

УДК 004.8:330.46

DOI: 10.20998/2411-0558.2020.01.10

В. В. МОРОЗ, канд. техн. наук, проф., Одеський національний університет, Одеса,

Д. ХЕЛВИГ, chief executive office, Black-Box. AI Inc.,

Д. В. МОРОЗ, студ., Одеський національний університет, Одеса,

П. П. ЖУКОВ, студ., Одеський національний університет, Одеса

АНАЛІЗ НЕЙРОМЕРЕЖЕВИХ МОДЕЛЕЙ LSTM ТА GMDH ДЛЯ ПРОГНОЗУВАННЯ КРИПТОВАЛЮТИ

Досліджується застосування нейромережових моделей для задачі прогнозування цін на криптовалюту. На відміну від класичних статистичних методів аналізу фінансових і економічних рядів, в основі яких є багатовимірний лінійний регресійний аналіз, пропонується модель з пам'яттю та адаптивна поліноміальна модель. Апробація моделей проводиться на даних криптовалютних ринків завдяки їх високій волатильності та низькій кореляції з традиційними активами. Для прогнозування застосовуються GMDH та LSTM нейронні мережі. Доведена перевага поліноміальної регресійної моделі GMDH за критерієм швидкість-точність прогнозування. Іл.: 5. Табл.: 1. Бібліогр.: 22 назв.

Ключові слова: поліноміальна регресійна модель; нейронна мережа; прогнозування; криптовалюта.

Постановка проблеми. Однією з найбільш актуальних і складних задач фінансової математики є задача прогнозування цін на активи. Аналіз сфери економічного і фінансового прогнозування показує, що з появою електронних грошей, інтерес до задачі тільки зростає. Криптовалюти, які з'явилися завдяки науково-технічному прогресу і комп'ютеризації усіх сфер діяльності людини, є підвидом електронних грошей. Такі криптовалюти, як Bitcoin (BTC), Ethereum (ETH) та інші, на відміну від традиційних валют, наприклад, USD та EUR, є децентралізованими цифровими валютами на основі Інтернет-протоколів open-to-peer (P2P). BTC та ETH використовують технологію блокчейн (blockchain) [1].

Загальна ринкова капіталізація криптовалют вимірюється сотнями мільярдів доларів. Криптовалюти відрізняються великою волатильністю, через що їх вартість значно змінюється у часі на відміну, наприклад, від акціонерного капіталу, та низькою кореляцією з традиційними активами. Також не існує явних базових принципів для криптовалютних активів, на основі яких можна було б побудувати фундаментальні торгові стратегії (наприклад, стратегії на основі вартості).

Для розробки нових стратегій або методів прогнозування торгівлі

криптовалютами необхідно застосовувати більш ефективні моделі, які ґрунтуються на глибокому навчанні, що дозволить зберігати складні закономірності і явища в даних.

Аналіз сучасних досліджень. Задачі прогнозування економічних і фінансових часових рядів приділяється висока увага протягом кількох десятиліть. Але дослідження в даній сфері проводилися більшою мірою з застосуванням інструментів економетрики – класичних статистичних методів. Найважливішим серед цих інструментів є багатовимірною лінійною регресією з оберненням коваріаційної матриці. Але сучасні фінанси оперують даними, для яких застосування лінійної регресії/авторегресії не обґрунтовано спрощує модель [2, 3]. До того ж, при прогнозуванні цін на фінансові активи, необхідно враховувати причинно-наслідкові зв'язки процесу ціноутворення. Якщо розглядати модель прогнозування як модель формування цін [4], ґрунтуючись на взаємодії суб'єктів ринку, які приймають економічні рішення на основі деяких показників або закономірностей, то базуючись на економічних законах або законах поведінкового фінансування доцільно застосовувати економетричні та балансові моделі. Моделі часових рядів та авторегресійні моделі враховують минулу динаміку, а стохастичні моделі враховують випадкові факти та події (політичні, пандемії, тощо), що ускладнює формальний опис причинно-наслідкових зв'язків.

Більшість досліджень при прогнозуванні часових рядів присвячені регресійним моделям (лінійна, множинна, нелінійна), які зараз майже не застосовуються на відміну від авторегресійних моделей. Найбільш популярними серед них є ARIMA [5] (Autoregressive Integrated Moving average) модель та її різноманітні модифікації [6, 7]. Для ARIMA моделей вхідні дані мають бути стаціонарним процесом, тобто має існувати порогове значення автокореляції. Для нестационарних процесів, якими і є процеси ціноутворення криптовалюти, застосовується дискретне диференціювання щоб зробити процес стаціонарним. Ще одним підвидом авторегресійних моделей є ARCH (Autoregressive Conditional Heteroscedasticity) моделі, які застосовують у випадку, коли в AR моделях враховується дисперсія помилок, та GARCH (Generalized Autoregressive Conditional Heteroskedasticity) модель [8-11], коли вона враховується в моделі ARMA. Моделі даного типу застосовуються при моделюванні фінансових часових рядів з високою волатильністю та кластеризацією коливань. Але через високий характер коливань, відсутність внутрішньої цінності та чіткого джерела походження всі відомі моделі дають прогнози з великим діапазоном варіацій або є помилковими або складними в обчисленнях.

Наступним етапом досліджень в прогнозуванні криптовалютних ринків стає розробка підходів на основі штучного інтелекту та штучних нейронних мереж ANN (Artificial Neural Network). Перші роботи в даному напрямку поєднали AR моделі та їх модифікації з ANN, що дало в результаті модель NNAR (Neural Network Autoregression). Аналіз результатів застосування NNAR показав, що дана модель з переоцінкою на кожному етапі може бути застосована у погодинних і похвилинних прогнозах [12].

Методика з використанням логістичної регресії та методу опорних векторів SVM (Support Vector Machine) для прогнозування ціни біткоіна запропонована в [16]. Модель на основі BNN (Bayesian Neural Network) досліджується в роботі [13], а в [14] пропонується застосування ансамбля нейронних мереж. Також в [15] дослідили прогнозування біткоіна за допомогою методики на основі штучної нейронної мережі ANN (Artificial Neural Network) в поєднанні з використанням ринкових технічних показників – технічного аналіз. Проте отримані результати виявилися неприйнятними через низьку продуктивність та час навчання.

Ціль статті. При застосуванні ANN виникає питання її структури та підготовки вхідних даних. В роботі досліджується дві нейронні мережі RNN-LSTM та GMDH-подібна. Велика кількість досліджень з застосування даних інструментів не дають сталої думки щодо ефективності їх як моделей для розв'язання задачі прогнозування криптовалютних часових рядів. Тому в даному дослідженні пропонується на вхід мережі подавати не історичні дані часового ряду, а його емпіричну модову декомпозицію – внутрішні модові функції.

Обґрунтування вибору ANN. Основою для вибору рекурентної мережі RNN стала її властивість зв'язувати попередню інформацію з поточною задачею, тобто попередні характерні коливання курсу криптовалюти з поточними (рис. 1). Але поточні коливання можуть відповідати характерним попереднім коливанням, які відбувалися у минулому на деякій часовій відстані.

Зі збільшенням цієї відстані RNN втрачає можливість зв'язувати інформацію. Теоретично це можливо, але практично призводить до постійної зміни параметрів мережі, що ускладнює її використання. Наприклад, на рис. 2 поточні коливання для заданого двох часового інтервалу могли мати місце кілька годин або днів, місяців тому.

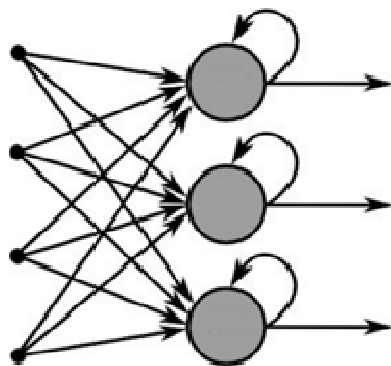


Рис. 1. Структура RNN



Рис. 2. Поточні коливання ціни біткоїна на заданому інтервалі

Навчання довгостроковим залежностям можливе завдяки застосуванню особливої архітектури RNN (рис. 3) – довгої короткострокової пам'яті LSTM (Long Short-Time Memory).

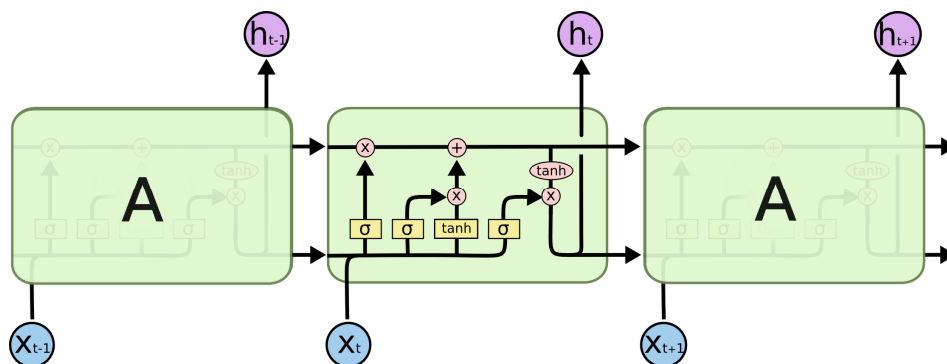


Рис. 3. Модуль LSTM [17]

Дана архітектура на відміну від RNN, яка має один шар нейронної мережі (рис. 1), включає чотири шари, що взаємодіють між собою (рис. 3). Через x_t і h_t позначені входи і виходи елемента мережі в момент часу t . Функції активації позначені через σ – сигмоїда та \tanh – гіперболічний тангенс.

Архітектура RNN, як і будь-якої ANN для глибокого навчання є відкритим питанням і є лише один висновок стосовно кількості шарів і нейронів в шарах – збільшення структурних розмірів мережі не покращує результат починаючи з деякої сталої архітектури. Це ж стосується і задачі прогнозування – ускладнення моделі не буде покращувати прогноз.

Будемо вважати, що ми не маємо уяви про структуру моделі, а фактори, які впливають на роботу моделі, носять оціночний характер і є нечіткими величинами та їх вплив явно є нелінійним. Це вимагає вибору моделі, яка буде підлаштовуватися під дані. Такі моделі можуть породжуватися за допомогою емпіричних методів моделювання – самоорганізації. Моделі, отримані за допомогою самоорганізації є ефективними для прогнозування.

Тому поставлена задача може бути роз'язана за допомогою Методу Групового Врахування Аргументів – GMDH (Group Method of Data Handling). Його особливість полягає в тому, що він дозволяє побудувати регресійні моделі і вибрати серед них модель з мінімальним числом параметрів [18 – 22].

Передбачається, що досліджуваний за допомогою GMDH об'єкт представляється кількома входами і одним виходом. Також він може бути змодельований певним набором компонентів базисної функції:

$$Y = (x_1, \mathbf{K}, x_n) = a_0 + \sum_{i=1}^M a_i f_i, \quad i = \overline{0, M},$$

де $x = (x_1, \dots, x_n)$ – входи; Y – вихід; a_i ($i = \overline{0, M}$) – коефіцієнти; f_i – елементарні базисні функції. Зв'язок між входами і виходом самоорганізуючої мережі може бути представлений нескінченим поліномом Вольтера-Колмогорова-Габора в формі:

$$y_n = \alpha_0 + \sum_{i=1}^M \alpha_i x_i + \sum_{i=1}^M \sum_{j=1}^M \alpha_{ij} x_i x_j + \sum_{i=1}^M \sum_{j=1}^M \sum_{k=1}^M \alpha_{ijk} x_i x_j x_k \dots$$

Відповідна мережа була побудована з послідовності простих поліномів. Тобто GMDH є поліноміальною нейронною мережею, в якій сигмоїдальна функція обробки вузла замінюється на поліноміальну. Входи визначаються як затримані (зсунуті на одиницю часу) елементи прогнозованого часового ряду. На відміну від ітеративного методу в

нейронних мережах зворотного поширення, мінімізації помилок в GMDH виконується шляхом послідовного наближення за допомогою найменших квадратів. Також навчання в даному випадку зводиться до розв'язання системи рівнянь. І ще одна перевага даної мережі – вона не може перетренуватися.

Для прогнозування ціни біткоїна з застосуванням RNN-LSTM та GMDH ANN були вибрані набори COLHV за місяць з інтервалом в одну хвилину. Історичні дані отримані за допомогою API торгової платформи Binance.com. Розмір часового ряду – 44640 елементів. Прогнозування виконувалося на 15 хвилин. Для прогнозування в мережі LSTM використовувалися бібліотеки Keras і TensorFlow без підтримки графічного процесора, GMDH реалізований з застосуванням псевдооберненої матриці методом Мура-Пенроуза. Дані для навчання мережі побудовані з часового ряду шляхом зсувів.

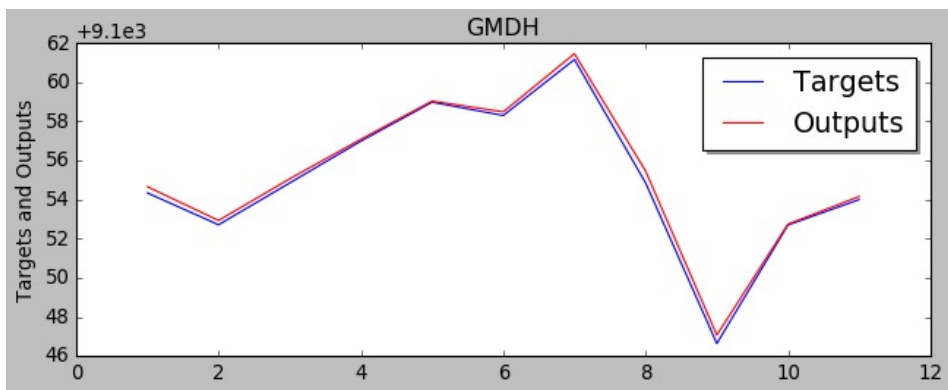


Рис. 4. Реальні і прогнозованні ціни за допомогою GMDH.

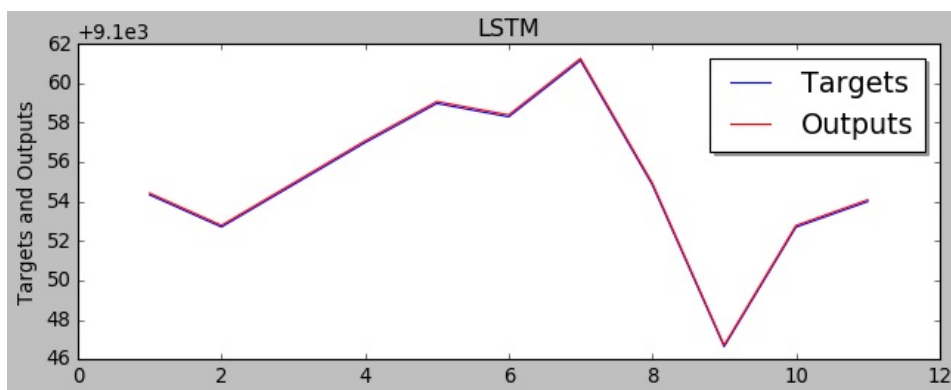


Рис. 5. Реальні і прогнозованні ціни за допомогою LSTM (100 epoch)

Для оцінки точності прогнозу застосуємо метрику середньої абсолютної похибки в процентах:

$$MAPE = \frac{1}{N} \sum_{i=1}^N \frac{Y_{i \text{ real}} - Y_{\text{forecasted}}}{Y_{i \text{ real}}} \cdot 100.$$

Точність прогнозу є доповненням MAPE до одиниці. Порівняльний аналіз за критерієм швидкість-точність прогнозування доводить перевагу індуктивної нейронної мережі.

Таблиця 1

	Реальні дані	GMDH	MAPE GMDH	LSTM-10	MAPE LSTM-10	LSTM-20	MAPE LSTM-20	LSTM-50	MAPE LSTM-50	LSTM-100	MAPE LSTM-100
	9154.34	9149.25	0.055602	9160.16	0.119245	9158.73	0.047955	9158.2	0.042166	9154.43	0.000983
	9152.71	9151.03	0.018355	9158.55	0.082177	9157.12	0.048182	9156.58	0.042283	9152.79	0.000874
	9154.85	9154.63	0.002403	9160.66	0.065868	9159.24	0.047953	9158.7	0.042054	9154.94	0.000983
	9156.99	9153.21	0.04128	9162.77	0.104444	9161.36	0.047723	9160.82	0.041826	9157.08	0.000983
	9158.98	9155.14	0.041926	9164.74	0.104859	9163.33	0.047494	9162.79	0.041599	9159.08	0.001092
	9158.3	9157.26	0.011356	9164.07	0.074367	9162.65	0.047498	9162.12	0.041711	9158.39	0.000983
	9161.16	9159.19	0.021504	9166.89	0.084069	9165.49	0.047265	9164.95	0.04137	9161.26	0.001092
	9154.85	9158.48	0.039651	9160.66	0.023803	9159.24	0.047953	9158.7	0.042054	9154.94	0.000983
	9146.64	9161.41	0.16148	9152.55	0.09671	9151.1	0.048761	9150.57	0.042967	9146.71	0.000765
	9152.7	9155.07	0.025894	9158.54	0.037902	9157.11	0.048183	9156.57	0.042283	9152.78	0.000874
	9154	9147.04	0.076032	9159.82	0.139717	9158.39	0.047957	9157.86	0.042167	9154.09	0.000983
	9155	9152.95	0.022392	9160.81	0.085874	9159.38	0.047843	9158.85	0.042054	9155.09	0.000983
Точність			0.961003		0.897441		0.952101		0.95789		0.999017
Час (с)		37.766		112.165		235.187		610.948		1217.287	

Висновки. Для прогнозування нелінійного нестационарного ряду на прикладі зміни цін на криптовалютних ринках для пари BTCUSDТT реалізовано дві нейронні мережі типу RNN-LSTM та GMDH. Порівняльний аналіз результатів прогнозу на 15 хвилин показав, що індуктивна нейронна мережа типу GMDH має суттєву перевагу в режимі реального часу і її точність поступається LSTM лише при навчанні останньої в 30 раз довше.

В подальших дослідженнях планується застосування індуктивної мережі з попередньою обробкою вхідних даних методами нелінійного аналізу і заміною псевдообернення Мура-Пенроуза більш ефективними методами.

References:

1. Nakamoto S. (2008), Bitcoin: A Peer-to-Peer Electronic Cash System, Available: <https://bitcoin.org/bitcoin.pdf>.
2. López de Prado, Marcos (2018), *Advances in Financial Machine Learning* (Chapter 1) (January 18, 2018). *Advances in Financial Machine Learning*, Wiley, 1st Edition. Available at SSRN: <https://ssrn.com/abstract=3104847>
3. Calkin, N., and M. López de Prado (2014), "The topology of macro financial flows: An application of stochastic flow diagrams", *Algorithmic Finance*, Vol. 3, No. 1, pp. 43-85.

4. Derbentsev, V., Datsenko, N., Stepanenko, O., and Bezkorovainyi, V. (2019), "Forecasting cryptocurrency prices time series using machine learning approach". *SHS Web Conf.* 65 02001. DOI: 10.1051/shsconf/20196502001.
5. Simon Stevenson (2007), "A comparison of the forecasting ability of ARIMA models", *Journal of Property Investment & Finance*, 25 (3), pp. 223-240.
6. Chen, An-Sing, Hung-Chou Chang, and Lee-Young Cheng (2019), "Time-varying Variance Scaling: Application of the Fractionally Integrated ARMA Model", *The North American Journal of Economics and Finance*, 47, pp. 1-12.
7. Hyndman, Rob J., and George Athanasopoulos (2018), *Forecasting: Principles and Practice*. Melbourne: OTexts, Available online: <https://otexts.com/fpp2/> (accessed on 07 May 2020).
8. Bollerslev, Tim, Russell, Jeffrey, and Watson, Mark (2010), "Chapter 8: Glossary to ARCH (GARCH)". *Volatility and Time Series Econometrics: Essays in Honor of Robert Engle (1st ed.)*. Oxford: Oxford University Press, pp. 137-163.
9. Klüppelberg, C., Lindner, A., and Maller, R. (2004), "A continuous-time GARCH process driven by a Lévy process: stationarity and second-order behaviour", *Journal of Applied Probability*. 41 (3): pp. 601-622. doi:10.1239/jap/1091543413.
10. Li, D., Zhang, X., Zhu, K., and Ling, S. (2018), "The ZD-GARCH model: A new way to study heteroscedasticity" (PDF). *Journal of Econometrics*. 202 (1), pp. 1-17. doi:10.1016/j.jeconom.2017.09.003.
11. Otto, P., Schmid, W., and Garthoff, R. (2018), "Generalised spatial and spatiotemporal autoregressive conditional heteroscedasticity". *Spatial Statistics*, 26 (1), pp. 125-145. doi:10.1016/j.spasta.2018.07.005.
12. Khashei, M., and Bijari, M. (2011), "A novel hybridization of artificial neural networks and ARIMA models for time series forecasting", *Applied Soft Computing*, Vol. 11, No. 2, pp. 2664-2675. doi: 10.1016/j.asoc.2010.10.015.
13. Jang, H., and Lee, J. (2018), "An Empirical Study on Modeling and Prediction of Bitcoin Prices With Bayesian Neural Networks Based on Blockchain Information", *IEEE ACCESS*, Vol. 6, pp. 5427-5437.
14. Sin, E., and Wang, L. (2017), "Bitcoin Price Prediction Using Ensembles of Neural Networks", in *2017 13th International conference on natural computation, fuzzy systems and knowledge discovery (ICNC FSKD)*, pp. 666-671.
15. Radityo, A., Munajat, Q., and Budi, I. (2017), "Prediction of Bitcoin exchange rate to American dollar using artificial neural network methods", in *Advanced Computer Science and Information Systems (ICACSIS)*, 2017 International Conference on, pp. 433-438.
16. Greaves A, and Au B (2015), "Using the Bitcoin transaction graph to predict the Price of Bitcoin". Available online (accessed 10 May 2020): <https://pdfs.semanticscholar.org/a0ce/864663c100582805ffa88918910da89add47.pdf>.
17. Olah, C. (2015), "Understanding LSTM Networks", Available online (accessed 20 May 2020): <https://colah.github.io/posts/2015-08-Understanding-LSTMs>.
18. Ivakhnenko, A.G. (1971), Polynomial theory of complex systems, *IEEE Trans. Systems, Man Cybernet.*, 1, pp. 364-378.
19. Madala, H. R. and Ivakhnenko, A. G. (1994), *Inductive Learning Algorithms for Complex Systems Modelling*, CRC Press, Boca Raton.
20. Nikolaev, N.Y. and Iba, H. (2003), "Polynomial harmonic gmdh learning networks for time series modeling", *Neural Netw.*, Vol. 16, pp. 1527-1540.
21. Onwubolu, G. (2009), *Hybrid Self-Organizing Modeling Systems*, Springer, Germany.
22. Zaichenko, Y.P., and Zaets, I.O. (2007), "Comparative analysis of the predictive models built using fuzzy algorithms and fuzzy GMDH using different algorithms for generating fuzzy prediction models", *Proceedings of the Intern. Conference on inductive modeling, National Academy of Sciences Ukraine, INNS inform.technology and systems*, no. 4, pp. 158-165.

*Статтю подав д-р техн. наук, проф. ОНУ імені І.І. Мечникова
Гунченко Ю.А.*

Надійшла (received) 23.05.2020

Moroz Volodymyr, PhD in Applied Mathematics, Professor
Odessa I.I. Mechnikov National University
Str Dvoryans'ka, 2, Odesa, Ukraine, 65072
Phone (+380) 67-484-6975, e-mail v.moroz@onu.edu.ua

Helvig Jonas, CEO
BlackBox Inc.
Street Nile, 64, N1 7SR, London
Phone (+44) 7-4807-24361, e-mail jonas@black-box.ai

Moroz Dmytro, stud.
Odessa I.I. Mechnikov National University
Str Dvoryans'ka, 2, Odesa, Ukraine, 65072
Phone (+380) 93-234-7827, e-mail Dmytro-Moroz@stud.onu.edu.ua

Zhukov Pavlo, stud.
Odessa I.I. Mechnikov National University
Str Dvoryans'ka, 2, Odesa, Ukraine, 65072
Phone (+380) 50-708-7489, e-mail Pavel-Zhukov@stud.onu.edu.ua

УДК 004.8:330.46

Аналіз нейромережевих моделей LSTM та GMDH для прогнозування криптовалюти / Мороз В.В., Хелвіг Д., Мороз Д.В., Жуков П.П. // Вісник НТУ "ХПІ". Серія: Інформатика та моделювання. – Харків: НТУ "ХПІ". – 2020. – № 1 (3). – С. 113 – 122.

Досліджується застосування нейромережевих моделей для задачі прогнозування цін на криптовалюти. На відміну від класичних статистичних методів аналізу фінансових і економічних рядів, в основі яких є багатовимірний лінійний регресійний аналіз, пропонується модель з пам'яттю та адаптивна поліноміальна модель. Апробація моделей проводиться на даних криптовалютних ринків завдяки їх високій волатильності та низькій кореляції з традиційними активами. Для прогнозування застосовуються GMDH та LSTM нейронні мережі. Доведена перевага поліноміальної регресійної моделі GMDH за критерієм швидкість-точність прогнозування. Іл.: 5. Табл.: 1. Бібліогр.: 22 назв.

Ключові слова: поліноміальна регресійна модель; нейронна мережа; прогнозування; криптовалюта.

УДК 004.8:330.46

Анализ нейросетевых моделей LSTM и GMDH для прогнозирования криптовалюты / Мороз В.В., Хелви́г Д., Мороз Д.В., Жуков П.П. // Вестник НТУ "ХПИ". Серія: Інформатика и моделирование. – Харьков: НТУ "ХПИ". – 2020. – № 1 (3). – С. 113 – 122.

Исследовано применение моделей нейронных сетей для решения проблемы прогнозирования цен на криптовалюту. В отличие от классических статистических методов анализа финансово-экономических рядов, основанных на многомерном линейном регрессионном анализе, предложена модель с памятью и адаптивная полиномиальная модель. Модели тестируются на примере прогнозирования криптовалюты из-за ее высокой изменчивости и низкой корреляции с традиционными активами. Для прогнозирования используют нейронные сети GMDH и LSTM. Доказано преимущество полиномиальной регрессионной модели GMDH по критерию прогнозирования скорости и точности. Ил.: 5. Табл.: 1. Библиогр.: 22 назв.

Ключевые слова: полиномиальная регрессионная модель; нейронная сеть; прогнозирование; криптовалюта.

UDC 004.8:330.46

Analysis of LSTM and GMDH network models for cryptocurrency forecasting / Moroz Volodymyr, Ceo Jonas Helvig, Moroz Dmytro, Zhukov Pavel // Herald of the National Technical University "KhPI". Series of "Informatics and Modeling". – Kharkov: NTU "KhPI". – 2020. – № 1 (3). – P. 113 – 122.

The application of neural network models for the problem of cryptocurrency price forecasting is investigated. Unlike classical statistical methods of financial and economic series analysis, which are based on multidimensional linear regression analysis, a memory model and an adaptive polynomial model are proposed. Models are being tested on cryptocurrency markets due to their high volatility and low correlation with traditional assets. GMDH and LSTM neural networks are used for forecasting. The advantage of the polynomial regression model GMDH is proved by the criterion of speed-accuracy forecasting. Figs.: 5. Tabl.: 1. Refs.: 22 titles.

Keywords: polynomial regression model; neural network; prognostication; cryptocurrency.