

以特徵挑選與球型複數類神經模糊系統 進行時間序列預測

Time Series Prediction with Feature Selection and Sphere Complex Neuro-fuzzy System

1. 緒論

表1：股票預測文獻彙整

作者 (年)	資料集來源	輸出目標	取樣期間	方法	表現指標
Li et al. (2014)	SSEC, NASDAQ	Stock price	2011-2012	SVM, EMD	RMSE, MAE, MAPE
Xi et al. (2014)	Chongqing Iron & Steel	Stock price	01.04.2012- 10.08.2012	RBF	RMSE
Bas et al. (2015)	BIST, TAIFEX	Stock price	10. 01.2010- 12. 23.2010	FFANN	RMSE
Ye and Wei (2015)	SSEC	Stock price	2012 -2014	WNN	RMSE, MAPE
Khuat et al. (2016)	Apple, Yahoo, Google	Stock price	2009-2013 2013-2014 2014-2015.	MLP	RMSE
Qiu and Song (2016)	Nikkei 225	Stock price	2007-2013	GA-ANN	Hit ratio
Chen et al. (2016)	TAIEX HSI	Stock price	1998-2006	ANFIS-based	RMSE, Wilcoxon test, Profitable unit
Zhang et al. (2017)	SSEC, TAIEX	Stock price	2000-2006 1990-1999	Type-2 FTS	RMSE, MAPE
Wei et al. (2017)	SSEC	Stock price	2009-2014	2RS-WNN	RMSE, MAD, MAPE, DS%
Chong et al. (2017)	KOSPI	Stock return	2010-2014	DNN	NMSE, RMSE, MAE, MI
Liu et al. (2017)	000573: Shenzhen	Stock volatility	2015-2016	RNNs	Accuracy
Chatzis et al. (2018)	39 Countries	Stock direction	1996-2017	LogR, RF, SVMs, NNs, CART,	Accuracy

				XG- Boost, MXNET	
Pang et al. (2018)	SHASHR, TMSE, TMBA, SINOPEC	Stock price	2006 -2016	ALSTM, ELSTM	MSE, DA
Lei (2018)	SSEC, All Ords, CSI 300, Nikkei 225, DJI	Stock price, Stock direction	2009-2014	BP-NN, RBF-NNAN, FIS-NN, SVM, WNN, RS-WNN, 2RS-WNN	RMSE, MAD, MAPE, DS%, CP%, CD%
Shastri et al. (2019)	Apple	Stock price	2013 - 2016	ANN	MAPE, Accuracy

3. 研究方法

表2：單目標預測資料矩陣

Candidate features					Target
$\left[\begin{array}{ccccc} f_1 & f_2 & \dots & f_s & \\ x_1 & x_2 & \dots & x_s & \\ x_2 & x_3 & \dots & x_{s+1} & \\ \vdots & \vdots & \dots & \vdots & \end{array} \right]$					$\left[\begin{array}{c} t \\ x_{s+1} \\ x_{s+2} \\ \vdots \end{array} \right]$

其中 s 為使用者設定的候選特徵的個數， f 為候選特徵變數， t 為欲進行預測的目標變數， x 為原始數據進行差分後的值。

表3：多目標預測資料矩陣

		Candidate features								Targets			
		$f_1 \quad f_2 \quad \dots \quad f_s \quad f_{s+1} \quad f_{s+2} \quad \dots \quad f_{s* TS }$								$t^{(1)} \quad t^{(2)} \quad \dots \quad t^{(TS)}$			
Data matrix		$x_1^{(1)}$	$x_2^{(1)}$	\dots	$x_s^{(1)}$	$x_1^{(2)}$	$x_2^{(2)}$	\dots	$x_s^{(TS)}$	$x_{s+1}^{(1)}$	$x_{s+1}^{(2)}$	\dots	$x_{s+1}^{(TS)}$
		$x_2^{(1)}$	$x_3^{(1)}$	\dots	$x_{s+1}^{(1)}$	$x_2^{(2)}$	$x_3^{(2)}$	\dots	$x_{s+1}^{(TS)}$	$x_{s+2}^{(1)}$	$x_{s+2}^{(2)}$	\dots	$x_{s+2}^{(TS)}$
		\vdots	\vdots	\dots	\vdots	\vdots	\vdots	\dots	\vdots	\vdots	\vdots	\dots	\vdots

其中目標變數的集合以 TS 表示之，且 $TS = \{t^{(j)}, j = 1, 2, \dots, |TS|\}$ ， $|TS|$ 為目標變數的個數。這些候選特徵的集合則被稱作候選特徵池（Candidate feature pool），以 CP 標記之，且 $CP = \{f_i, i = 1, 2, \dots, |CP|\}$ ， $|CP|$ 為所有候選特徵變數的個數。

表4：影響資訊矩陣

		f_1	f_2	\dots	$f_{ CP }$	$t^{(j)}$
f_1		0	$I_{f_1 \rightarrow f_2}$	\dots	$I_{f_1 \rightarrow f_{ CP }}$	$I_{f_1 \rightarrow t^{(j)}}$
f_2		$I_{f_2 \rightarrow f_1}$	0	\dots	$I_{f_2 \rightarrow f_{ CP }}$	$I_{f_2 \rightarrow t^{(j)}}$
\vdots		\vdots	\vdots	0	\vdots	\vdots
$f_{s* TS }$		$I_{f_{ CP } \rightarrow f_1}$	$I_{f_{ CP } \rightarrow f_2}$	\dots	0	$I_{f_{ CP } \rightarrow t^{(j)}}$
$t^{(j)}$		$I_{t^{(j)} \rightarrow f_1}$	$I_{t^{(j)} \rightarrow f_2}$	\dots	$I_{t^{(j)} \rightarrow f_{ CP }}$	0

其中 $t^{(j)}$ 為第 j 個目標變數，且 $j=1, 2, \dots, |TS|$ 。

4. 實驗內容

4.1 特徵的擷取與影響

表5：單目標預測特徵多寡之效能比較（RMSE）

	RMSE				
	2 Features	4 Features	6 Features	8 Features	Average
Random selection	46.0499	42.7427	44.6477	60.6420	48.5206
Ranking selection	45.0909	40.2567	42.5588	40.3005	42.0517

表6：四目標預測特徵多寡之效能比較（RMSE）

		RMSE				
		SSEC	HSI	Nikkei 225	DJI	Average
2 Features	Random selection	37.6763	356.2660	238.3009	276.9093	227.2881
	Ranking selection	39.5780	349.6290	218.3735	285.6585	223.3098
4 Features	Random selection	35.2426	479.6879	229.3043	283.7045	256.9848
	Ranking selection	36.9934	350.8079	214.9728	276.6222	219.8491
6 Features	Random selection	36.4123	332.8108	217.3304	272.3843	214.7345
	Ranking selection	35.1617	320.4717	206.9819	278.2265	210.2105
8 Features	Random selection	41.8079	330.8052	206.0468	343.5975	230.5644
	Ranking selection	36.3643	306.8784	211.2294	276.0410	207.6282

4.2 中國股市預測與比較

表7：實驗 2 單目標預測實驗設定

	2000	2001	2002	2003	2004	2005	2006
Number of original data	259	245	260	260	261	259	240
Number of data pairs	229	215	230	230	231	229	210
Number of training data	191	179	191	191	192	191	175
Number of rules	3	3	3	3	3	3	3
Number of premise parameters	24	10	36	40	38	6	24
Number of consequent parameters	33	9	33	33	33	15	33

表8：實驗 2 SSEC 之效能比較 (RMSE)

	RMSE							Average
	2000	2001	2002	2003	2004	2005	2006	
Huang and Yu (2005)	23.9147	31.9274	31.9575	21.9938	21.7138	14.6053	75.0643	31.5967
Cheng et al. (2008)	29.4617	33.5855	33.4515	21.6367	32.0092	12.3227	64.0943	32.3659
Chen (2002)	40.768	43.009	57.6315	32.2600	28.4259	16.4664	62.6612	40.1746
Lee et al. (2006)	30.5366	48.4292	45.2494	24.1420	22.3151	12.0581	82.0055	37.8194
Egrioglu et al. (2011)	17.9911	24.0736	26.3361	18.1261	12.5963	5.9938	114.9601	31.4396
Wang et al. (2013)	43.0975	34.0014	26.4196	17.8860	20.1084	11.8674	379.5415	75.9888
Bas et al. (2015)	35.1766	55.1909	55.0887	66.6560	37.5188	27.9020	221.1243	71.3955
Yolcu et al. (2016)	34.0485	51.7665	56.8118	65.4207	33.7176	24.0424	226.9612	70.3955
Zhang et al. (2017)	16.2662	20.3227	18.0470	17.7821	13.7292	9.0226	36.5687	18.8198
Proposed method	17.3292	22.1362	15.4426	12.467	15.4890	13.8235	22.7848	16.8633

4.3 中國與國際市場的相互作用

表9：實驗 3-1 四目標預測實驗設定

	2002	2003
Number of original data	227	226
Number of data pairs	197	196
Number of training data	158	163
Number of rules	3	3
Number of premise parameters	36	32
Number of consequent parameters	33	33

表10：實驗 3-1 四目標預測之效能比較（RMSE）

		RMSE			
		Hsieh et al. (2011)	Chen et al. (2016)	Zhang et al. (2017)	Proposed method
2002	SSEC	-	-	18.0470	20.0222
	HSI	-	118.27	-	105.8136
	Nikkei 225	141	-	-	136.6241
	DJI	132	-	-	129.2588
2003	SSEC	-	-	17.7821	15.1384
	HSI	-	132.67	-	101.6561
	Nikkei 225	177	-	-	127.4672
	DJI	89	-	-	75.5405

表11：模型平均之效能比較（RMSE）

		RMSE			
		Hsieh et al. (2011)	Chen et al. (2016)	Zhang et al. (2017)	Proposed method
	SSEC	-	-	17.9146	17.5803
	HSI	-	125.4700	-	103.7349
	Nikkei 225	159.0000	-	-	132.0457
	DJI	110.5000	-	-	102.3997

表12：實驗 3-2 單目標預測實驗設定

	SSEC	HSI	Nikkei 225	DJI
Number of original data	242	245	245	249
Number of data pairs	212	215	215	219
Number of training data	176	179	179	182
Number of rules	3	3	3	3
Number of premise parameters	10	10	10	10
Number of consequent parameters	9	9	9	9

表13：實驗 3-2 四目標預測實驗設定

Number of original data	217
Number of data pairs	187
Number of training data	156
Number of rules	3
Number of premise parameters	40
Number of consequent parameters	33

表14：單目標與多目標預測之效能比較（RMSE）

	RMSE			
	SSEC	HSI	Nikkei 225	DJI
1 Target	32.6480	332.7996	238.5166	251.3554
4 Targets	31.6853	306.8782	200.4556	245.1469