

УДК 004.032.26

ПОБУДОВА МОДЕЛЕЙ ПРОГНОЗУВАННЯ НА БАЗІ ПАКЕТУ MATLAB NEURAL NETWORKS

СКНАР І. І.

УДК 004.032.26

Скнар І. І. Побудова моделей прогнозування на базі пакету Matlab Neural Networks

Дослідження спрямоване на розкриття можливостей використання пакету Matlab Neural Networks для побудови прогнозуючих моделей економічних і соціальних процесів.

Ключові слова: імітаційна модель, нейронні мережі, прогнозування, смертність немовлят, фондові показники, стан економіки, рівень життя.

Рис.: 7. Табл.: 1. Бібл.: 5.

Скнар Ілля Іванович – студент, Криворізький економічний інститут Київського національного економічного університету ім. В. Гетьмана (вул. К. Маркса, 64, Кривий Ріг, Дніпропетровська обл., 50000, Україна)

E-mail: ilayya@mail.ru

УДК 004.032.26

UDC 004.032.26

Скнар И. И. Построение моделей прогнозирования на базе пакета Matlab Neural Networks

Sknar I. I. The Construction Of Prediction Models Based on the Package Matlab Neural Networks

Исследование направлено на раскрытие возможностей использования пакета Matlab Neural Networks для построения прогнозных моделей экономических и социальных процессов.

The research focused on the potentialities of the package Matlab Neural Networks to build predictive models of economic and social processes.

Ключевые слова: имитационная модель, нейронные сети, прогнозирование, младенческая смертность, фондовые показатели, состояние экономики, уровень жизни.

Key words: simulation model, neural networks, prognosis, infant mortality, financial highlights, state of the economy, living standard.

Рис.: 7. Табл.: 1. Библ.: 5.

Рис.: 7. Табл.: 1. Библ.: 5.

Скнар Илья Иванович – студент, Криворожский экономический институт Киевского национального экономического университета им. В. Гетьмана (вул. К. Маркса, 64, Кривой Рог, Днепропетровская обл., 50000, Украина)

Sknar Ilya I. – Student, Kryvyi Rig Economic Institute of the Kiev National Economic University named after V. Getman (vul. K. Marksa, 64, Kryvyi Rig, Dnipropetrovska obl., 50000, Ukraine)

E-mail: ilayya@mail.ru

E-mail: ilayya@mail.ru

В останні десятиліття у світі відбувається зростання інтересу до математичного моделювання, побудови багатофакторних моделей та прогнозування на їх основі. На сьогодні це перспективний напрямок, оскільки використання таких моделей дозволяє ефективно досліджувати поведінку більшості складних динамічних систем, а прогноз, побудований на їх основі, є найбільш точним. Використання декількох різних моделей дозволяє зробити достовірну експертну оцінку. Проте можлива ситуація, коли зв'язок факторів, що впливають та результуючий показник, не очевидний, але є необхідність його дослідити, побудувати модель чи зробити прогноз.

Питанням дослідження даних засобами нейронних мереж займалися вітчизняні та іноземні вчені: Хайкін С., Руденко О. Г., Бодяньський Є. В., Осовський С. та інші. Зокрема сферою використання нейронних мереж в середовищі Matlab Neural Networks займалися Медведєв В. С., Потьомкін В. Г., Д'яконов В. та Круглов В. Питаннями аналізу часових рядів та їх прогнозуванням займалися Афанас'єв В. Н., Юзбашев М. М., Лукашин Ю. П. та інші. Вивченням смертності немовлят як демографічного показника, що характеризує стан соціального, економічного та культурного розвитку країни чи регіону, займалися Салліван А., Шефферін С. М. Проте в цих роботах не досліджено аспекти впливу неявно взаємопов'язаних факторів і можливість їх прогнозування.

Таким чином поставлена мета статті – дослідити залежність між часовими рядами, зв'язок між якими не є очевидним, побудувати адекватні моделі та спрогнозувати поведінку показників на кілька наступних періодів, використовуючи пакет Matlab Neural Networks.

В економічній літературі немає однозначного визначення поняття рівня життя населення [3]. У дослідженнях цього питання зазвичай розраховується велика кількість показників для забезпечення адекватної оцінки, наприклад: показники доходів населення; показники витрат і споживання населення матеріальних благ і послуг; заощадження; показники накопиченого майна та забезпеченості населення житлом; показники диференціації доходів населення, рівня та меж бідності; соціально-демографічні характеристики; узагальнюючі оцінки рівня життя населення. Але деякі дослідники виділяють показник смертності немовлят, що тісно корелює з вищевказаною сукупністю показників [5]. Тому показник смертності немовлят (смертності дітей віком до 1 року, включаючи перинатальну смертність) використано як результуючий при побудові моделі.

Для дослідження вищезазначеної проблеми та побудови моделі були обрані показники рівня життя населення та смертності немовлят. Результуючим показником в моделі буде використаний показник смертності немовлят [5].

Факторами, що впливають на рівень життя, обрано показники реального ВВП на душу населення, індекс споживчих цін та індекс Доу-Джонса. Зв'язок результуючого показника та цих факторів є не просто не очевидним, а навіть суперечливим.

Для забезпечення високої адекватності моделі та точності прогнозу необхідно мати достовірну та повну

інформацію. Використання інструментарію нейронних мереж також накладає свої умови: для побудови якісної мережі необхідна велика вибірка даних. Тому за базу дослідження використані офіційні статистичні дані з сайтів <http://finance.yahoo.com/>, <http://www.measuringworth.com/> (промисловий індекс Доу-Джонса, реальний ВВП на душу населення, індекс споживчих цін), <http://www.cdc.gov/>, <http://www.census.gov/>, <http://www.sidscenter.org/>, <https://www.cia.gov/> (статистичні дані по смертності немовлят).

Підготуємо отриманні дані для побудови моделі:

- ✦ для кращої апроксимації згладжуємо ряди *GDP* та *CPI* за допомогою експоненційного згладжування з коефіцієнтом 0,7;
- ✦ для прогнозування часового ряду за допомогою нейронних мереж створюємо плаваюче вікно (табл. 1), в якому перші чотири стовпця є входами нейронної мережі, останній – вихід. Таким чином, на основі чотирьох попередніх значень *D-n* прогнозується таке значення *D-(n+1)*.

Таблиця 1

Впорядкування часового ряду по типу плаваючого вікна

Hist1	Hist2	Hist3	Hist4	Result
D-1	D-2	D-3	D-4	D-5
D-2	D-3	D-4	D-5	D-6
D-3	D-4	D-5	D-6	D-7
...

Аналогічним чином впорядковуємо досліджувані дані, але враховуємо присутність результуючого показника, що має бути спрогнозований на *N*-періодів у відповідності до трьох факторів.

На сьогоднішній момент існує кілька десятків структур нейронних мереж. Оскільки всі штучні нейронні мережі базуються на концепції нейронів, з'єднань та передатних функцій, існує подібність між різними структурами нейронних мереж. Більшість змін походять з різних правил навчання.

Для процесу навчання необхідно мати модель зовнішнього середовища, у якій функціонує нейронна мережа – потрібну для вирішення задачі інформацію. Також, необхідно визначити, як модифікувати вагові параметри мережі. Алгоритм навчання означає процедуру, у якій використовуються правила навчання для налаштування ваг.

Існують три загальні парадигми навчання: «з учителем», «без учителя» (самонавчання) і змішана [4]. У першому випадку нейромережа має у своєму розпорядженні правильні відповіді (виходи мережі) на кожен вхідний приклад, що відповідає умовам поставленої задачі. Тому саме ця концепція обрана для реалізації моделі.

Загальних правил вибору архітектури та умов навчання мережі для конкретного завдання не існує, тому було побудовано декілька нейронних мереж різної архітектури, кількості нейронів і правил зміни ваг.

Перші адекватні результати отримані за допомогою мережі архітектури Feed-forward Backpropagation, що має три прошарки (на 75, 15 та 1 нейрон відповідно) з функцією активації *tansig* на кожному прошарку, трьома вхідними та одним вихідним нейроном (рис. 1).

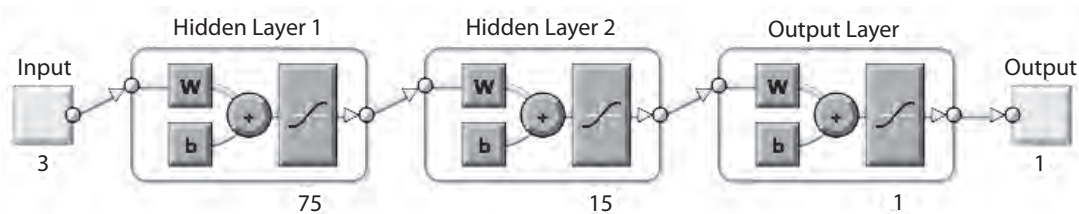


Рис. 1. Нейронна мережа архітектури Feed-forward Backpropagation

При побудові цієї мережі виявилось, що існує велика дисперсія прогнозованих та реальних значень смертності немовлят у період 1915 – 1937 рр. Тому було вирішено скоректувати вхідний масив інформації та надалі використовувати дані з 1937 по 2011 рр. – 75 значень, 70 з яких беруться для навчання і 5 – для прогнозування.

Результат роботи мережі зображено на рис. 2.

Результат роботи цієї мережі наведено на рис. 4.

У математичній моделі помітні аномальні значення: 1978 – 1990 рр. У 2004 – 2005 рр. модель має значне аномальне відхилення, що ставить під сумнів достовірність наступного прогнозу. Прогнозні значення смертності 6,9442; 6,0684; 6,0763; 6,1038; 6,117. Отримане значення моделі на 2011 р. і перше прогнозоване значення

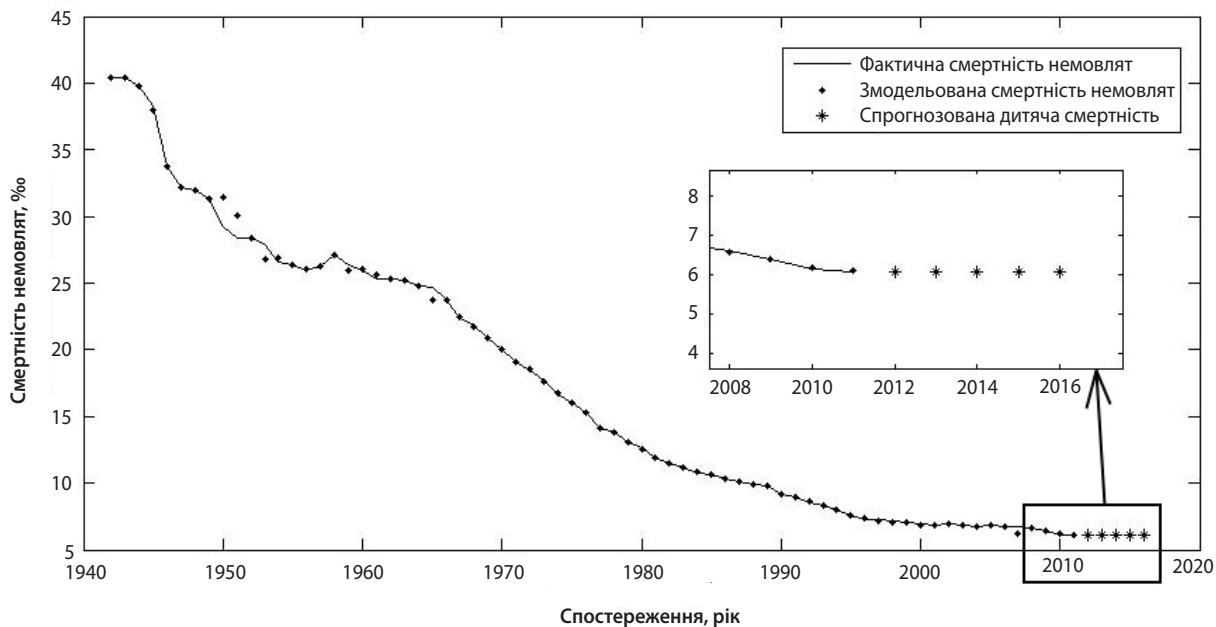


Рис. 2. Прогноз з використанням нейронної мережі архітектури Feed-forward Backpropagation

На графіку видно, що побудована модель високого рівня достовірності. Помітні кілька незначних аномальних значень на проміжку 1950 – 1955 та 2007 рр. Прогнозні значення смертності 6,0748; 6,0627; 6,0645; 6,067; 6,0657. Для оцінки якості побудованої моделі візьмемо значення скоректованого коефіцієнту детермінації [2], що дорівнює 0,9985, тобто достовірність моделі висока.

Друга нейронна мережа має архітектуру Cascade-Forward-backpropagation та має 2 прошарки з 14 нейронами на 1 прошарку і 1 на другому. Функція активації – tansig на кожному прошарку. Структуру наведено на рис. 3.

(2012 рік) також виявили аномальне відхилення. Для даної моделі скоректований коефіцієнт детермінації дорівнює 0,9794. Його зниження пояснюється наявністю аномальних значень, але все ж достовірність моделі достатньо висока.

Третя нейронна мережа має архітектуру NARX з 2-ма прошарками (10 та 1 нейронів відповідно) та функцією активації tansig (рис. 5).

Результат роботи мережі наведено на рис. 6.

Прогнозні значення смертності 6,0617; 6,0622; 6,064; 6,0617; 6,0608. Скоректований коефіцієнт детермінації дорівнює 0,9966, що дещо нижче, ніж у першій моделі, проте вище, ніж у другій. Достовірність даної моделі можна також визнати дуже високою.

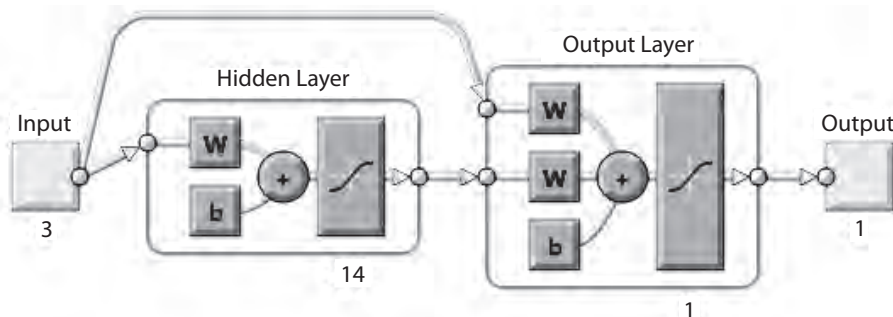


Рис. 3. Нейронна мережа архітектури Cascade-Forward-backpropagation

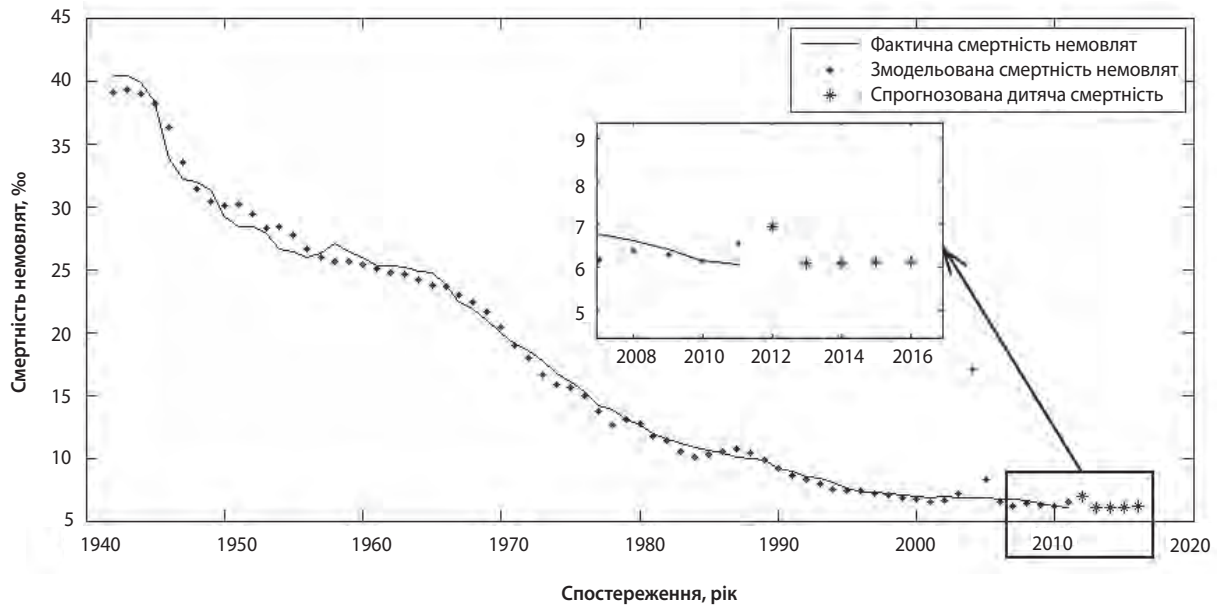


Рис. 4. Прогноз з використанням нейронної мережі архітектури Cascade-Forward-backpropagation

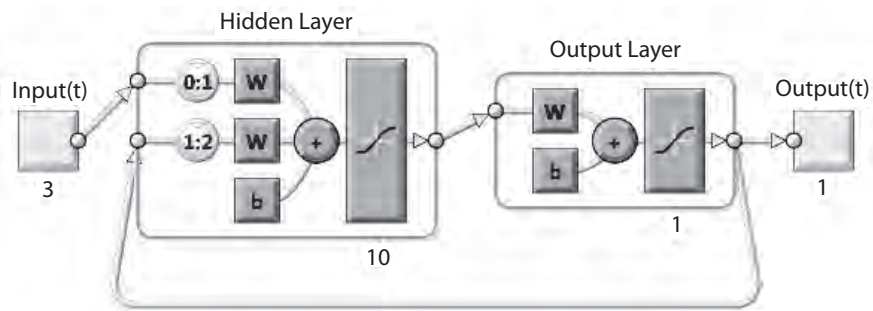


Рис. 5. Нейронна мережа архітектури NARX

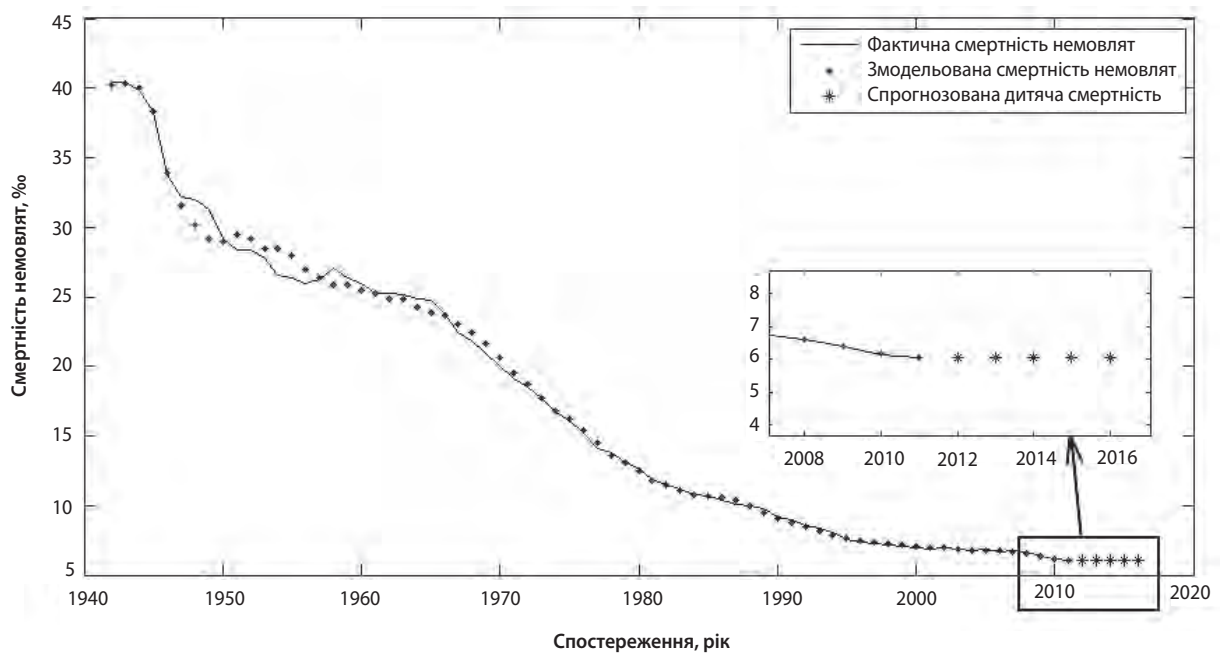


Рис. 6. Прогноз з використанням нейронної мережі архітектури NARX

Для порівняння була побудована багатофакторна модель [2] з експоненційним розподілом у середовищі MS Excel (рис. 7).

Прогнозні значення смертності 5,626; 5,524; 5,844; 5,723; 5,57. Скоректований коефіцієнт детермінації дорівнює 0,97714, що нижче, ніж у всіх попередніх моделей. З 1950 по 2011 рр. модель показує значення, близькі до еталонних. Але при збільшенні масштабу стає помітно, що побудована модель після 2004 р. показує значення смертності нижчі за реальні, тому прогнозовані значення після 2011 р. вважати адекватними неможливо.

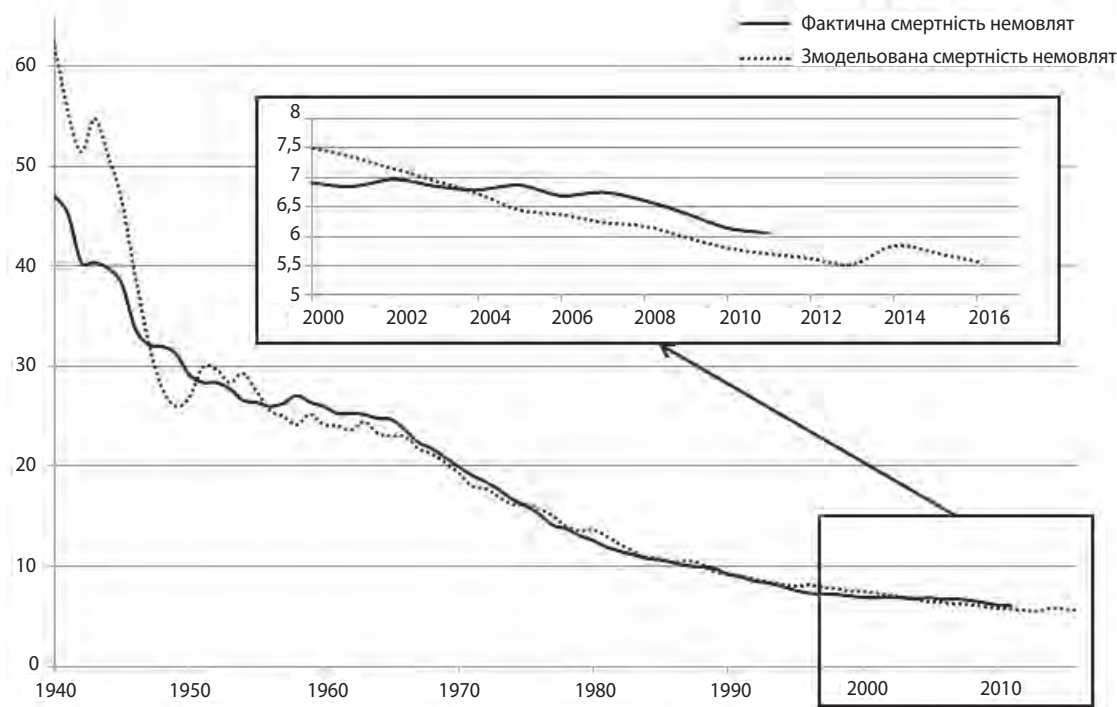


Рис. 7. Прогноз з використанням багатофакторної математичної моделі в середовищі MS Excel

ВИСНОВКИ

Порівнюючи моделі, можна зазначити, що реалізація багатофакторної математичної моделі значно простіша та більш передбачувана і займає менший об'єм часу. Але в умовах, коли зв'язок впливаючих і результуючих показників не очевидний, не дає достатньої точності прогнозу, тим більше, якщо необхідно зробити прогноз більше, ніж на 1 крок. За таких умов доцільніше використовувати більш досконалі інструментарій, такий як штучні нейронні мережі. У результаті вдалося отримати адекватні моделі, що з високим рівнем достовірності вирішують дане питання. Подібний досвід надалі можливо використати в будь-яких соціально-економічних, статистичних та інших дослідженнях на всіх рівнях. Таким чином, штучні нейронні мережі надають великі можливості для побудови багатофакторних моделей при дослідженні економічних і соціальних показників. ■

ЛІТЕРАТУРА

1. **Борисов В. В.** Искусственные нейронные сети. Теория и практика / В.В. Круглов, В. В. Борисов. – М.: Горячая линия – Телеком, 2001. – 382 с.

2. **Лещинський О. Л.** Економетрія : Навч. посіб. для студ. вищ. навч. закл. / О. Л. Лещинський, В. В. Рязанцева, О. О. Юнькова. – К.: МАУП, 2003.

3. **Романов О. М.** Рівень життя населення / О. М. Романов, В. М. Жеребін. – М.: Юнити-Дана, 2002.

4. **Хайкин С.** Нейронные сети / С. Хайкин. – М.: Вильямс, 2006. – 1103 с.

5. **Sullivan, Arthur; Steven M. Sheffrin** (2003). Economics: Principles in action. Upper Saddle River, New Jersey 07458: Pearson Prentice Hall. pp. 474. ISBN 0-13-063085-3.

Рецензент: зав. кафедрою інформатики та інформаційних технологій Криворізького економічного інституту ДВНЗ КНУ
Зеленський О. С.