

Г. С. Іващенко, В. О. Понамарьов, В. О. Холєв

Харківський національний університет радіоелектроніки, Харків, Україна

## КОРОТКОСТРОКОВЕ ПРОГНОЗУВАННЯ НЕСТАЦІОНАРНИХ ЧАСОВИХ РЯДІВ З ВИКОРИСТАННЯМ МОДЕЛЕЙ MLP ТА LSTM

**Анотація.** **Актуальність.** Рішення завдання прогнозування відіграє важливу роль у процесах стратегічного планування та оперативного управління у різних сферах господарчої діяльності. Формою прогнозування є прогнозування часових рядів, при якому актуальною проблемою залишається вибір доцільного методу серед сучасних засобів обчислювального інтелекту, таких як штучні нейронні мережі. Проблема вибору обумовлена великою кількістю параметрів та налаштувань, які залежать від особливостей прогнозованого часового ряду, та суттєво впливають на якість отриманого прогнозу. **Метою даної роботи** є аналіз методів короткострокового прогнозування нестационарних часових рядів з використанням моделей штучних нейронних мереж, таких як багатошаровий перцептрон та довга короткострокова пам'ять. **Об'єктом дослідження** є процес прогнозування часових рядів. **Предметом дослідження** є використання моделей штучних нейронних мереж для короткострокового прогнозування. **Результати.** Експериментальні дослідження показали, що середня помилка при прогнозуванні за допомогою запропонованих засобів на 2-6% нижче у порівнянні з використанням поширених традиційних моделей. **Висновок.** Отримані результати підтверджують перевагу використання моделей MLP та LSTM перед прогнозуванням на основі методів, обраних для аналізу у M3-Competition.

**Ключові слова:** прогнозування часових рядів, обчислювальний інтелект, штучна нейронна мережа, багатошаровий перцептрон, довга короткострокова пам'ять, M3-Competition.

### Вступ

Завдання прогнозування зустрічається практично у будь-яких предметних областях та галузях діяльності. Поширене використання прогнозування у економіці, наприклад, для визначення майбутнього обсягу продажу товарів та зміни цінової політики, виходячи з припущеної зміни попиту. З урахуванням прогнозів щодо майбутнього розвитку процесів та подій можна гнучкіше планувати діяльність підприємств та приймати виважені управлінські рішення.

Окремим випадком завдання прогнозування є прогнозування часових рядів, як однієї з популярних форм представлення інформації о процесах реального світу. Значення параметра, що описують зміни у предметній області, фіксуються через певні часові інтервали та утворюють часовий ряд. У такому випадку прогноз є припущенням щодо майбутніх значень часового ряду на основі наявних даних. Наприклад, підприємства у процесі своєї діяльності накопичують великий обсяг інформації, що може бути представлена у вигляді часових рядів та використана для отримання прогнозу щодо подальших перспектив розвитку бізнесу.

При вирішенні завдань прогнозування часових рядів широкою популярністю набуло використання статистичних методів та підходів на основі засобів обчислювального інтелекту [1, 2].

Частина існуючих моделей та відповідних методів використовуються на окремих етапах процесу прогнозування, таких як підготовка вихідних даних, обробка присутніх у реальних даних викривлень та інших. Результат прогнозування часового ряду залежить від багатьох параметрів, таких як обсяг та достовірність необхідної початкової інформації, урахування можливості зміни середовища, в якому протікає процес, доступність необхідних обчислювальних засобів [3, 4].

Проблема вибору методу прогнозування часових рядів широко представлена у сучасних наукових роботах. Особливе місце займають штучні нейронні мережі (ШНМ). На відміну від багатьох статистичних методів, нейронні мережі більш стійкі до зашумлених даних і здатні використовувати необмежену кількість незалежних змінних.

У дослідженні [5] проведено емпіричне порівняння використання засобів обчислювального інтелекту, зокрема, методу k-ближчих сусідів, і окрема увага приділена аналізу моделей ШНМ, таких як багатошаровий перцептрон (MLP), RBF, узагальнено-регресійні ШНМ, ШНМ на основі ядерної регресії та регресії гаусівського процесу. Для експериментальних досліджень були обрані часові ряди щомісячних даних з M3-Competition [6]. Найкращі результати серед розглянутих засобів показали багатошаровий перцептрон та штучна нейронна мережа на основі регресії гаусівського процесу.

У [7] обґрунтовано застосування багатошарового перцептрон для прогнозування часових рядів розповсюдження захворювання COVID-19.

Емпіричні дослідження, описані у [8], показують, що алгоритми прогнозування на основі ШНМ глибокого навчання, такі як модель довгої короткострокової пам'яті (LSTM), перевершують традиційні підходи, які засновані на моделі ARIMA.

Дослідження [9] обґрунтовує використання моделей MLP та LSTM, доводячи, що вони демонструють найкращу ефективність при прогнозуванні нестационарних рядів на один крок вперед.

Виходячи з аналізу сучасних досліджень і публікацій можна зробити висновок, що завдання прогнозування за допомогою моделей ШНМ є актуальним, і тому доцільним є дослідження ефективності використання популярних моделей, таких як MLP та LSTM, при прогнозуванні часових рядів з M3-Competition.

**Метою цієї роботи** є аналіз методів короткострокового прогнозування нестационарних часових рядів на основі засобів обчислювального інтелекту.

### Постановка завдання

Часовий ряд представлений вектором, який відображає значення, що характеризують протікання якогось процесу протягом останніх  $N$  інтервалів часу,  $x_t = (x_{t-N}, x_{t-N+1}, x_{t-i}, \dots, x_{t-2}, x_{t-1})$ , де  $t$  – мітка поточного часу,  $0 \leq t \leq N$ .

Ряд, що залишається у рівновазі щодо постійного середнього рівня є стаціонарним [10]. Відповідність вхідних даних до умови стаціонарності дозволяє отримати точніші дані прогнозу, на відміну від прогнозування нестационарних часових рядів, які є більш поширеними у економіці і тому проблема прогнозування саме нестационарних рядів є більш актуальною [10].

Короткостроковість визначається значенням горизонту прогнозування. Горизонт прогнозування – це часовий інтервал, в межах якого прогноз виконується із заданою точністю [11, 12].

Необхідно прогнозувати майбутні значення у нестационарних часових рядах, наявних у M3-Competition, з значеннями горизонту прогнозування, відповідними до представлених у [6] результатів. Дослідження передбачає використання моделей штучних нейронних мереж, таких як багатошаровий перцептрон та довга короткострокова пам'ять та порівняння отриманих результатів з результатами прогнозування в M3-Competition.

### Основна частина

Багато з використовуваних підходів до прогнозування побудовані на невеликій кількості моделей, та відрізняються окремими прийомами та послідовністю їх застосування.

Використовувані у роботі моделі багатошарового перцептрон та довгої короткострокової пам'яті мають один прихований шар вузлів та вихідний шар, який використовується для отримання результатів прогнозування. Однією з проблем використання моделей є підготовка даних для створення навчальної, тестової та контрольної вибірок.

Моделі MLP та LSTM відображають послідовність минулих спостережень, у якості вхідних даних, у вихідне значення. Таким чином, послідовність спостережень має бути перетворена на кілька прикладів, на яких модель може навчатися.

Звичайна послідовність вихідних даних у випадку використання часового ряду має вигляд  $x_1, x_2, x_3, \dots, x_N$ . Для подальшого використання при навчанні ШНМ таку вихідну послідовність необхідно розділити для створення навчальної вибірки на кілька шаблонів введення та виведення, званих зразками, де перші значення часового ряду використовуються як вхідні дані, а наступне значення (чи декілька значень, залежно від горизонту прогнозування) використовується як вихідні дані для однокрокового (чи багатокрокового) прогнозування.

Розділена послідовність має вигляд:

$$\begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ x_2 & x_3 & x_4 & x_5 \\ x_3 & x_4 & x_5 & x_6 \\ x_4 & x_5 & \dots & x_N \end{pmatrix}. \quad (1)$$

Значення вхідної змінної може бути числовим або категоріальним і може набувати кількох значень. Усі вони повинні бути нормовані та приймати значення в діапазоні від нуля до одиниці. Нормалізацію значень можна виконати за допомогою наступного виразу:

$$x_t = \frac{0.1 - 0.9}{x_{\min} - x_{\max}} x_t + \frac{0.9 * x_{\min} - 0.1 * x_{\max}}{x_{\min} - x_{\max}}, \quad (2)$$

де  $x_{\min}$ ,  $x_{\max}$  мінімальне та максимальне значення вхідних значень.

Використана модель навчання з вчителем, приклади навчальної вибірки, виходячи з обсягу наявних у M3-Competition часових рядів, містять по вісім значень часового ряду (вхідні значення для ШНМ) та по одному коректному значенню результату, яке може бути використане при корегуванні ваг зв'язків між вузлами ШНМ у процесі навчання. Таким чином, навчання виконуються шляхом покрокового прогнозу на підставі восьми попередніх значень часового ряду.

Модель MLP складається із вхідного, прихованого та вихідного шарів. Кожен шар складається з деякої множини вузлів. Вага зв'язків між вузлами є позитивним числом, якщо один вузол стимулює інший або негативним, якщо один вузол пригнічує інший. Вузли з вищими значеннями ваги мають більший вплив інші вузли [13].

Для прогнозу з використанням багатошарового перцептрон використовується мережа з  $n$  вузлами на вході (дорівнює кількості значень часового ряду, на основі яких будується прогноз),  $m$  внутрішніми вузлами та одним вузлом на виході (у випадку прогнозування на один крок уперед) [14].

На вході кожного внутрішнього вузла значення стимулу дорівнює:

$$net_j(x_t) = \sum_{i=1}^n w_{ij} x_{t-1} + b_j, \quad (3)$$

де  $w_{ij}$  – вага між вузлом  $i$  та вузлом  $j$ ,  $b_j$  – зміщення внутрішнього вузла  $i$ .

На виході внутрішнього вузла значення стимулу дорівнює:

$$y_j = f \left( \sum_{i=1}^n w_{ij}^h x + b_j \right), \quad (4)$$

де  $w_{ij}^h$  – вага між вузлом  $i$  та вузлом  $j$ ,  $b_j$  – зміщення внутрішнього вузла  $i$ .

На вході вихідного вузла значення стимулу дорівнює:

$$net^o(x_t) = \sum_{i=1}^n w_i^o x_{t-1} + b^o, \quad (5)$$

де  $w_i^o$  – вага між вузлом  $i$  та вузлом  $j$ ,  $b^o$  – зміщення внутрішнього вузла  $i$ .

На виході вихідного вузла значення стимулу дорівнює:

$$y_t = f \left( \sum_{i=1}^n w_i^o * f \left( \sum_{j=1}^m w_{jk} + b_j \right) + b^o \right), \quad (6)$$

де  $w_i^o$  – вага між вузлом  $i$  та вузлом  $j$ ,  $w_{jk}$  – вага між вузлом  $i$  та вихідним вузлом,  $b_j$  – зміщення внутрішнього вузла  $i$ ,  $b^o$  – зміщення вихідного вузла.

Нейрони у моделі LSTM складають клітини пам'яті, блок входу, блок виходу, вхідний шлюз та вихідний шлюз [15].

Для прогнозу з використанням довгої короткострокової пам'яті використовується мережа з  $n$  блоків обробки та  $m$  входів [16].

Блок входу поєднує поточний вхід та вихід цього блоку в останній ітерації:

$$z^{(t)} = g \left( w_z x^{(t)} + r_z y^{(t-1)} + b_z \right), \quad (7)$$

де  $w_z$  та  $r_z$  – вагові коефіцієнти  $x^{(t)}$  та  $y^{(t-1)}$ ,  $b_z$  – вектор зміщення.

Вхідний шлюз поєднує поточний вхід, вихід цього блоку та значення клітинки пам'яті на останній ітерації:

$$i^{(t)} = \sigma \left( w_i x^{(t)} + r_i y^{(t-1)} + p_i * c^{(t-1)} + b_i \right), \quad (8)$$

де  $w_i$ ,  $r_i$  та  $p_i$  – вагові коефіцієнти  $x^{(t)}$ ,  $y^{(t-1)}$  та  $c^{(t-1)}$ ,  $b_i$  – вектор зміщення.

Значення активації пропускового шлюзу обчислюються на основі поточного входу, виходів та стану клітинки пам'яті:

$$f^{(t)} = \sigma \left( w_f x^{(t)} + r_f y^{(t-1)} + p_f * c^{(t-1)} + b_f \right), \quad (9)$$

де  $w_f$ ,  $r_f$  та  $p_f$  – вагові коефіцієнти для  $x^{(t)}$ ,  $y^{(t-1)}$  та  $c^{(t-1)}$ ,  $b_f$  – вектор зміщення.

Значення клітинки пам'яті поєднує значення вхідного блоку, вхідного шлюзу, та значення пропускового шлюзу із попереднім значенням клітинки пам'яті:

$$c^{(t)} = z^{(t)} * i^{(t)} + c^{(t-1)} * f^{(t)}. \quad (10)$$

Вихідний шлюз поєднує поточний вхід, вихід цього блоку та значення клітинки пам'яті на останній ітерації:

$$o^{(t)} = \sigma \left( w_o x^{(t)} + r_o y^{(t-1)} + p_o * c^{(t-1)} + b_o \right), \quad (11)$$

де  $w_o$ ,  $r_o$  та  $p_o$  – вагові коефіцієнти для  $x^{(t)}$ ,  $y^{(t-1)}$  та  $c^{(t-1)}$ ,  $b_o$  – вектор зміщення.

Блок виходу поєднує поточне значення клітинки пам'яті та значення вихідного шлюзу:

$$y^{(t)} = g \left( c^{(t)} \right) * o^{(t)}. \quad (12)$$

Щоб знайти оптимальні ваги для використовуваних моделей, використовується алгоритм зворотного розповсюдження помилок.

Для функції активності  $f'$  виходом є:

$$o_j = f' \left( net_j \right), \quad (13)$$

Для похідної  $f'$  отримуємо:

$$\frac{\partial o_j}{\partial net_j} = f' \left( net_j \right); \quad (14)$$

$$\frac{\partial f}{\partial x} = \frac{-\alpha e^{-\alpha x}}{1 + e^{-\alpha x}} = \alpha f (1 - f); \quad (15)$$

$$\frac{\partial o_i}{\partial net_i} = \frac{-e^{-\alpha net_i}}{\left( 1 + e^{-\alpha net_i} \right)^2} = o_i (1 - o_i). \quad (16)$$

Таким чином, помилка в мережі на вихідному вузлу розраховується таким чином:

$$\delta_j = \left( t_j - o_j \right) * \alpha * o_j * \left( 1 - o_j \right). \quad (17)$$

Для обчислення помилки внутрішнього  $j$  шару:

$$\delta_j = o_j * \left( 1 - o_j \right) * \sum_k \delta_k w_{jk}, \quad (18)$$

де  $w_{jk}$  – вага зв'язку між вузлом внутрішнього шару та вузлом наступного шару.

Нові значення будуть обчислені згідно:

$$w'_{ij} = w'_{ij} + \Delta w_{ij} = w'_{ij} + \eta \delta_j o_i. \quad (19)$$

При проектуванні архітектури нейронної мережі необхідно враховувати, що кількість нейронів у прихованих шарах повинна залежати від обсягу навчальної вибірки.

Кількість прихованих шарів та кількість нейронів вибираються такими, щоб вони утворювали втричі менше зв'язків, ніж кількість навчальних прикладів.

Коефіцієнт навчання  $\eta$  – значення суттєво менше одиниці. Визначення коефіцієнта навчання згідно:

$$\eta = \frac{\max [x] - \min [x]}{N}. \quad (20)$$

Якщо значення  $\eta$  дуже мало, то налаштування вагових коефіцієнтів може займати багато часу для досягнення збіжності. Але великі значення призводять до пропусків глобальних екстремумів та передчасної збіжності.

Для отримання набору вагових коефіцієнтів необхідно виконати такі кроки:

1. Ініціалізація ваг з випадковими значеннями.
2. Для кожного внутрішнього вузла обчислити значення сигналу на вході та виході.
3. Обчислити значення сигналу на вході та виході вихідного вузла.
4. Якщо різниця між отриманим значенням сигналу у вихідному вузлі і реальним значенням не більше допустимого, то на цьому закінчити, інакше перейти до кроку 5.
5. Обчислити помилку та розрахувати нові вагові коефіцієнти.
6. Скоригувати величину відповідно до результату перевірки: чи збігається напрям змін прогнозованих значень з реальними значеннями.
7. Повторити кроки 2, 3, 4, 5, 6.

Обсяг навчальної вибірки залежить кількості входів штучної нейронної мережі та коефіцієнтів навчання.

### Результати порівняльного аналізу

В ході експериментальних досліджень виконано короткострокове прогнозування на вісім кроків уперед для часових рядів, що використовуються в M3-Competition та проведено порівняння отриманих даних з наявними результатами прогнозування за допомогою Vox-Jenkins, ARARMA, ForecastPro, Auto-ANN, SMARTFCS, RBF.

Прогнозування за допомогою розглянутих MLP та LSTM відбувалося по одному кроку вперед, приймаючи результати минулого кроку як частину вихідних даних.

Графічне відображення результатів прогнозування наведено на рис. 1.

Для оцінки короткострокового прогнозу часового ряду використовується значення середньої абсолютної помилки (MAE), та, оскільки для оцінки якості отриманого прогнозу доцільним є використання критеріїв, що дозволяють оцінити помилку у відсотковому відношенні, значення середньої абсолютної помилки у відсотках (MAPE).

Середня абсолютна помилка розраховується за формулою:

$$MAE = \sum_{i=1}^k |y_i - x_i| / k, \quad (21)$$

де  $y_i$  – фактичне значення,  $x_i$  – прогнозне значення,  $k$  – кількість прогнозованих значень.

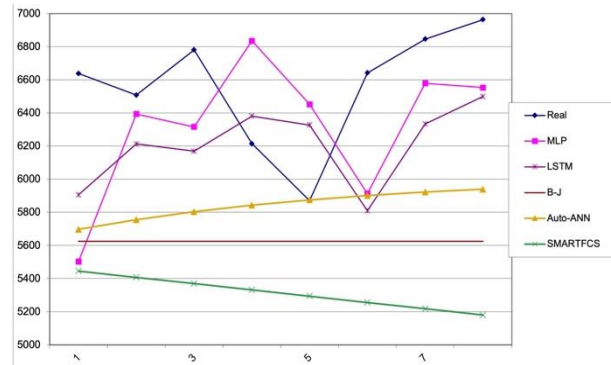


Рис. 1. Результати прогнозування

Середня абсолютна помилка у відсотках розраховується згідно:

$$MAPE = \frac{100\%}{k} \sum_{t=1}^k \left| \frac{A_t - F_t}{A_t} \right|, \quad (22)$$

де  $A_t$  – фактичне значення,  $F_t$  – прогнозне значення,  $k$  – кількість прогнозованих значень.

Середня абсолютна помилка та середня помилка у відсотках при використанні різних методів короткострокового прогнозування часових рядів наведена у табл. 1. Для аналізу були обрані з M3-Competition ряди різного походження довжиною від 39 до 72 значень. Суттєвого впливу кількості значень прогнозованого часового ряду на ефективність прогнозу не виявлено. Результати прогнозування часових рядів підтверджують суттєву перевагу використання моделей багатозарового перцептронів та довгої короткострокової пам'яті.

Таблиця 1 – Середня абсолютна помилка та середня помилка у відсотках при використанні різних методів короткострокового прогнозування часових рядів

Метод	Критерії	N700, (44)	N800, (44)	N900, (72)	N1000, (52)	N1100, (52)	N1200, (52)	N1300, (39)	N1400, (48)
B-J	MAE	932	353	111	175	402	123	470	232
	MAPE, %	8,3	4,85	1,73	2,54	3,81	0,46	1,63	0,92
ARARMA	MAE	1072	370	203	325	430	91	292	84
	MAPE, %	7,72	9,57	1,87	1,44	4,69	0,73	5,45	1,27
Auto-ANN	MAE	717	810	97	180	484	74	435	127
	MAPE, %	13,98	11,37	2,14	2,57	7,57	1,75	8,53	2,74
ForecastPro	MAE	984	233	109	350	448	43	470	227
	MAPE, %	16,11	11,63	3,93	4,84	7,94	1,3	5,49	0,99
SMARTFCS	MAE	1245	251	87	244	615	39	253	144
	MAPE, %	10,69	20,91	1,87	2,71	9,69	1,05	7,9	1,5
RBF	MAE	919	701	176	185	1056	50	340	228
	MAPE, %	14,77	5,51	2,11	5,2	7,91	0,62	8,54	2,68
MLP	MAE	540	174	89	171	201	51	118	78
	MAPE, %	18,75	7,73	1,7	3,64	10,84	0,56	4,88	1,7
LSTM	MAE	509	273	96	96	267	50	284	107
	MAPE, %	13,77	21,95	3,42	2,77	18,51	0,72	6,29	2,69

### Висновки

В роботі досліджено використання моделей багатозарового перцептронів та довгої короткостроков-

ої пам'яті для короткострокового прогнозування нестационарних часових рядів. Проведено порівняння отриманих результатів з наявними результатами прогнозування в M3-Competition.

Досліджувані моделі показали перевагу над засобами, що використовуються для прогнозування часових рядів M3-Competition.

З урахуванням того, що розглянуті ШНМ передбачають отримання прогнозу лише на один крок уперед (мають один нейрон у вихідному шарі), у процесі прогнозування з заданим горизонтом прогнозування послідовно використовувалися дані, які були отримані у якості результату на попередньому кроці, для використання їх на вході ШНМ. Це дозволяє досягти

меншої помилки, ніж при створенні прогнозу відразу на задану кількість значень.

Це також дозволяє забезпечити гнучкість використання ШНМ, оскільки не потребує перебудови архітектури при зміні потрібного горизонту прогнозування.

Середня помилка у відсотках при прогнозуванні часових рядів досліджуваних моделей на 2-6% нижче при порівнянні з використанням традиційних моделей.

#### СПИСОК ЛІТЕРАТУРИ

- Iorio C., Frasso G., Ambrosio A. (2016), "Parsimonious time series clustering using P-splines", *Expert Systems with Applications*, № 52, pp. 26-38.
- Yang J. (2006), "Power System Short-term Load Forecasting: Thesis for Ph.d degree", *Elektrotechnik und Informationstechnik der Technischen Universitat, Darmstadt, Germany*, P. 139.
- Korablev N.M., Ivaschenko G.S. (2014), "Parallel immune algorithm of short-term forecasting based on model of clonal selection", *Radio Electronics, Computer Science, Control*, № 2. pp. 73-78, doi: [doi.org/10.15588/1607-3274-2014-2-11](https://doi.org/10.15588/1607-3274-2014-2-11)
- Ivaschenko G.S., Korablev N.M. (2014), "Time series forecasting on the basis of the case-based reasoning using the models of artificial immune systems", *System technologies*, № 6(95), pp. 43-51, available at: [http://nbuv.gov.ua/UJRN/st\\_2014\\_6\\_8](http://nbuv.gov.ua/UJRN/st_2014_6_8)
- Ahmed N., Atiya A., Gayar N. (2010), "An Empirical Comparison of Machine Learning Models for Time Series Forecasting", *Econometric Reviews*, № 29, pp. 594-621.
- Makridakis S. (2000), "The M-3 Competition: Results, Conclusions and Implications", *International Journal of Forecasting*, № 16, pp. 451-476.
- Borghi P., Zakordonets O., Teixeira J. (2021), "A COVID-19 time series forecasting model based on MLP ANN", *Procedia Computer Science*, № 181, pp. 940-947.
- Namini S., Tavakoli N., Namin A. (2018), "A Comparison of ARIMA and LSTM in Forecasting Time Series", *17th IEEE International Conference on Machine Learning and Applications*.
- Cao J., Li Z., Li J. (2019), "Financial time series forecasting model based on CEEMDAN and LSTM", *Physica A: Statistical Mechanics and its Applications*, № 519, pp. 127-139.
- Vasily D., Matviychuk A., Datsenko N., Bezkorovainyi V., Azaryan A. (2020), "Machine learning approaches for financial time series forecasting", *Proceedings of the Selected Papers of the Special Edition of International Conference on Monitoring, Modeling & Management of Emergent Economy*, № 2713, Odessa, Ukraine, pp. 434-450.
- Hemmati A., Abdoos M., Akbar A. (2015), "Short term load forecasting using a hybrid intelligent method", *Knowledge-Based System*, № 76, pp. 139-147.
- Mohammed A.S., Ivashchenko H., Filimonchuk T., Ivanisenko I., Barkovska O. (2020), "Green Hybrid Models Based on Clonal Selection and Case-based Reasoning for Short-term Time Series Forecasting", *Journal of Green Engineering*, Vol. 10, Issue 5. pp. 2139-2154, available at: <http://www.jgenn.com/volume10-issue5.php>
- Hamzacebi C., Akay D., Kutay F. (2009), "Comparison of direct and iterative artificial neural network forecast approaches in multi-periodic time series forecasting", *Expert Systems with Applications*, № 36, pp. 3839-3844.
- Voyant C., Nivet M., Paoli C., Muselli M., Notton G. (2014), "Meteorological time series forecasting based on MLP modelling using heterogeneous transfer functions", *Journal of Physics: Conference Series*, № 574, Madrid, Spain, pp. 28-31.
- Greff, K., Srivastava, R., Koutnik, J., Steunebrink, B., Schmidhuber, J. (2017), "LSTM: A Search Space Odyssey", *IEEE Transactions on Neural Networks and Learning Systems*, № 28, pp. 2222-2232.
- Su Y., Kuo C. (2019), "On extended long short-term memory and dependent bidirectional recurrent neural network", *Neurocomputing*, № 356, pp. 151-161.

Received (Надійшла) 09.01.2023

Accepted for publication (Прийнята до друку) 01.03.2023

#### Short-term of non-stationary time series forecasting using MLP and LSTM models

Heorhii Ivashchenko, Vladyslav Ponamarov, Vladyslav Kholiev

**Abstract. Topicality.** The solution of the problem of forecasting takes an important role in the processes of strategic planning and operational management in various areas of economic activity. A widespread form of forecasting is time series forecasting, in which the actual problem is the choice of an appropriate method among modern means of computational intelligence, such as artificial neural networks. The problem of choice is due to a large number of parameters and settings that depend on the features of the predicted time series and have a significant impact on the forecast quality. **The goal of this work** is to analyze methods for short-term forecasting of non-stationary time series using artificial neural network models such as multilayer perceptron and long short-term memory. **The object of research** is the process of time series forecasting. **The subject of research** is the use of artificial neural network models for short-term forecasting. **Results.** Experimental research has shown that the average error in forecasting using the proposed tools is 2-6% lower compared to using common traditional models. **Conclusions.** The obtained results confirm the advantage of using the MLP and LSTM models over forecasting based on the methods chosen for the M3-Competition analysis.

**Keywords:** time series forecasting, computational intelligence, artificial neural network, multilayer perceptron, long short-term memory, M3-Competition