

Долгих А.О.

Дніпровський національний університет імені Олеся Гончара

Байбуз О.Г.

Дніпровський національний університет імені Олеся Гончара

ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ АНСАМБЛЕВОГО ПРОГНОЗУВАННЯ З ВИКОРИСТАННЯМ ПАРАЛЕЛЬНИХ ОБЧИСЛЕНЬ

Сьогодні ансамблі моделей прогнозування є одним із перспективних напрямів розв'язання задач, де прогностична точність важливіша, ніж можливість інтерпретувати модель. Головна ідея ансамблів полягає в навчанні кількох базових моделей та агрегації результатів їх роботи. Емпіричні дослідження показують, що комбінації прогнозів у середньому частіше дають кращі прогнози, ніж методи, які базуються на виборі лише однієї моделі прогнозування. При побудові ансамблів особливої актуальності набуває питання забезпечення різноманітності моделей та ефективного навчання моделей-членів ансамблю. Стаття присвячена розв'язанню цих питань.

У роботі набув подальшого розвитку метод побудови ансамблів моделей прогнозування часових рядів. Запропонований підхід включає побудову таких моделей: моделі авторегресії-ковзкого середнього, моделі нейронної мережі, моделі сингулярного спектрального розкладання, адаптивних моделей, моделі групового урахування аргументів і моделі нечіткої логіки. Для організації навчання та оцінки моделей ансамблю використовуються засоби організації паралельних обчислень фреймворку .NET. Відбір моделей для побудови ансамблевого прогнозу здійснюється за допомогою багатокритеріальної методики. Вона дає змогу оцінювати як точність прогнозування на навчальній часовій послідовності, так і результати аналізу залишків. Крім цього, описано реалізований на базі цих розробок програмний комплекс. Надано діаграму класів і розкриті деталі реалізації програмного продукту.

Засобами розробленого програмного забезпечення був проведений аналіз часових рядів фінансової природи. Показано, що використання запропонованого ансамблевого методу дає змогу підвищити точність прогнозування та зменшити час, необхідний для тренування й оцінки моделей-членів ансамблю. Сформульовані напрями та перспективи подальших досліджень.

Ключові слова: часові ряди, ансамблі моделей прогнозування, адаптивні моделі, нейронна мережа, ARIMA, SSA, МГУА, паралельні обчислення, якість прогнозування.

Постановка проблеми. Сьогодні задача прогнозування часових рядів набуває особливої актуальності в різних областях людської діяльності: у природничих науках – для прогнозування кількості опадів, стану забруднення водних ресурсів, деяких біологічних і біохімічних показників; в економіці – для прогнозування щоденних біржових котирувань, курсів валют, щотижневих і щомісячних обсягів продажів, річних обсягів виробництва, тощо.

Незважаючи на велике різноманіття моделей прогнозування, яке існує сьогодні [1, с. 20; 2, с. 8], завжди існує вірогідність обрати модель, неприйнятну для досліджуваного процесу. Навіть пошук серед моделей, найбільш придатних для певних вузьких підмножин часових рядів, може не дати бажаних результатів: «Although forecasting expertise can be found in the literature, these sources often fail to adequately describe conditions under which a method

is expected to be successful» [3, с. 2]. Щоб уникнути ризику вибору невідповідного методу, сьогодні часто використовують ансамблі моделей прогнозування. У зв'язку з цим особливої актуальності набуває питання розробки ефективних алгоритмів ансамблевого прогнозування.

Аналіз останніх досліджень і публікацій.

Головна ідея ансамблів полягає в навчанні кількох базових моделей та агрегації результатів їх роботи. Емпіричні дослідження показали, що комбінації прогнозів у середньому частіше дають кращі прогнози, ніж методи, які базуються на виборі лише однієї моделі прогнозування [3, с. 3].

При побудові ансамблів особливої актуальності набуває питання забезпечення різноманітності моделей [4, с. 1–2; 5, с. 1]. Зазвичай його вирішують шляхом побудови досить великої кількості моделей із різних класів [6, с. 2–3; 7, с. 14]. У таких умовах актуальним напрямом дослідження є використання

паралельних обчислень для тренування та оцінки моделей прогнозування. Це дає змогу збільшити швидкість побудови моделей прогнозування, ефективно розподіляти ресурси комп'ютера, підвищити масштабованість програми. Окрім цього, певні задачі, такі як аналіз і прогнозування фінансових часових рядів у ризик- і портфоліо-менеджменті, можуть вимагати виконання великої кількості обчислень для досягнення потрібної точності. Застосування паралелізму в цих випадках вважається новим, перспективним напрямом обчислювальної математики [8, с. 123].

Постановка завдання. Комбінований прогноз при побудові ансамблів – це лінійна комбінація окремих прогнозів, отриманих різними моделями. Припустимо, що ми маємо сукупність M моделей прогнозування. Для заданого часового ряду $\{u_t; t=1, N\}$ і певного прогнозного горизонту h , j -й метод прогнозування дає такий прогноз:

$$y_{N+1}^{(j)}, y_{N+2}^{(j)}, \dots, y_{N+h}^{(j)}, j = \overline{1, M}, \quad (1)$$

тоді комбінований прогноз будують за такою формулою:

$$y_{N+i} = \frac{1}{\sum_{j=1}^M w_j} \cdot \sum_{j=1}^M w_j \cdot y_{N+i}^{(j)}, i = \overline{1, h}, \quad (2)$$

де $w_i \in [0, 1]$ – вага певного методу, $\sum_{j=1}^M w_j = 1$.

Тренування моделей прогнозування, що входять до складу ансамблю, можна розглядати як колекцію незалежних одна від одної задач, які добре підлягають паралелізації. Тому проведені дослідження мали за меті:

- побудувати достатньо широкий спектр моделей часових рядів з різних класів;
- поширити різноманіття ансамблю за рахунок зміни параметрів моделей;
- прискорити час, необхідний для тренування моделей і їх оцінки, за допомогою паралельних обчислень;
- дослідити ефективність застосування такого підходу й порівняти з результатами, отриманими при використанні звичайних, послідовних обчислень;
- оцінити якість побудованої моделі й точність отриманих прогнозів.

Виклад основного матеріалу. *Реалізовані моделі часових рядів.* Як зазначалося раніше, при побудові ансамблю важливо забезпечити достатнє різноманіття моделей. Згідно з роботою [6, с. 1], використання методів з різних класів дає змогу підвищити якість прогнозування. У процесі виконання дослідження розроблений програмний комплекс на платформі .NET, який містить реалізацію адаптивних моделей прогнозування, моделей авторегресії-ковзкого середнього, нейронних мереж, нечіткої логіки, спектрального сингулярного аналізу та групового урахування аргументів. Це досить потужні методи прогнозування, які добре зарекомендували себе в аналізі й прогнозуванні різних за природою та складністю процесів. Нижче наведено їх короткий опис.

Адаптивні моделі. При використанні адаптивних моделей часовий ряд подають у вигляді функції:

$$u_t = f(a_{1t}, a_{2t}, \dots, a_{pt}, t) + e_t, \quad (3)$$

під час побудови якої відслідковують величину відхилень прогнозних значень від реальних [9, с. 50]. Для ідентифікації оптимальних значень параметрів цих моделей, так званих коефіцієнтів згладжування, був використаний алгоритм, описаний у роботі [10, с. 16–17].

Модель ARIMA. За підходу ARIMA, поінтегрованої моделі авторегресії-ковзкого середнього, часовий ряд представляють у вигляді 4:

$$\Delta^d X_t = c + \varepsilon_t + \sum_{i=1}^p \alpha_i \Delta^d X_{t-i} + \sum_{j=0}^q b_j \varepsilon_{t-j}, \quad (4)$$

де c, α_i, b_j – параметри моделі, Δ^d – оператор різниці часового ряду порядку d , p – порядок авторегресії, q – порядок ковзкого середнього, ε_t – незалежні, нормально розподілені похибки з нульовим середнім та однаковою дисперсією. Головною проблемою, яка виникає при роботі із цими моделями, є задача підбору оптимальних значень порядку інтегрування, авторегресії та ковзкого середнього. У розроблюваному авторами програмному комплексі ця проблема для розв'язку цієї задачі використовувалися процедури, описані в роботі [11, с. 10–11].

Модель нейронної мережі (ANN). У ході виконання дослідження була реалізована модель нейронної мережі, запропонована в праці [12, с. 131–133]. За такого підходу на вхід мережі подають $u_{t-1}, u_{t-2}, \dots, u_{t-p}$ значення ряду, де p – кількість попередніх значень ряду, які використовуються для обчислення прогнозів. Програмою автоматично обирається таке значення p , за якого значення інформаційного критерію Акаїке, AIC , є найменшим із можливих. Навчання мережі проводилося з використанням алгоритму Бройдена – Флетчера – Гольдфарба – Шанно (BFGS) [13]. Програмне забезпечення дозволяє побудувати M нейронних мереж із різними початковими значеннями ваг і зсувів нейронів і видає як прогнозу середнє результатів, отриманих кожною з них.

Модель сингулярного спектрального розкладання (SSA). Модель SSA належить до потужних

непараметричних методів аналізу часових рядів. Базовий алгоритм методу складається з двох етапів, які доповнюють один одного: розкладання вихідного часового ряду на адитивні складники та його відновлення. Модифікації базового алгоритму дають змогу побудувати прогноз відносно майбутніх значень ряду на один або декілька кроків уперед [9, с. 62–63].

Модель нечіткої логіки (FL). При використанні нечіткої логіки в аналізі даних будується база правил нечітких продукцій, за допомогою якої можна отримувати прогнози стосовно значень досліджуваного процесу. У ході виконання дослідження реалізована процедура нечіткого виводу, описана в роботі [14, с. 3–6]. Вона включає формування універсуму, тобто множини всіх можливих значень ряду:

$$U = [V_{min} - D_1; V_{max} + D_2], \quad (5)$$

де V_{max} та V_{min} – максимальне та мінімальне значення ряду, D_1 та D_2 – додатні константи та поділ його на m інтервалів однакової довжини.

Після цього проводять процедуру фазифікації, тобто кожне значення часового ряду перетворюють у «нечіткий» вигляд. Для цього розраховують значення функції належності рівню до кожного з інтервалів універсуму. Далі проводять виведення «нечітких» прогнозних значень ряду. На останньому етапі, етапі дефазифікації, переходять від нечіткої форми до «чітких» прогнозів.

Модель групового урахування аргументів (GMDH). При використанні моделей групового урахування аргументів найчастіше використовують поліноми форми 6:

$$y = a + \sum_{i=1}^m b_i x_i + \sum_{i=1}^m \sum_{j=1}^m c_{ij} x_i x_j, \quad (6)$$

де m – число змінних, a, b, c, d, \dots – це коефіцієнти змінних у поліномі, які також називають вагами, Y – це майбутнє значення ряду, x_i та x_j – попередні значення часового ряду.

Останнього часу все більшого поширення набувають нейронні мережі типу МГУА. При їх використанні в прогнозуванні часових рядів розглядають усі попарні поєднання попередніх значень ряду із затримкою (лагом) P . Кожну таку комбінацію подають на вхід нейронів і проводять навчання мережі, під час якого знаходять оптимальні моделі нейронів.

Процес навчання такої нейронної мережі складається з обчислення оптимальних значень коефіцієнтів нейронів і вибору їх функцій активації. У програмному комплексі використовуються нейронні мережі МГУА, описані в роботі [15, с. 380–382].

Використання паралельних обчислень для побудови ансамблю

Програмне забезпечення для побудови ансамблю прогнозування розроблено на мові C# з використанням можливостей фреймворку .NET. За допомогою інтерфейсу програми користувач має змогу змінювати типи та кількість моделей, які входять до складу ансамблю. За замовчуванням, до ансамблю входять моделі з усіх класів, описаних вище.

Після того, як склад ансамблю сформований, проводиться навчання моделей прогнозування та їх оцінка. Блок-схему алгоритму навчання моделей проілюстровано на рис. 1. Реалізований підхід використовує паралелізм задач.

Починаючи з версії 4 .NET-фреймворку, найкращим способом застосування паралельних обчислень у C#-програмі вважається бібліотека паралельних задач TPL, яка являє собою набір загальнодоступних типів у просторах імен System.Threading і System.Threading.Tasks. Її використання дає змогу підвищити ефективність і масштабованість коду, водночас спрощуючи процес додавання паралелізму в програму для розробника.

Для навчання моделей ансамблю та їх оцінки використовувався метод Parallel.ForEach<TSource>(IEnumerable<TSource>, Action<TSource>) із простору імен System.Threading.Tasks. Він дає проводити тренування й оцінку якості моделей ансамблю паралельно за допомогою пулу потоків System.Threading.ThreadPool. Перевагами

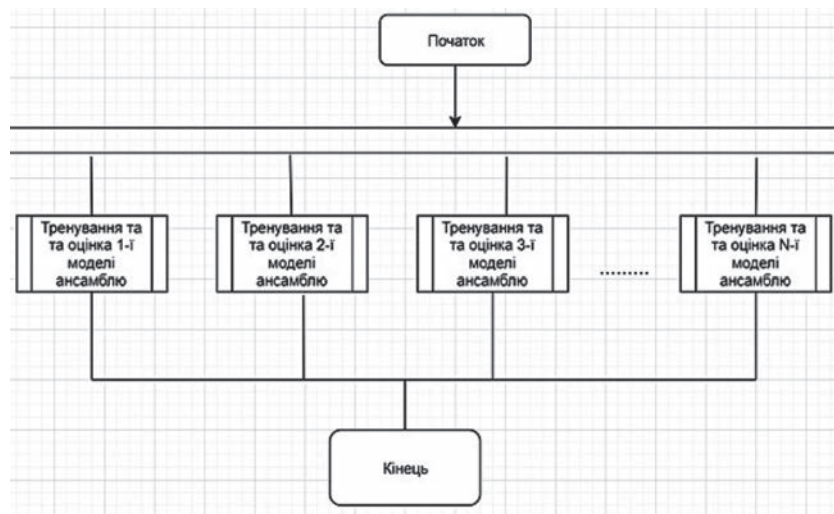


Рис. 1. Алгоритм навчання моделей ансамблю

використання `Parallel.ForEach` є те, що цей підхід, окрім забезпечення паралельності обчислень, ефективно розподіляє список моделей, які потрібно натренувати, між робочими потоками, тобто нові потоки створюються лише за необхідності, а не автоматично окремий потік для кожного завдання. Це може забезпечити значно кращі показники загальної продуктивності, коли кількість моделей в ансамблі невелика [16].

Моделі, реалізовані в програмному комплексі, належать до єдиної ієрархії та є наслідниками класу `GeneralForecastingModel`. Діаграму класів розробленого комплексу наведено на рис. 2.

Для навчання моделей використовується віртуальний метод `FitModel`. Його базова реалізація, що знаходиться у класі `GeneralForecastingModel`, є порожньою. Кожна з моделей-нащадків реалізує власний алгоритм навчання. Для забезпечення виконання тренування моделей і їх оцінки в правильному порядку використовувалася технологія `Continuation Task`, або задачі продовження. Задачі продовження – це асинхронні задачі, які викликаються попередніми задачами після свого завершення. Задачі продовження створюються за допомогою методу `ContinueWith` класу `Task` [17]. Як попередні задачі використовуються методи навчання моделей, а задачі продовження – це методи оцінки їх якості. Оцінка якості моделей проводиться з використанням критерію Акаїке й

аналізу залишків моделі. Після цього кожна модель прогнозування отримує певний рейтинг за процедурою, описаною в роботі [18, с. 133–135]. Моделі з найвищим рейтингом, за замовчуванням, 10% з усіх побудованих, беруть участь у побудові прогнозів щодо майбутніх значень ряду. Для агрегації результатів використовується середнє арифметичне прогнозів, отриманих кожною моделлю окремо.

Результати роботи запропонованого алгоритму побудови ансамблю моделей прогнозування. Тестування розробленого програмного забезпечення проводилося на фінансових часових рядах, які являють собою щоденні коливання цін на акції американських компаній у період із січня 2017 року по січень 2018 року, розташованих на ресурсі фінансової інформації «Yahoo!» [19].

У поточному розділі наведені результати апробації запропонованого підходу для прогнозування щоденних цін на акції компанії `Diebold Incorporated, DBD` на момент відкриття біржі (рис. 3). `Diebold Incorporated` є виробником комплексних систем самообслуговування й безпеки для банків і фінансових інститутів. Попередній аналіз часового ряду свідчить про те, що в ряді наявна тенденція до спадання та періодичний складник. Для перевірки ряду на наявність тренду використовувався ранговий критерій Спірмена, а для тестування наявності сезонної компоненти – критерії екстремальних точок і серіальної кореляції.

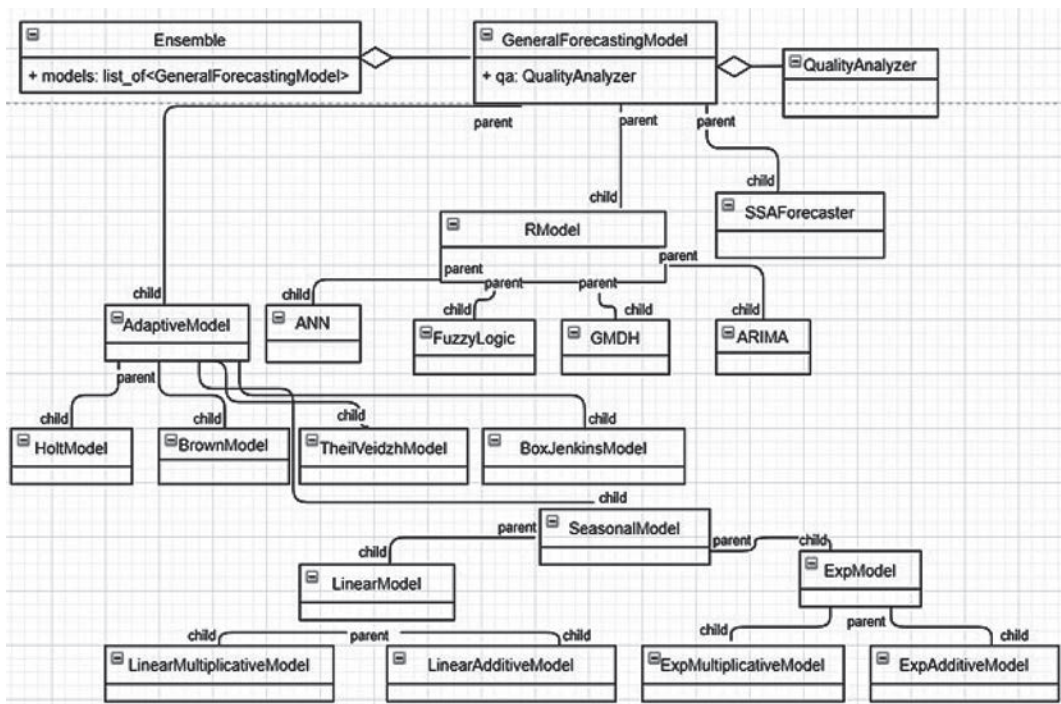


Рис. 2. Діаграма класів розробленого програмного комплексу

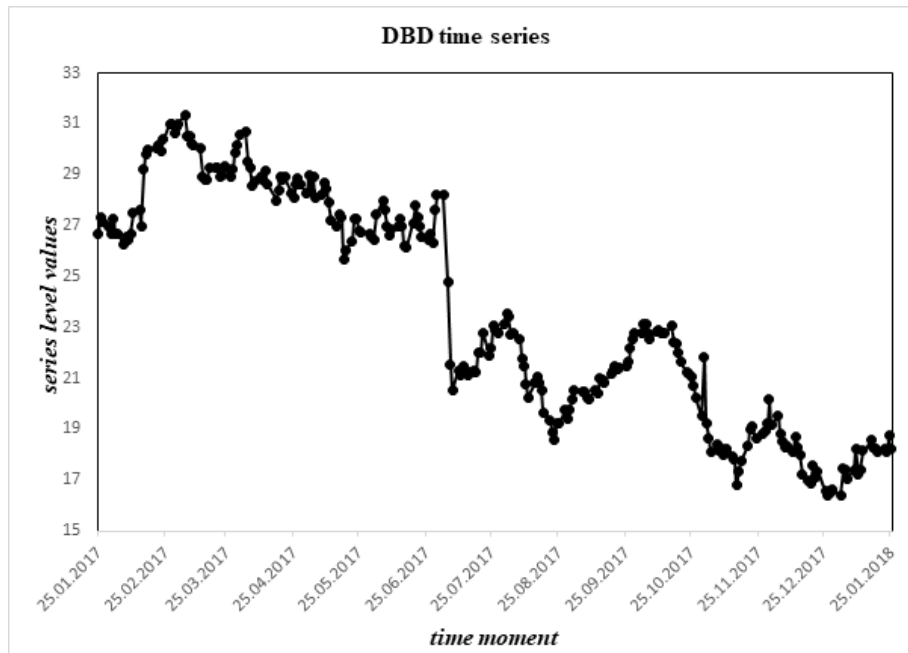


Рис. 3. Часовий ряд цін на акції компанії DBD

Таблиця 1

Середній час побудови ансамблю при використанні паралельних і послідовних обчислень для ряду DBD

Кількість моделей	$T_1(n)$	$T_p(n)$	p	$S_p(n)$	$E_p(n)$
25	27766,2	4612,6	28	6,0196	0,2149
50	46475,8	7597,6	29	6,1171	0,2109
75	67669,2	10766,8	27	6,2877	0,2328
100	110879,4	13568,2	29	8,1720	0,2817
125	115078,4	15799,2	29	7,2838	0,2511

Для прогнозування майбутніх рівнів ряду застосовано ансамблевий підхід. Результати використання паралельного алгоритму побудови ансамблю порівнюються з результатами послідовного підходу за допомогою таких характеристик, як прискорення й ефективність (таблиця 1, рис. 4).

Прискорення, яке ми отримуємо при використанні паралельного алгоритму для p процесорів, у порівнянні з послідовним варіантом виконання обчислень визначається формулою 7:

$$S_p(n) = \frac{T_1(n)}{T_p(n)}, \quad (7)$$

де $T_1(n)$ – час рішення задачі на скалярній електронно-обчислювальній машині, $T_p(n)$ – час виконання паралельного алгоритму.

Ефективність використання паралельним алгоритмом процесорів обчислюється за формулою:

$$E_p(n) = \frac{S_p(n)}{p}, \quad (8)$$

де $S_p(n)$ – прискорення, p – кількість процесорів. Величина ефективності визначає середню частку

часу виконання алгоритму, протягом якої процесори реально використовуються для вирішення завдання.

Тестування проводилося на комп'ютері з оперативною пам'яттю 8Гб та процесором Intel(R) Core(TM) i5-6200U з частотою 2.40 Гц, двома ядрами й чотирма логічними процесорами.

З таблиці 1 і рис. 4 видно, що використання паралельних обчислень дає змогу зменшити час, необхідний для тренування моделей, у 6–8 разів. При цьому в разі збільшення кількості моделей прогнозування в ансамблі збільшується прискорення й ефективність запропонованого методу.

Під час проведення експерименту прогнози, отримані прогнозування з використанням ансамблю, порівняні з результатами роботи окремих моделей, а саме: ARIMA, ANN, FL, GMDH, SSA – та адаптивною моделлю Хольта. У таблиці 2 наведено прогнозні значення ряду, побудовані реалізованими моделями, і реальні значення ряду для горизонту прогнозування $h = 1..5$. У таблиці 3 показано, як змінюється значення середньоквадратичної похибки прогнозу, MSE залежно від горизонту прогнозування для ансамблю й окремих його членів.

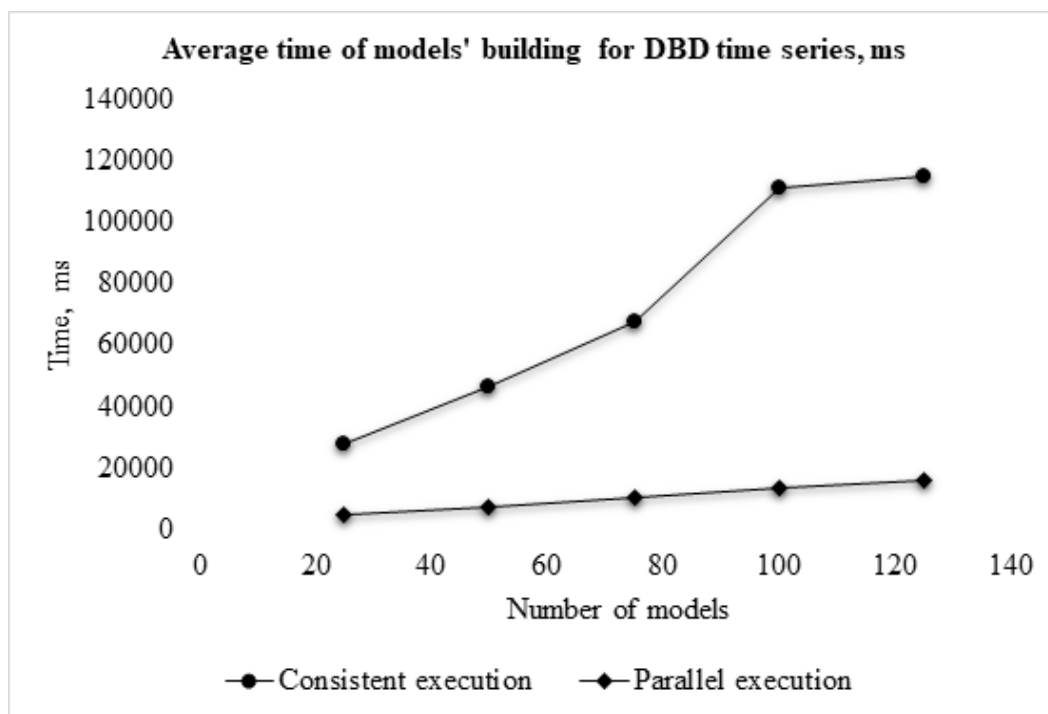


Рис. 4. Середньочасові затрати на побудову ансамблю прогнозування з використанням послідовного та паралельного підходів

Таблиця 2

Прогнозні значення ряду, отримані ансамблем моделей і кожною моделлю окремо

Модель/прогнознi значення	h = 1	h = 2	h = 3	h = 4	h = 5
ARIMA	18,2165	18,1829	18,1494	18,1159	18,0823
Модель FL	17,725	17,2	16,675	16,15	15,625
Модель ANN	18,2707	18,2898	18,3076	18,324	18,3393
Модель GMDH	18,4502	18,4542	18,5216	18,5798	18,6375
Модель SSA	17,9079	17,6457	17,379	17,1079	16,8326
Модель Хольта	18,2627	18,2163	18,1009	18,1655	18,1401
Ансамбль прогнозування	18,256	18,1919	18,1774	18,1818	18,1469
Реальні значення ряду	18,1	18,05	18,55	18,5	18,3

Таблиця 3

Порівняння якості прогнозування з використанням середньоквадратичної похибки, MSE

Модель/значення MSE	h = 1	h = 2	h = 3	h = 4	h = 5
ARIMA	0,0136	0,0156	0,0639	0,0848	0,0773
Модель FL	0,1406	0,4316	1,4596	2,4753	3,4114
Модель ANN	0,0291	0,0433	0,0485	0,0441	0,0356
Модель GMDH	0,1226	0,143	0,0956	0,0733	0,0814
Модель SSA	0,0369	0,1002	0,5239	0,8774	1,1326
Модель Хольта	0,0265	0,0271	0,0853	0,0919	0,0787
Ансамбль прогнозування	0,0243	0,0222	0,0611	0,0711	0,0616

З приведених результатів видно, що при використанні ансамблю моделей у більшості випадків вдається досягти підвищення точності прогнозування. Однак існують випадки, коли при використанні окремої моделі, наприклад, ANN для гори-

зонтів $h = 3$, $h = 4$, $h = 5$ та ARIMA для $h = 1$, $h = 2$, похибка прогнозу є меншою, ніж при використанні агрегованого прогнозу. У таблиці 4 показано, наскільки вдалося у відсотках зменшити значення MSE при використанні ансамбле-

вого підходу в порівнянні з окремими моделями. На рис. 5 зображено реальні значення ряду та прогнози, побудовані за допомогою ансамблевого підходу.

Фактичні значення ряду для всього горизонту прогнозування $h = \overline{1..5}$ знаходяться в межах довірчого інтервалу, що свідчить про якість побудованої ансамблевої моделі.

Висновки. При побудові ансамблю прогнозування особливо важливим стає питання забезпечення різноманіття моделей, що входять до його складу. У роботі запропоновано розширити спектр моделей за рахунок побудови моделей різних класів і зміни їх параметрів. У ході виконання дослідження реалізовані такі моделі часових рядів: адаптивні, SSA, FL, GMDH, ANN та ARIMA. Для пришвидшення виконання обчислень, необхідних для побу-

дови ансамблю, розроблено алгоритм паралельного навчання й оцінювання моделей із використанням бібліотеки TPL .NET-фреймворку. Він дає змогу істотно скоротити час, необхідний для побудови ансамблю. З аналізу результатів роботи запропонованого методу випливає, що в більшості випадків комбінований прогноз дає більш точні результати, ніж результати окремих моделей. Однак існують також ситуації, коли результати окремої моделі показують менше значення похибки прогнозу. У зв'язку з цим як напрями майбутніх досліджень можна виділити такі:

- дослідження нових методів відбору моделі прогнозування для заданого часового ряду;
- розробка засобів агрегації результатів прогнозування, отриманих різними моделями.

Таблиця 4

Зменшення *MSE* у відсотках при побудові прогнозів із використанням ансамблю й окремих моделей

Модель/кількість прогнозів	h = 1	h = 2	h = 3	h = 4	h = 5
ARIMA	N/A	N/A	4.3818	16.1557	20.3105
FL	82,7169	94,8564	95,8139	97,1276	98,1943
ANN	16,4949	48,7298	N/A	N/A	N/A
GMDH	80,1795	84,4756	36,0879	3,0013	24,3243
SSA	34,1463	77,8443	88,3375	91,8965	94,5612
Модель Хольта	8,3019	18,0812	28,3705	22,6333	21,7281

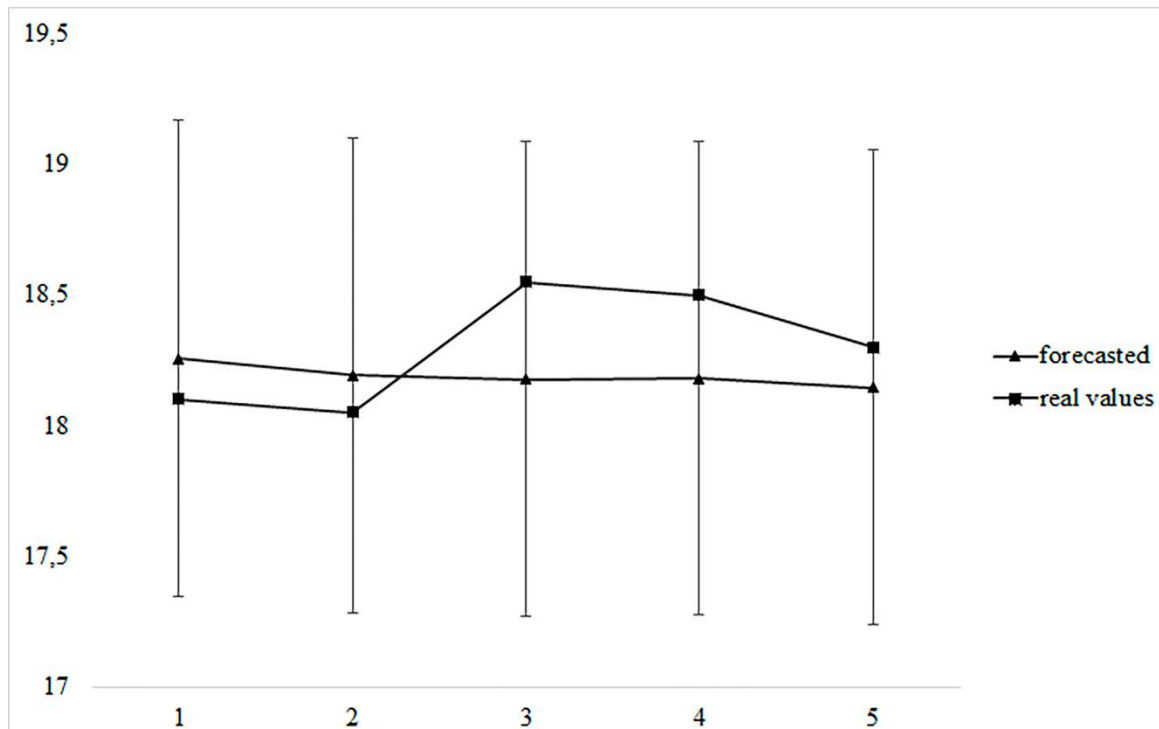


Рис. 5. Фактичні та прогнозні значення ряду DBD для горизонту прогнозування $h = \overline{1..5}$

Список літератури:

1. Чучуєва І.А. Модель прогнозування часових рядів за вибіркою максимальної правдо подоби : дис. ... канд. техн. наук : 05.13.00 / Московський державний технічний університет. Москва, 2012. 154 с.
2. Тихонов Э.Е. Методы прогнозирования условий рынка : учебное пособие. Невинномиск : Северокавказский государственный технический университет, 2006. 221 с.
3. Armstrong J.S., Adya M., Collopy F. Rule-based forecasting: using judgement in time series extrapolation. *Principles of Forecasting: A Handbook for Researchers and Practitioners* / Ed. J.S. Armstrong. Norwell. MA : Kluwer Academic Publishers, 2001. P. 257–283.
4. David F. Hendry, Michael P. Clements. Pooling of forecasts. *Econometrics Journal*. 2004. Vol. 7. P. 1–31.
5. Mariana Oliveira, Luis Torgo. Ensembles for Time Series Forecasting. *Proceedings of Machine Learning Research: Workshop and Conference Proceedings*. 2015. Vol. 39. P. 360–370.
6. Wichard J.D., Ogorzalek M. Time series prediction with ensemble models. *Neurocomputing*. 2007. Vol. 70. № 13–15. P. 2371–2378.
7. David Shaub, Peter Ellis. Package 'forecastHybrid'. URL: <https://cran.r-project.org/web/packages/forecastHybrid/forecastHybrid.pdf> (дата звернення: 25.01.2019).
8. Норкін Б.В. Системний імітаційний аналіз та оптимізація страхового бізнесу. *Кібернетика і системний аналіз*. 2014. Т. 52. № 2. С. 112–125.
9. Білобородько О.І., Ємел'яненко Т.Г. Аналіз динамічних рядів : навчальний посібник. Дніпро : РВВ ДНУ, 2014. 80 с.
10. Долгіх А.О., Білобородько О.І., Байбуз О.Г. Знаходження оптимальних значень параметрів адаптивних моделей прогнозування часових рядів з використанням генетичного алгоритму. *Актуальні проблеми автоматизації та інформаційних технологій*. 2016. № 20. С. 11–22.
11. Hyndman R.J., Khandakar Y. Automatic Time Series Forecasting: The forecast Package for R. *Journal of Statistical Software*. 2008. Vol. 27. № 3.
12. Hyndman R.J. Forecasting: Principles and Practice. University of Western Australia, 2014.
13. Борисов Э.С. Про методи навчання багатосарових нейронних мереж прямого розповсюдження. Частина 3 : Градієнтні методи другого порядку. URL: <http://mechanoid.su/neural-net-backprop3.html> (дата звернення: 25.01.2019).
14. Abbasov A.M., Mamedova M.H. Application of fuzzy time series to population forecasting. *CORP*. 2003.
15. Dag O., Yozgatligil C. GMDH: An R Package for Short Term Forecasting via GMDH-Type Neural Network Algorithms. *The R Journal*. 2016. Vol. 8. № 1. P. 379–386.
16. Partitioner Class. URL: <https://docs.microsoft.com/en-us/dotnet/api/system.collections.concurrent.partitionerr?view=netframework-4.8> (дата звернення: 31.10.2019).
17. Chaining Tasks by Using Continuation Tasks. URL: <https://docs.microsoft.com/en-us/dotnet/standard/parallel-programming/chaining-tasks-by-using-continuation-tasks> (дата звернення: 31.10.2019).
18. Долгіх А.О., Байбуз О.Г. Порівняльний аналіз ефективності моделей прогнозування часових рядів за допомогою багатокритеріальної процедури на прикладі фінансових показників. *Вісник Житомирського державного політехнічного університету. Серія «Технічні науки»*. 2019. № 1 (83). С. 131–141.
19. YAHOO! Finance. URL: <https://finance.yahoo.com/> (дата звернення: 31.10.2019).

Dolhikh A.O., Baibuz O.G. INFORMATION TECHNOLOGY OF ENSEMBLES FORECASTING USING PARALLEL COMPUTING

Nowadays, ensembles of forecasting models are one of the promising areas of task solving, where predictive accuracy is more important than the ability to interpret the model. The main idea of ensembles is to teach several basic models and to aggregate the results of their work. Empirical studies show that combinations of forecasts on average often give better forecasts than methods based on the choice only one forecasting model. In the construction of ensembles, the issue of ensuring the diversity of models and the issue of effective training of ensemble members become especially actual. The article is devoted to solving these issues.

The method of building ensembles of time series forecasting models has been further developed in the current research. The proposed approach includes the construction of the following models: autoregressive integrated moving average model, the neural network model, the singular spectrum analysis model, adaptive models, the group method of data handling and the fuzzy logic model. To organize the training and quality evaluation of ensemble models, the means of organizing parallel calculations of the .NET framework have been used. The selection of models for the construction of the ensemble forecast is carried out using multi-criteria method. It allows to assess both the accuracy of forecasting on the training time sequence, and the results of the residual analysis. In addition, the article describes the software package implemented on the basis of these developments. The class diagram and details of realization of a software product have been presented.

By means of the developed software the analysis of time series of the financial nature has been carried out. It is shown that the use of the proposed ensemble method allows to increase the accuracy of forecasting and to reduce the time required for training and evaluation of ensemble member models. Directions and prospects of further research have been formulated.

Key words: time series, ensembles of forecasting models, adaptive models, artificial neural network, ARIMA, SSA, GMDH, parallel computing, forecasting quality.