

## **ПРОГНОЗУВАННЯ КУРСУ НА ОСНОВІ МОДЕЛІ ARIMA**

**Макарчук Олег, Петрикова Ірина**

**Науковий керівник: канд.-ф.-м. наук, доцент Макарчук О.П.**

*Центральноукраїнський державний педагогічний університет імені*

*Володимира Винниченка, м. Кропивницький, Україна*

*В статті представлено оцінки точності прогнозування на основі ARIMA моделі. Найкращий спосіб вивчити ринок — вибрати найпопулярніші валюти та витратити час на їх дослідження. Нереально очікувати відстеження кожного типу, який стає доступним, особливо враховуючи, наскільки легко створити стартап криптовалюти. Намагаючись дізнатися більше про різні типи доступної цифрової валюти, це також допомагає зрозуміти, як вона розробляється, і мати обмежене розуміння того, як працює структура ціноутворення. Проведено прогнозування трьох криптовалют, а саме: Zcash, Dash, Monero, зібравши щомісячні дані з 2018-2021 роки. Визначено подальші кроки дослідження.*

*Ключові слова: числові ряди, короткострокове прогнозування, модель ARIMA, криптовалюта, оцінка точності, авторегресивна модель, лінійне експоненціальне згладжування, блокчейн,*

## **Prognostication of course is on basic of model of ARIMA**

**O Makarchuk, I Petrykova**

**Scientific supervisor: Candidate of Physics and Mathematics Science**

**Makarchuk O.P.**

*Volodymyr Vynnychenko Ukrainian State Pedagogical University,*

*Kropyvnytsky, Ukraine*

*The article presents estimates of forecasting accuracy based on the ARIMA model. The best way to study the market is to choose the most popular currencies and spend time researching them. It is unrealistic to expect tracking of each type that becomes available,*

*especially considering how easy it is to create a cryptocurrency startup. Trying to learn more about the different types of digital currency available also helps to understand how it is developed and to have a limited understanding of how pricing works.*

*Keywords: numerical series, short-term forecasting, ARIMA model, cryptocurrency, accuracy estimation, autoregressive model, linear exponential smoothing, blockchain,.*

## **1. Постановка проблеми.**

У процесі миттєвого розвитку інформаційних технологій, кожного дня з'являються все нові та нові інструменти управління економікою, до яких можна пов'язати й електронні гроші. У більшості країн, економіка розвивається та підвищує свій рівень, а люди постійно перебувають у пошуку більш безпечного, найзручнішого та кращого. Це і привило до появи на ринку аналог звичайних валют — криптовалюти. Криптовалютою називають цифрові активи, які функціонують як валюта, використовуючи криптографію для забезпечення безпеки. Головною відмінністю криптовалюти від інших електронних грошей – це анонімність користування нею та абсолютна децентралізація. Найпоширенішою криптовалютою є саме Bitcoin. Деякі країни світу її вже легалізували та офіційно використовують. Що стосується України, то хоча уряд не визнав криптовалюту офіційно легальною, але наша країна входить до десяти країн світу за чисельністю користування електронною валютою. Кожного дня, розмір торгів криптовалютою використовуючи гривні, наближується до 1.9 млн. доларів. Для активного використання різних видів криптовалют, питання розвитку та застосування цього платіжного засобу в Україні потребують подальших досліджень.

Робота присвячена проблемі знаходження кращої та точної короткострокової моделі для різних видів криптовалют.

**Об'єкт дослідження:** є динаміка зміни курсу криптовалюти.

**Предмет дослідження:** є методи та модель прогнозування часових рядів на основі використання ARIMA–моделей.

**Мета дослідження:** побудова математичної моделі ряду, за допомогою якої можна пояснити поведінку ряду і здійснити прогноз на майбутні періоди.

## **2. Прогнозування курсу на основі моделі ARIMA.**

Рівняння прогнозування ARIMA(p,d,q): моделі ARIMA теоретично є найзагальнішим класом моделей для прогнозування часового ряду, який можна зробити «стаціонарним» шляхом диференціювання (якщо необхідно), можливо, у поєднанні з нелінійними перетвореннями, такі як журналювання або дефляція (якщо необхідно). Випадкова величина, яка є часовим рядом, є стаціонарною, якщо всі її статистичні властивості є постійними в часі. Стаціонарний ряд не має тенденції, його варіації навколо середнього мають постійну амплітуду, і він постійно хитається, тобто його короткострокові випадкові часові моделі завжди виглядають однаково в статистичному сенсі. Остання умова означає її автокореляцію (кореляції з його власними попередніми відхиленнями від середнього) залишаються постійними протягом часу, або еквівалентно, що його спектр потужності залишається постійним протягом часу. Випадкову величину цієї форми можна розглядати (як зазвичай) як комбінацію сигналу та шуму, а сигнал (якщо він є очевидним) може бути зразком швидкої або повільної реверсії середнього, або синусоїдального коливання, або швидкої зміни знака, а також може мати сезонну складову. Модель ARIMA можна розглядати як «фільтр», який намагається відокремити сигнал від шуму, а потім сигнал екстраполюється в майбутнє для отримання прогнозів.

Рівняння прогнозування ARIMA для стаціонарного часового ряду являє собою лінійне (тобто рівняння типу регресії), в якому провісники

складаються із затримок залежної змінної та/або лагів помилок прогнозу.

Тобто:

Передбачуване значення  $Y$  = константа та/або зважена сума одного чи кількох останніх значень  $Y$  та/або зважена сума одного чи кількох останніх значень помилок.

Прогнозне рівняння будується таким чином. Спочатку нехай  $y$  позначає  $d$ -ю різницю  $Y$ , що означає:

Якщо  $d=0$ :

$$y_t = Y_t.$$

Якщо  $d=1$ :

$$y_t = Y_t - Y_{t-1}.$$

Якщо  $d=2$ :

$$y_t = (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}) = Y_t - 2Y_{t-1} + Y_{t-2}.$$

Зауважимо, що друга різниця  $Y$  (випадок  $d=2$ ) не є відмінністю від 2 періодів тому. Швидше, це перша різниця-від-першої різниці, яка є дискретним аналогом другої похідної, тобто локальне прискорення ряду, а не його локальний тренд.

У термінах  $y$  загальне рівняння прогнозування виглядає так:

$$\hat{y}_t = \mu + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q}.$$

Тут параметри ковзного середнього ( $\theta'_s$ ) визначаються таким чином, що їх знаки в рівнянні є від'ємними, згідно з умовою, введеною Боксом і Дженкінсом. Деякі автори та програмне забезпечення (включаючи мову програмування R) визначають їх так, щоб замість них були знаки плюс. Коли фактичні числа підключені до рівняння, немає двозначності, але

важливо знати, яку конвенцію використовує ваше програмне забезпечення, коли ви читаєте вихідні дані. Часто параметри позначаються там AR(1), AR(2), ... і MA(1), MA(2), ... тощо.

Щоб визначити відповідну модель ARIMA для  $Y$ , ви почнете з визначення порядку диференціації ( $d$ ) для стаціоризації ряду та видалення загальних ознак сезонності, можливо, у поєднанні з перетворенням, що стабілізує дисперсію, таким як каротаж або дефляція. Якщо ви зупинитесь на цій точці і передбачите, що різниця рядів є постійною, ви просто підібрали модель випадкового блукання або випадкової тенденційної моделі. Однак стаціонаризовані ряди все ще можуть мати автокорельовані помилки, що свідчить про те, що деяка кількість термінів AR ( $p \geq 1$ ) та/або деяка кількість членів MA ( $q \geq 1$ ) також необхідні в рівнянні прогнозування.

Процес визначення значень  $p$ ,  $d$  і  $q$ , які є найкращими для даного часового ряду, буде обговорено в наступних розділах приміток (посилання на які розміщені у верхній частині цієї сторінки), але попередній перегляд деяких типів з сезонності моделей ARIMA, які зазвичай зустрічаються наводиться нижче.

Наприкладі розглянемо модель :

ARIMA(1,1,0) = диференційована авторегресивна модель першого порядку: якщо помилки моделі випадкового блукання є автокорельованими, можливо, проблему можна виправити, додавши одну затримку залежної змінної до рівняння передбачення, тобто за допомогою регресія першої різниці  $Y$  на себе відстає на один період. Це дасть таке рівняння прогнозу:

$$\hat{Y}_t - Y_{t-1} = \mu + \phi_1 (Y_{t-1} - Y_{t-2})$$

$$\hat{Y}_t - Y_{t-1} = \mu,$$

які можна переставити

$$\hat{Y}_t = \mu + Y_{t-1} + \phi_1 (Y_{t-1} - Y_{t-2}).$$

Це авторегресивна модель першого порядку з одним порядком несезонної різниці та постійним членом, тобто модель ARIMA(1,1,0).

Реалізація електронних таблиць: моделі ARIMA, подібні до описаних вище, легко реалізувати в електронній таблиці. Рівняння прогнозу — це просто лінійне рівняння, яке посилається на минулі значення вихідних тимчасових рядів і минулі значення помилок. Таким чином, ви можете налаштувати електронну таблицю прогнозування ARIMA, зберігаючи дані в стовпці A, формулу прогнозу в стовпці B і помилки (дані мінус прогнози) у стовпці C. Формула прогнозування в типовій клітинці в стовпці B буде просто лінійний вираз, що відноситься до значень у попередніх рядках стовпців A і C, помножених на відповідні коефіцієнти AR або MA, що зберігаються в клітинках в іншому місці електронної таблиці.

Для короткострокового прогнозування були вибрані 3 види криптовалюти, а саме:

- Zcash. Zcash був запущений в останній половині 2016 року, і хоча він відносно новий на ринку, він уже має багато прихильників в інвестиційній спільноті. Zcash пишається своєю надійною безпекою та конфіденційністю.
- Monero. Monero — це криптовалюта з відкритим вихідним кодом, запущена в квітні 2014 року. Вона є безпечною, приватною та невідстежуваною та швидко викликає інтерес у спільноті. На відміну від

інших типів цифрової валюти, розробка Monero була виключно на основі пожертвувань і керувалася спільнотою.

- Dash. Колись відомий як Darkcoin, Dash схожий на Zcash тим, що він також пропонує більше анонімності, ніж Bitcoin. Він працює в децентралізованій мережі mastercode.

Були взяті середньомісячні дані за 4 роки ( від 2018 р. по 2021 р. ).

Задача: визначити кращу модель короткострокового прогнозування на основі ARIMA модель для трьох криптовалют.

Для прогнозування будуть використовуватись щомісячні дані з 2018-2021 роки по трьом криптовалютам. Дані були зібрані з сайту Coinmarketcap [4, 5]. Маємо таблицю, дані представлені у доларах.

Таблиця 1. Щомісячні дані з 2018 р. по 2021 р. Zcash - криптовалюта

Zcash							
місяці	дані	місяці	дані	місяці	дані	місяці	дані
1	61,07	10	44,83	19	52,81	28	137,72
2	54,98	11	36,93	20	60,45	29	217,03
3	51,59	12	34,24	21	83,56	30	226,11
4	54,05	13	29,29	22	60,16	31	131,13
5	67,77	14	45,04	23	63,85	32	104,99
6	69,6	15	64,14	24	64,93	33	140,54
7	96,15	16	35,99	25	69,09	34	130,53
8	84,1	17	39,65	26	85,06	35	138,41
9	54,84	18	45,62	27	129,79		

Проведемо прогнозування за ARIMA (1,1,0). За формулою:

$$\hat{Y}_t - Y_{t-1} = \mu + \phi_1 (Y_{t-1} - Y_{t-2})$$

$$\hat{Y}_t - Y_{t-1} = \mu,$$

які можна переставити

$$\hat{Y}_t = \mu + Y_{t-1} + \phi_1 (Y_{t-1} - Y_{t-2}).$$

Отримали короткостроковий прогноз на три місяці.

Таблиця 2. Прогнозування на три місяці за ARIMA (1,1,0). Zcash - криптовалюта

Zcash	
місяці	дані
33	137,11
34	134,07
35	141,91

Щоб розрахувати прогнози на правильність, були взяті фактичні (тобто справжні) дані на три місяці:

Таблиця 3. Фактичні дані трьох місяців. Zcash - криптовалюта

Zcash	
місяці	дані
33	140,54
34	130,53
35	138,41

Висновки та перспективи подальшого дослідження. Для короткострокового прогнозування було розглянуто три криптовалюти, а саме: Zcash, Monero, Dash. За допомогою ARIMA модель, були отримані прогнози з усіх трьох валют.

Важливою частиною залишається побудова графіків та оцінка точності, яка дасть змогу визначити найточнішу модель прогнозування на майбутній період для кожної з трьох криптовалют.

### Список літератури

1. Бутенко Е.Д. Біткоїн. Стан і перспекиви розвитку криптовалюта // Фінанси і кредит. – 2014. – №23 (599). – С. 44–47.
2. Ден Л.Л. Глибоке навчання: методи та застосування / Ден Л. та Ю.



Д. //Основи та тенденції обробки сигналів, 7(3-4) – 2014. – pp. 197-387.

3. Єгошин А.В. Аналіз та прогнозування складних стохастичних сигналів на основі методів ведення меж реалізацій динамічних систем: Автореферат дисертації... канд. техн. наук. Санкт-Петербург, 2009. 19 с.

4. Топ-100 Криптовалюти з ринкової капіталізації. - URL:  
<https://coinmarketcap.com/ru>

5. Zcash - URL: <https://coinmarketcap.com/ru/currencies/zcash/markets/>