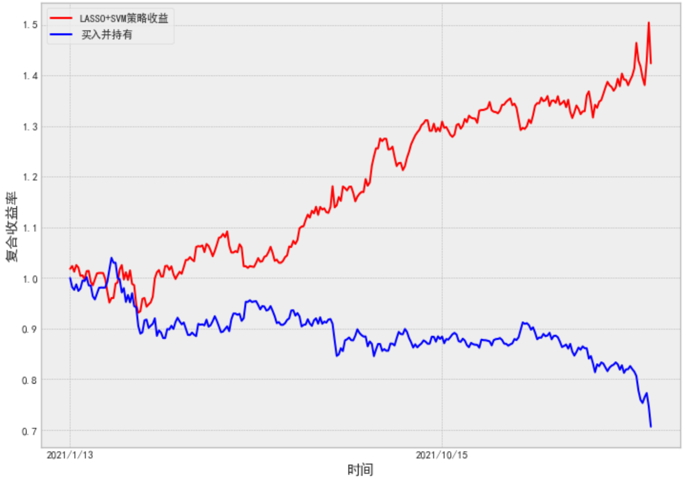


Back-Test On CSI 300 Index Futures

UCLA Economics
Master of Quantitative Economics



LASSO+LSTM
vs
Buy and Hold



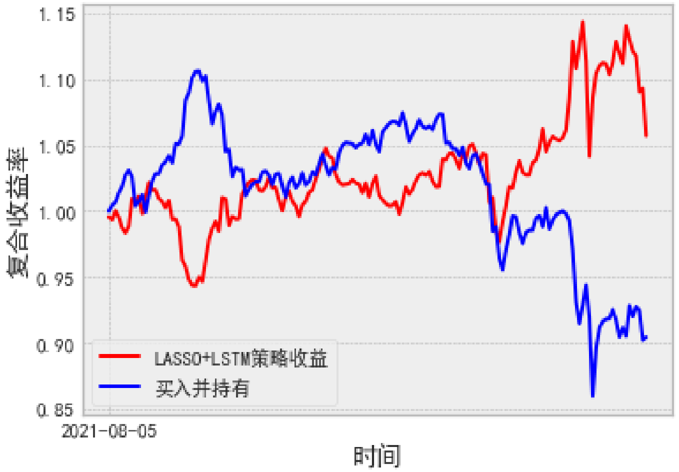
LASSO+SVM
vs
Buy and Hold

LASSO+LSTM+SVM
vs
Buy and Hold

	LASSO+LSTM	LASSO+SVM	LASSO+LSTM+SVM
Annual excess rate of return	8.77%	33.58%	21.34%
Annual volatility	0.2008	0.1998	0.1506
Sharpe ratio	0.44	1.68	1.42
IR	0.4365	1.6806	1.4174
Number of transactions	155	85	60

Back-Test On CSI 500 Index Futures

UCLA Economics
Master of Quantitative Economics



LASSO+LSTM
vs
Buy and Hold

LASSO+SVM
vs
Buy and Hold

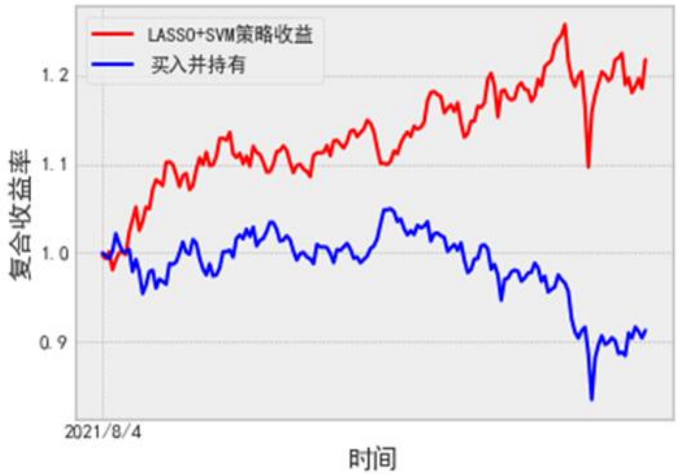
LASSO+LSTM+SVM
vs
Buy and Hold

	LASSO+LSTM	LASSO+SVM	LASSO+LSTM+SVM
Annual excess rate of return	10.53%	25.40%	18.81%
Annual volatility	0.1932	0.1921	0.1133
Sharpe ratio	0.55	1.32	1.66
IR	0.5448	1.3223	1.6602
Number of transactions	71	59	32

Back-Test On SSE 50 Index Futures

UCLA Economics
Master of Quantitative Economics

LASSO+LSTM
vs
Buy and Hold



LASSO+SVM
vs
Buy and Hold

LASSO+LSTM+SVM
vs
Buy and Hold

	LASSO+LSTM	LASSO+SVM	LASSO+LSTM+SVM
Annual excess rate of return	26.52%	32.51%	18.13%
Annual volatility	0.2059	0.205	0.1713
Sharpe ratio	1.29	1.59	1.06
IR	1.288	1.586	1.0583
Number of transactions	87	59	45

UCLA Economics
Master of Quantitative Economics

05

Conclusion

Conclusion and Deficiency

UCLA Economics
Master of Quantitative Economics

- The strategy can obtain high excess returns while keeping low excess volatility, and the quantitative trading strategy based on LASSO+LSTM+SVM model can give consideration to the stability of both high returns and low volatility.
- The accuracy of the trained signal is low.
- Parameter optimization of LSTM and SVM.
- Transaction cost

Thanks for listening

UCLA Economics
Master of Quantitative Economics

06

Appendix

References

UCLA Economics
Master of Quantitative Economics

[1]T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. Journal of econometrics, 31(3):307–327, 1986.

[2]Q. Cao, K. B. Leggio, and M. J. Schniederjans. A comparison between fama and french's model and artificial neural networks in predicting the chinese stock market. Computers & Operations Research, 32(10):2499–2512, 2005.

[3]K. Chen, Y. Zhou, and F. Dai. A lstm-based method for stock returns prediction: A case study of china stock market. In 2015 IEEE international conference on big data (big data), pages 2823–2824. IEEE, 2015.

[4]Y. Chen, S. Mabu, K. Hirasawa, and J. Hu. Trading rules on stock markets using genetic network programming with sarsa learning. In Proceedings of the 9th annual conference on Genetic and evolutionary computation, pages 1503–1503, 2007.

[5]G. Ding and L. Qin. Study on the prediction of stock price based on the associated network model of lstm. International Journal of Machine Learning and Cybernetics, 11(6):1307–1317, 2020.

[6]G. Grudnitski and L. Osburn. Forecasting s&p and gold futures prices: An application of neural networks. Journal of Futures Markets, 13(6):631–643, 1993.

[7]S. Hochreiter and J. Schmidhuber. Long short-term memory. Neural computation, 9(8):1735– 1780, 1997.

[8]W. Huang, Y. Nakamori, and S.-Y. Wang. Forecasting stock market movement direction with support vector machine. Computers & operations research, 32(10):2513–2522, 2005.

[9]M. A. Kaboudan. Genetic programming prediction of stock prices. Computational Economics, 16(3):207–236, 2000.

[10]M. Kaucic. Predicting eu energy industry excess returns on eu market index via a constrained genetic algorithm. Computational Economics, 34(2):173–193, 2009.

[11]K.-j. Kim. Financial time series forecasting using support vector machines. Neurocomputing, 55 (1-2):307–319, 2003.

[12]T. Kimoto, K. Asakawa, M. Yoda, and M. Takeoka. Stock market prediction system with modular neural networks. In 1990 IJCNN international joint conference on neural networks, pages 1–6. IEEE, 1990.

[13]D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980, 2014.

[14]H. Li, Y. Shen, and Y. Zhu. Stock price prediction using attention-based multi-input lstm. In Asian conference on machine learning, pages 454–469. PMLR, 2018.

[15]Z. Liao and J. Wang. Forecasting model of global stock index by stochastic time effective neural network. Expert Systems with Applications, 37(1):834–841, 2010.

[16]I. Matsuba. Application of neural sequential associator to long-term stock price prediction. In [Proceedings] 1991 IEEE International Joint Conference on Neural Networks, pages 1196–1201. IEEE, 1991.

[17]D. Wei. Prediction of stock price based on lstm neural network. In 2019 International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM), pages 544–547. IEEE, 2019.

[18]H. White. Economic prediction using neural networks: The case of ibm daily stock returns. In ICNN, volume 2, pages 451–458, 1988.

[19]Q. Zhuge, L. Xu, and G. Zhang. Lstm neural network with emotional analysis for prediction of stock price. Engineering letters, 25(2), 2017.

[20]邓晓卫 and 章铖斌. 基于混合神经网络模型预测下的统计套利研究. 统计与决策, 1, 2019.3

Indicators

Index type	Name	Categories	Output
Average line	MACD, DEA, DIF, EMA, SMA, WMA, TRIX	7	10
Trend	DEMA, ADX, AROON, CCI	4	8
Momentum	RSI, ATR, ROC, MOM, CMO, MFI	6	8
Turnover	OBV	1	1
Channel	STOCH, Williams AD	1	4
Other	SAR, DELTA	2	11
Total	——	21	42

Model Evaluation

UCLA Economics
Master of Quantitative Economics

Method	Calculation
MSE	$MSE = \frac{\sum(Y_i - \widehat{Y}_i)^2}{T}$
MAE	$MAE = \sqrt{\frac{\sum\lvert Y_i - \widehat{Y}_i\rvert}{T}}$
MAPE	$MAPE = \frac{100\%}{T}\sum\left\lvert\frac{Y_i - \widehat{Y}_i}{Y_i}\right\rvert$
SMAPE	$SMAPE = \frac{100\%}{T}\sum\frac{\lvert Y_i - \widehat{Y}_i\rvert}{(\lvert Y_i\rvert + \lvert\widehat{Y}_i\rvert)/2}$

		Prediction			
		Buy	Hold	Sell	
Real	Buy	$n_{b,b}$	$n_{b,h}$	$n_{b,s}$	$N_{b,.}$
	Hold	$n_{h,b}$	$n_{h,h}$	$n_{h,s}$	$N_{h,.}$
	Sell	$n_{s,b}$	$n_{s,h}$	$n_{s,s}$	$N_{s,.}$
		$N_{.,b}$	$N_{.,h}$	$N_{.,s}$	N

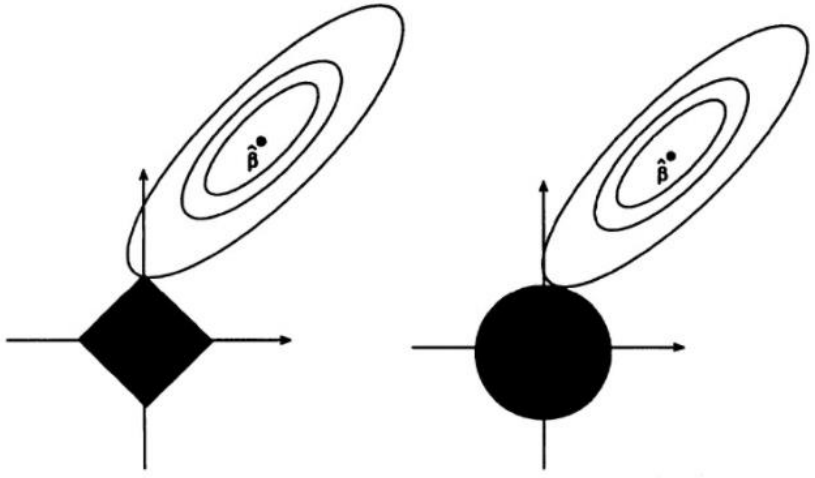
$$Prec = \frac{n_{s,s} + n_{b,b}}{N_{.,s} + N_{.,b}}$$

$$\text{Rec} = \frac{n_{s,s}+n_{b,b}}{N_{s,.}+N_{b,.}}$$

LASSO

Least absolute shrinkage and selection operator

The degree of regression complexity adjustment is controlled by the parameter λ. The larger λ is, the greater the punishment will be to the linear model with more variables, so as to finally obtain a representative variable comb



(a) (b)

$$Lasso: |\beta_1| + |\beta_2| \leq s$$

$$Ridge: \beta_1^2 + \beta_2^2 \leq s$$

Result-1

UCLA Economics
Master of Quantitative Economics

	AIC	BIC	LASSO
Selection	Step by step forward and backward, with AIC as the criterion.	Step by step forward and backward, with BIC as the criterion.	Take the minimum λ plus the maximum value in the range of a standard error Se as the criterion.
Multicollinearity	VIF = 492902, AIC = -1171.26	VIF = ∞, AIC = -1030.39	VIF = 6.91
In Sample			
MSE	71.87	71.11	5.79
MAE	6.54	6.51	0.0053
MAPE	6409.09	6520.69	∞
SMAPE	172.23	171.82	83.37
Accuracy	0.51	0.47	0.73
Retrieval rate	0.07	0.06	0.55
Variables	27	25	9

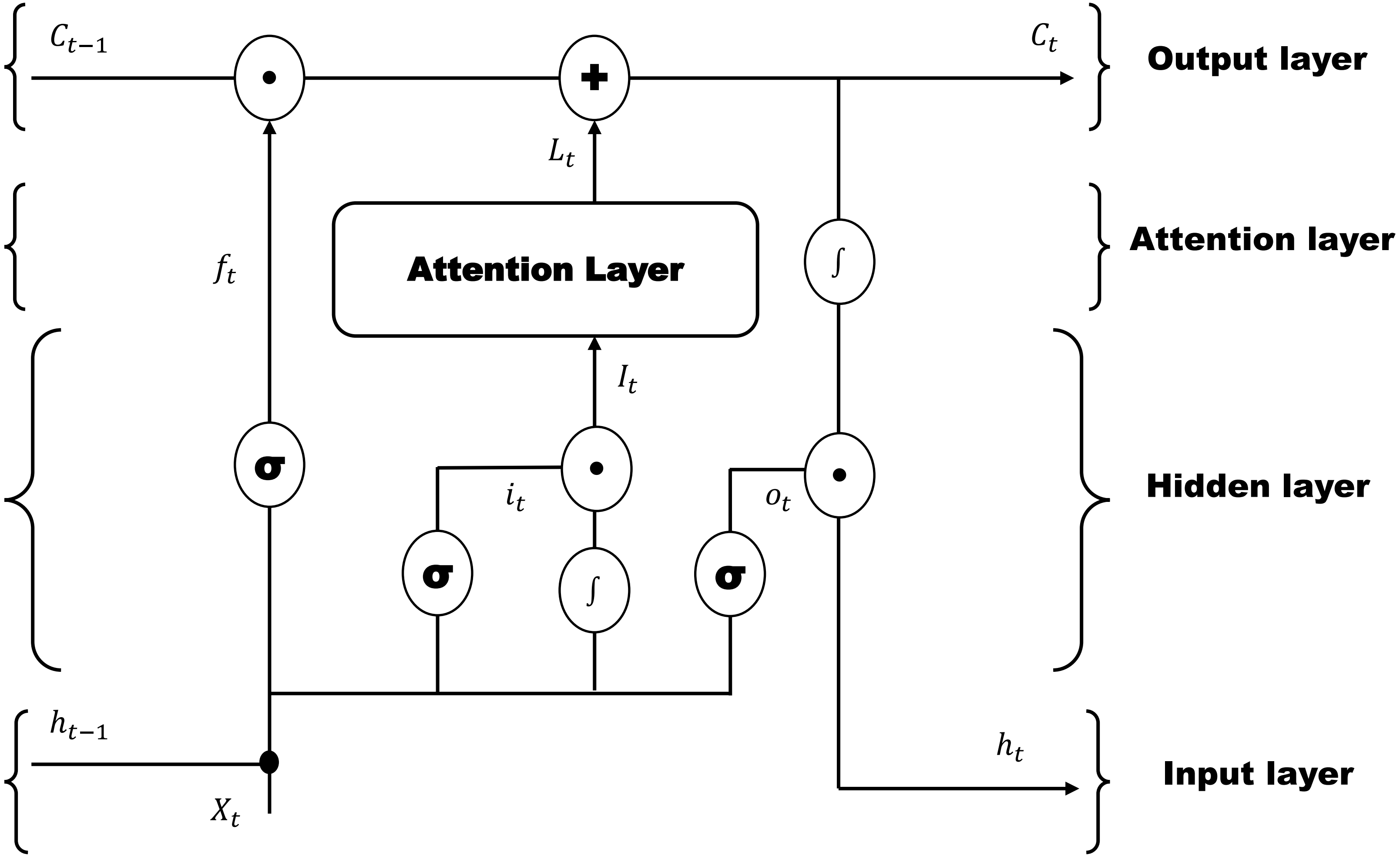
Result-2

UCLA Economics
Master of Quantitative Economics

Name	Calculation	Remark
DIF	$DIF = EMA(CLOSE,12) - EMA(CLOSE,26)$	
MACD	$MACD = 2 * (DIF - DEA)$	
CMO	$CMO = \frac{(SU - SD) * 100}{(SU + SD)}$	**SU SD is the sum of the absolute value of the closing price difference between the rising day and the falling day in 10 days, respectively.**
CMOMA	$CMOMA = MA(CMO,3)$	
ROC	$ROC_i = CLOSE_i - CLOSE_{i-20}$	
STOCH_K	$K_i = \frac{2}{3}K_{i-1} + \frac{1}{3}RSI_i$	
WAD	$WAD_I = A/D_i + WAD_{i-1}$	
MOM	$MOM_i = CLOSE_i - CLOSE_{i-10}$	
Delta1	$Delta_1 = CLOSE_{i-1}$	

LSTM

UCLA Economics
Master of Quantitative Economics



SVM

UCLA Economics
Master of Quantitative Economics

For the input training data set $X = (X_1, X_2,...,X_T)$,Let $\phi(X)$ represent the feature vector of X to be mapped, so in the high-dimensional feature space, the equation corresponding to the division of hyperplane can be expressed as:

$$f(x) = W^T\phi(X) + b$$

Therefore, the input dependent variable $\mathrm{y} = (\mathrm{y}_1, \mathrm{y}_2,...,\mathrm{y}_T)$,can be obtained, so there is a minimization function:

$$\min_{w,b} \frac{1}{2}\|W\|^2, \qquad s.t. \quad y_i(W^T\phi(X) + b) \geq 1 \qquad (i = 1,2, \dots T)$$

The dual problem is:

$$\max_{\alpha} \sum_{i=1}^{T} \alpha_i - \frac{1}{2}\sum_{i=1}^{T}\sum_{j=1}^{T} \alpha_i\alpha_j y_i y_j\, \phi(X_i)^T\phi\left(X_j\right)$$

$$s.t. \quad \sum_{i=1}^{T} \alpha_i y_i = 0, \ \ \alpha_i \geq 0, \qquad i = 1,2, \dots, T$$

Because it is difficult to calculate the inner product of $\phi(X_i)^T\phi\left(X_j\right)$ after the samples X_i and X_j are mapped to the high-dimensional feature space, the inner product after the samples X_i and X_j are mapped to the feature space can be approximately calculated by their function values calculated by the function $K\left(X_i, X_j\right)$ in the original sample space, so:

$$\max_{\alpha} \sum_{Ti=1}^{T} \alpha_i - \frac{1}{2}\sum_{i=1}^{T}\sum_{j=1}^{T} \alpha_i\alpha_j y_i y_j\, K\left(X_i, X_j\right)$$

$$s.t. \quad \sum_{i=1} \alpha_i y_i = 0, \ \ \alpha_i \geq 0, \qquad i = 1,2, \dots, T$$

SVM

Solve the previous equation to the partition hyperplane:

$$f(x) = W^T\phi(X) + b$$

The solved equation to the partition hyperplane:

$$f(x) = W^T\phi(X) + b = \sum_{i=1}^{T} \alpha_i y_i\, \phi(X_i)^T \phi(X_j) + b$$
$$= \sum_{i=1}^{T} \alpha_i y_i\, K(X_i, X_j) + b$$