Supplementary file

For article:

Manuscript Title: A Hybrid Approach for a Novel Dynamic Trading System to Produce Robust Cryptocurrency Portfolios

Published in the journal:

Scientia Iranica, Transaction E: Industrial Engineering

Authors:

Ehsan Mohammadian Amiri, Akbar Esfahanipour*

Department of Industrial Engineering and Management Systems, Amirkabir University of Technology, Tehran, Iran

Introduction

This file contains four parts including 1) Eigenvector Centrality Feature Selection (ECFS) as the feature selection method that we have applied in our proposed system, 2) Long Short-Term Memory (LSTM) model to predict indicators for building the input matrix of the CNN, 3) Results of RMSE for fitting the LSTM model in the LSTM-CNN FS-based model, and 4) Results for variations of LSTM-CNN-FS-ZZ model accuracy as depicted.

S.1 Eigenvector Centrality Feature Selection (ECFS)

The ECFS method is among the methods with a classification-based feature selection filter approach that can rank features without using a specific classification algorithm. This capability is due to the design of its graph-based method, and it enables the use of this algorithm to prepare features for any classification method. Roffo & Melzi [23] have extensively tested their proposed ECFS method on seven different datasets selected from different application

* Corresponding author Email address: esfahaa@aut.ac.ir
Address: No. 350, Hafez Ave, Valiasr Square, Tehran, Iran 1591634311, Department of Industrial Engineering & Management Systems, Amirkabir University of Technology Tel: +98(21)6454-5369; Mobile: +98(912)347-9906; Fax:+98(21)6695-4569; Website: www.aut.ac.ir/esfahaa

areas of object recognition, handwritten recognition, biological data, and synthetic testing datasets. The ECFS method achieved top performances against seven competitor models of Fisher, FSV, Inf-FS, MI, LS, Relief, and REF.

Suppose $x = \left\{x^{(1)},\ldots,x^{(n)}\right\}$ is a set of features, and the graph *G=(V, E)* is an undirected graph, such that *V* is a set of vertices corresponding to each variable *x*, and *E* is a codified (weighted) edges among features. The adjacency matrix *A*, associated with graph *G*, specifies the state of weighted edges. That is, each element a_{ij} of matrix *A* represents a pairwise potential relation. These potential relations can be presented in the form of a binary function $\varphi(x(i),x^{(j)})$ of nodes $x^{(k)}$, as shown in Eq. (15).

$$a_{ij} = \varphi(x^{(i)}, x^{(j)}) \tag{15}$$

Different methods can balance the graph. Therefore, the φ function can be used manually in the method or be automatically trained from the data. Each feature distribution $x^{(i)}$ is normalized to sum to 1. Here, the following process takes place. First, Fisher's criterion is used as Eq. (16).

$$f_i = \frac{|\mu_{i,1} - \mu_{i,2}|^2}{\delta_{i,1}{}^2 + \delta_{i,2}{}^2} \tag{16}$$

Where $\mu_{i,c}$ and $\sigma_{i,c}^2$ are the mean and standard deviation of the *i-th* feature when the samples of the c-th class are considered. A higher value of f_i indicates more resolution of the *i-th* feature. Since there are labels of the classes, it's better to keep only the attributes that belong to these classes (rank). Therefore, mutual information is used to rank superior traits that achieve high features with high-class prediction power as Eq. (17).

$$m_i = \sum_{y \in Y} \sum_{z \in x^{(i)}} p(z,y)\log(\frac{p(z,y)}{p(z)p(y)}) \tag{17}$$

Where *Y* is a collection of class labels and p *is* a joint probability distribution. Then the kernel k is obtained by multiplying the matrices f and m^T as Eq. (18):

$$k = (f.m^T) \tag{18}$$

Where *f* and *m* are normalized n×1 column vectors in the range of 0 and 1, and k is an n×n matrix. To improve the performance, a secondary feature evaluation measure is the standard deviation, which expresses the variability of the features from the average, as shown in Eq. (19).

$$\sum(i,j) = \max(\sigma^{(i)}, \sigma^{(j)}) \tag{19}$$

Where σ is the standard deviation for *x* samples and ∑ is the inverted $n \times n$ matrix with values between *0* and *1*. Finally, the adjacency matrix *A* for graph *G* is obtained as follows:

$$A = \alpha k + (1-\alpha)\Sigma \tag{20}$$

Where α is a coefficient between *0* and *1*. The basic idea behind the ECFS method is to find the eigenvector v_0 of matrix *A* that corresponds to the largest eigenvalue. The values reflect the degree to which each node is in contact with the other nodes. Since the limit of A^l as *l* approaches a large positive number, *L* converges to v_0 as shown by Eq. (21).

$$\lim_{l \to L} \left[A^l{}_e \right] = \upsilon_0 \tag{21}$$

Where *e* does not conform to the principal vector v_0 of *A*. Using eigenvectors as a basis with a coefficient $\beta=0$ for v_0, it is always possible to decompose *e*. Hence we can write the Eqs. (22) and (23).

$$e = \beta_0 \upsilon_{0+} \beta_1 \upsilon_1 + \ldots + \beta_n \upsilon_n \tag{22}$$

Eventually,

$$A^l e = A^l (\beta_0 \upsilon_{0+} \beta_1 \upsilon_1 + \ldots + \beta_n \upsilon_n) = \beta_0 A^l \upsilon_{0+} \beta_1 A^l \upsilon_{1+} \ldots + \beta_n A^l \upsilon_{n=} \beta_0 \lambda_0 \upsilon_{0+} \beta_1 \lambda_1 \upsilon_{1+} \ldots \beta_n \lambda_n \upsilon_n \tag{23}$$

S.2 Long Short-Term Memory (LSTM)

The LSTM network architecture includes a sequence input gate, forget gate, output gate, memory cell candidate, memory cell, and shadow state that are explained as follows:

$$i_t = \sigma(x_i U_i + h_{t-1} W_i + b_i), \quad \text{Input gate} \tag{24}$$

$$f_t = \sigma(x_t Ut + h_{t-1} W_f + b_f), \quad \text{Forget gate} \tag{25}$$

$$o_t = \sigma(x_t U_o + h_{t-1} W_o + b_o), \quad \text{Output gate} \tag{26}$$

$$\tilde{c}_t = \tanh(x_t U_g + h_{t-1} W_g + b_g), \quad \text{Memory cell candidate} \tag{27}$$

$$c_t = f_t \circ c_{t-1} + i_t \circ \tilde{c}_t \quad \text{Memory cell} \tag{28}$$

$$h_t = \tanh(\tilde{c}t) \circ o_t \quad \text{Shadow state} \tag{29}$$

Where x_t is the input vector at time *t*, i_t is the input/update gate's activation vector at time *t*, f_t is the forget gate's activation vector at time *t*, o_t output gate's activation vector at time *t*, h_t is hidden state vector also known as output vector of the LSTM unit at time *t*, $\tilde{c}_t$ is cell input activation vector at time *t*, c_t is cell state vector at time *t*, *Wi*, W_f, W_o, W_g and b_i, b_f , b_o, b_g are weights and biases, respectively. *σ(.)* is the logistic function, and *tanh(.)* is the hyperbolic tangent function.

S.3 Variations of LSTM-CNN-FS-ZZ model accuracy

Table S.1 shows the Variations of LSTM-CNN-FS-ZZ model accuracy (%) with changes of the Zigzag parameter (PPC) in training periods for cryptocurrencies.

Please insert Table S.1 about here.

S.4 RMSE for fitting the LSTM model

Please insert Table S.2 about here.

Table 3 shows the learning ability of the LSTM model in predicting the A/D line indicator with an average RMSE of 0.042, which is better than other indicators. After that, the A/D oscillator and on-balance volume indicators have the least root mean square error with 0.045 and 0.080, respectively. Also, the average values of cryptocurrencies RMSEs is 0.056 that the lowest and highest values of RMSEs belong to ETC and XLM, respectively.

List of Table captions:

Table S.1 Variations of LSTM-CNN-FS-ZZ model accuracy (%) with changes of the Zigzag parameter (PPC) in training periods for cryptocurrencies

Table S.2 RMSE for fitting the LSTM model in the LSTM-CNN FS-based model

Table S.1 Variations of LSTM-CNN-FS-ZZ model accuracy (%) with changes of the Zigzag parameter (PPC) in training periods for cryptocurrencies

	PPC Parameter	Training Period											
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
ADA	**0.02**	72.24655	66.0294	64.44	67.2136	64.0416	76.7766	84.1956	81.2636	77.6247	83.4844	77.4486	78.3064
	0.04	74.51485	63.8121	66.8128	68.5482	70.4127	75.9232	84.6885	82.2515	78.7127	84.6296	80.7272	79.2837
	0.06	75.27445	64.6402	68.6438	71.0507	70.8715	80.4172	84.7821	77.9228	78.9	85.118	81.843	80.2028
	0.08	79.22905	66.5193	75.4918	72.4711	75.6367	82.4792	86.1838	83.9613	81.909	87.6097	79.2553	83.2184
	0.10	80.8163	77.0835	73.0851	76.1289	74.5696	84.5736	86.9388	85.1984	84.7181	85.5776	80.4083	80.5724
	0.12	77.96865	79.3231	76.3249	75.9722	78.3971	85.3061	85.0247	86.1225	86.9388	86.9254	81.6996	82.2129
	0.14	77.94505	78.422	81.4897	77.551	78.8534	79.5888	83.465	85.1176	85.2626	88.9796	82.9118	82.3166
	0.16	73.88965	82.0408	82.5591	75.7738	82.0408	77.1795	82.6103	84.2814	83.7728	87.4853	83.5911	83.3249
	0.18	72.18585	80.9152	83.6735	72.4612	75.8667	74.011	81.4977	82.7485	81.7491	83.6279	84.898	84.0816
BNB	**0.02**	80.757	76.2703	76.8289	54.1913	71.5708	72.8237	70.7999	68.5984	70.1905	69.8815	72.9488	76.0488
	0.04	81.2875	79.5494	81.6122	54.588	75.0307	73.1554	73.821	70.8242	71.1789	69.8515	73.1821	77.3716
	0.06	81.6144	78.3722	83.5992	53.9227	73.7567	75.8058	79.9676	71.9466	72.2625	68.3155	73.1841	77.8614
	0.08	82.9346	82.164	85.3329	56.7527	76.5075	77.1429	80.8493	74.0807	72.0901	70.2645	74.8154	79.9273
	0.10	84.4898	83.1229	85.7143	57.8967	79.6035	76.9644	80.7511	80.1786	74.3893	71.8853	75.6782	80.7331
	0.12	81.9879	84.4898	85.2429	58.038	76.9642	74.032	83.534	84.5659	75.2131	71.7728	81.6211	80.1921
	0.14	81.4836	81.795	78.3534	63.2484	79.3558	73.4491	84.898	84.898	75.5102	72.5586	82.1462	79.6001
	0.16	82.8213	83.0136	78.8643	64.7781	81.2245	72.9464	84.0615	83.228	73.5986	75.9184	83.4731	84.346
	0.18	79.3155	77.2897	78.5941	66.1225	80.2605	73.9305	81.2851	82.1806	69.2018	76.004	84.4898	86.5306
BTC	**0.02**	70.1716	78.4414	71.7297	73.1627	78.6333	78.7369	79.0073	67.8007	66.5611	80.9215	80.9969	77.0013
	0.04	74.2072	77.1752	71.2005	53.9841	79.4693	79.8627	80.1601	68.3323	67.8842	82.7611	83.3484	78.4194
	0.06	74.0859	77.0436	70.3892	57.225	80.0201	80.4801	81.3342	71.7255	68.9638	82.6085	84.8642	78.0996
	0.08	75.4348	77.3376	72.0912	65.2687	81.183	83.6331	81.77	71.4173	70.1994	83.7003	86.5306	79.5918
	0.10	77.0597	79.04	72.4993	69.6528	81.6743	82.7481	82.6473	72.8745	70.0101	87.3469	85.6392	78.0859
	0.12	79.8773	82.4983	74.8687	64.4488	82.2172	82.2726	83.4784	74.5665	71.7998	85.958	84.3069	78.7913
	0.14	80.8163	83.9061	75.0363	64.254	86.1225	84.3175	85.1617	76.3265	72.3213	85.8441	83.2065	76.3065
	0.16	78.5379	84.898	77.7074	57.4	84.9388	86.5306	86.1225	75.7005	72.4173	83.3934	83.6925	77.0089
	0.18	76.2891	82.7661	78.3674	73.8776	84.7649	85.7427	84.867	75.1413	73.0612	83.3627	82.5804	76.747
ETH	**0.02**	78.0516	67.6699	70.0384	86.1159	78.6682	67.9645	85.9075	77.9559	70.9979	70.4268	69.8191	81.5384
	0.04	77.4512	71.9698	70.2882	88.0674	76.7744	69.863	86.484	79.4238	71.1385	69.9434	71.4867	80.7584
	0.06	79.1837	72.5227	71.2056	87.227	77.5422	69.9147	86.1296	79.5366	72.5663	72.3985	70.5626	81.822
	0.08	78.4099	73.0612	72.3569	88.1633	76.7693	70.6812	87.138	81.3402	72.964	71.3677	71.3225	75.465
	0.10	77.7862	71.5711	73.0612	87.2855	77.904	70.9351	87.6618	81.4873	74.7887	71.723	71.3758	83.0818
	0.12	76.7743	69.7643	70.3313	85.9589	79.4364	71.9264	88.5714	81.7792	74.0137	73.0144	72.2449	82.5046
	0.14	77.1432	70.2531	69.5106	85.4734	80.4082	72.2449	87.6069	81.7581	73.8436	73.6534	71.9671	83.6735
	0.16	76.3994	67.2634	68.5832	86.3227	79.2	71.8527	86.8224	83.1634	74.2857	74.3499	70.3137	82.9159
	0.18	75.2078	67.6144	68.1642	85.4273	78.2331	70.3619	86.2825	83.2653	73.2425	74.6939	68.1726	81.6593
Avg. 30Cryptos	**0.02**	74.3687	72.0768	69.8633	68.6519	71.7505	73.0974	79.8876	73.4947	70.4566	75.9526	73.6204	76.4697
	0.04	78.7212	73.8866	72.1644	68.0119	75.4778	76.6551	81.5624	76.3840	70.5716	78.5104	78.0091	80.6833
	0.06	78.5726	73.8597	75.0165	68.1954	75.7126	78.3275	84.9474	77.1609	74.1082	78.7791	78.8375	80.8755
	0.08	80.4871	76.5645	78.2982	71.3360	77.6631	79.9111	84.8873	79.6869	75.3076	80.0926	78.9770	80.0536
	0.10	80.0970	77.8194	76.5950	73.0030	80.0899	78.8243	86.1528	80.6897	77.7926	79.6512	78.7044	81.5433
	0.12	80.3810	80.8179	76.9340	72.1915	80.8417	79.9163	86.2221	82.2915	77.6734	79.9747	80.1481	82.4372
	0.14	80.7650	79.0131	77.9005	73.4487	82.8060	78.8681	86.6269	83.1791	77.7804	82.1999	80.3049	82.4122
	0.16	79.2781	80.0340	78.0405	72.8687	83.2690	77.5483	86.8182	81.7223	77.5096	81.1838	81.0636	82.4989
	0.18	76.2886	77.8994	77.8328	75.8732	80.9693	77.6585	84.0551	82.2659	75.5057	81.3661	81.0232	82.7166

Table S.2 RMSE for fitting the LSTM model in the LSTM-CNN FS-based model

Training Period	**Training Period 1**			**Training Period 2**			**Training Period 3**		
Crypto/Indicator	**OBV**	**A/D line**	**A/D OSC**	**OBV**	**A/D line**	**A/D OSC**	**OBV**	**A/D line**	**A/D OSC**
ADA	0.078	0.051	0.042	0.07	0.037	0.033	0.059	0.036	0.038
BNB	0.068	0.031	0.031	0.069	0.034	0.032	0.067	0.038	0.045
BTC	0.082	0.04	0.037	0.084	0.034	0.044	0.085	0.056	0.043
ETH	0.069	0.036	0.039	0.076	0.025	0.048	0.068	0.03	0.044
Avg. 30Cryptos	0.077	0.042	0.041	0.079	0.035	0.044	0.075	0.043	0.047
Training Period	**Training Period 4**			**Training Period 5**			**Training Period 6**		
ADA	0.065	0.036	0.033	0.073	0.043	0.039	0.068	0.034	0.051
BNB	0.065	0.034	0.036	0.074	0.05	0.044	0.078	0.036	0.04
BTC	0.074	0.036	0.031	0.076	0.034	0.031	0.084	0.058	0.043
ETH	0.066	0.027	0.042	0.061	0.032	0.033	0.079	0.033	0.047
Avg. 30Cryptos	0.068	0.035	0.039	0.073	0.044	0.037	0.080	0.043	0.046
Training Period	**Training Period 7**			**Training Period 8**			**Training Period 9**		
ADA	0.065	0.033	0.044	0.074	0.04	0.052	0.08	0.033	0.039
BNB	0.082	0.034	0.045	0.106	0.052	0.048	0.142	0.048	0.048
BTC	0.08	0.038	0.036	0.082	0.028	0.029	0.077	0.04	0.045
ETH	0.074	0.029	0.038	0.086	0.038	0.041	0.094	0.039	0.051
Avg. 30Cryptos	0.077	0.038	0.042	0.091	0.040	0.046	0.103	0.043	0.047
Training Period	**Training Period 10**			**Training Period 11**			**Training Period 12**		
ADA	0.079	0.046	0.055	0.065	0.033	0.041	0.058	0.026	0.058
BNB	0.109	0.058	0.033	0.081	0.038	0.039	0.057	0.041	0.051
BTC	0.061	0.035	0.05	0.066	0.044	0.049	0.093	0.045	0.061
ETH	0.09	0.042	0.03	0.074	0.041	0.04	0.114	0.064	0.049
Avg. 30Cryptos	0.086	0.048	0.046	0.072	0.041	0.047	0.084	0.047	0.055

All Training Period

Crypto/Indicator	**OBV**	**A/D line**	**A/D OSC**	**Average of Indicator**
ADA	0.070	0.037	0.044	0.050
BNB	0.083	0.041	0.041	0.055
BTC	0.079	0.041	0.042	0.054
ETH	0.079	0.036	0.042	0.052
Avg. 30Cryptos	0.080	0.042	0.045	0.056

This part shows thee average of 12 RMSEs of the training periods for each indicator.