

РОЗДІЛ 9. МАТЕМАТИЧНІ МЕТОДИ, МОДЕЛІ
ТА ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ В ЕКОНОМІЦІПРОГНОЗУВАННЯ ЧАСОВИХ РЯДІВ СОЦІАЛЬНИХ МЕРЕЖ
НА ОСНОВІ НЕЙРОМЕРЕЖЕВОГО ПІДХОДУFORECASTING TIME SERIES OF SOCIAL NETWORKS
BASED ON THE NEURAL NETWORK APPROACH

У статті представлено прогнозну модель для аналізу стану та поведінки об'єктів соціальних мереж в динаміці та для побудови передбачень поведінки користувачів. Це відкриває можливості для вибору найбільш перспективного просування об'єктів в соціальних мережах, найкращого розміщення реклами, можливості створення самостійних проектів. У статті запропоновано підхід до формалізації процесів зросту/падіння популярності контенту під впливом зовнішніх факторів в рамках математичної прогнозної моделі на рекурентній нейронній мережі. Розроблена модель може використовуватися для прогнозування зросту репостів в соціальних мережах. Описано перелік зовнішніх і внутрішніх змінних, які впливають на поведінку вихідних даних в кожен момент часу. На основі отриманих з соціальних мереж вхідних даних вивчається поведінка часових рядів. Особливістю моделі є можливість працювати з масивами типу Big Data, які мають в більшості випадків мають не повну, або «зашумлену» інформацію. Результатом реалізації поставлених в дослідженні завдань є динамічна прогнозна модель, яка надає можливість створювати передбачення по таким показникам, як репости у соцмережі, тренди музики, популярність блогерів. Для прогнозування репостів по вибраним об'єктам дослідження запропоновано використати рекурентну штучну мережу з динамічним нейроном. Нейрон обробляє вхідні сигнали і розраховує прогнозний показник кількості репостів. Продемонстровано процес навчання нейронної мережі за період моделювання один місяць. Розкрити перспективи подальшої розробки моделі з метою покращення точності прогнозів соціально-економічних процесів.

Ключові слова: економіко-математична модель, прогнозування, соціальна мережа, нейронна мережа, часовий ряд.

The article presents a predictive model for analyzing the state and behavior of social network objects in dynamics and for building predictions of user behavior. This opens up opportunities for choosing the most promising promotion of objects in social networks, the best placement of advertising, the possibility of creating independent projects. The article proposes an approach to the formalization of the processes of growth/decrease in the popularity of content under the influence of external factors within the framework of a mathematical predictive model based on a recurrent neural network. The purpose of the model is mathematical processing and analysis of stochastic information flows, graphical presentation of forecasts of music trends, changes in the popularity of bloggers, video content, and more. The developed model can be used to predict the growth of reposts in social networks. The list of external and internal variables that affect the behavior of the output data at each moment of time is described. The behavior of time series is studied on the basis of input data obtained from social networks (nomenclature of indicators, number of reposts of researched objects, simulation period). A feature of the model is the ability to work with Big Data arrays, which in most cases have incomplete or "noisy" information. The dependence between input and output is proposed to be specified using a system of constraints and external coefficients that change over time. Control parameters are predictors of the model. Features of building a predictive model are shown, a list of the main limitations of the mathematical model is described. The result of the implementation of the tasks set in the research is a dynamic predictive model that provides an opportunity to create predictions based on such indicators as reposts on social networks, music trends, and the popularity of bloggers. It is proposed to use a recurrent artificial network with a dynamic neuron to predict reposts on selected research objects. The neuron processes input signals and calculates a predictive indicator of the number of reposts. The learning process of the neural network during the simulation period of one month is demonstrated. To reveal the prospects of further development of the model in order to improve the accuracy of forecasts of socio-economic processes.

Key words: economic-mathematical model, forecasting, social network, neural network, time series.

УДК 004.94

DOI: <https://doi.org/10.32782/infrastruct70-34>

Івченко І.Ю.

к.е.н., доцент,
доцент кафедри економічної
кібернетики та інформаційних технологій,
Національний університет
«Одеська політехніка»

Лінгур Л.М.

к.е.н., доцент,
доцент кафедри економічної
кібернетики та інформаційних технологій,
Національний університет
«Одеська політехніка»

Івченко О.І.

аспірант,
Національний університет
«Одеська політехніка»

Ivchenko Iryna

Odessa Polytechnic National University

Linhur Liubov

Odessa Polytechnic National University

Ivchenko O.I.

Odessa Polytechnic National University

Постановка проблеми. В сучасних умовах стрімкого розвитку інформаційних технологій все частіше для досліджень та отримання непрямої інформації використовуються соціальні мережі. Вони виступають джерелами великих масивів даних, формують політичну, економічну, культурну компоненту суспільства. Аналітика інформаційних масивів даних соціальних мереж надає можливість підприємцям розробляти оптимальну стратегію

розвитку бізнесу, підвищувати конкурентоздатність на сучасному ринку товарів та послуг [1].

Вивчення соціальних мереж виконується за допомогою збору первинних даних, аналізу отриманих вибірок та з подальшою побудовою різноманітних моделей: оптимізаційних, імітаційних, прогнозних, і т. ін. На основі отриманих вихідних даних будується аналітика стану та поведінки користувачів соціальних мереж (наприклад, у

мережі Facebook, Twitter, Instagramm, TikTok) та розробляється прогноз на подальшу поведінку досліджуваного процесу.

Для ефективного управління соціальними процесами та економічними явищами в дослідженні розробляється прогнозна нейромережева модель поведінки реальних об'єктів на базі часових рядів соціальних мереж. Особливість моделі в тому, що вона працює з інформаційними масивами типу BigData, які в більшості випадків мають не повну, або «зашумлену» інформацію.

Прогнозування часових рядів в умовах впливу зовнішніх факторів спрямоване на вирішення задачі передбачення поведінки користувачів з метою відбору найперспективнішого контенту (блогерів, музики, відео тощо) та подальшого інвестування їх діяльності, просування в соціальних мережах, розміщення реклами, можливості створення самостійних проектів за межами онлайн-продуктів.

Аналіз останніх досліджень і публікацій.

Класичні принципи, методи аналізу та прогнозування часових рядів (ЧР) розглядають різні світові науковці: Балабанов О. С., Калініна І. О., Снитюк В. Є., Cai Q., Zhang D., Zheng, W., Serikov T., Wang, L. та ін. [1–7]. Найбільш розповсюдженими є статистичні методи кластерного та регресійного аналізу, але всі вони мають деякі недоліки. Так за допомогою цих методів будують переважно лінійні математичні моделі, які недостатньо докладно можуть описувати досліджуваний процес або систему; моделі парного регресійного аналізу практично не дають можливості аналізувати великі масиви даних; моделі множинного аналізу враховують більшу кількість чинників впливу, але часто мають неякісні оцінки параметрів [2]. Все це значно звужує можливості їх використання для прогнозування реальних соціально-економічних процесів, які мають в основі інформаційні масиви типу BigData.

Прогнозування є невід'ємною частиною задач прийняття управлінських рішень. Прогнозування соціально-економічних процесів повинно опиратися на математичне моделювання та методи суміжних галузей наукових досліджень. В [3, с. 61–68] автори Cai Q., Zhang D., Zheng W., Leung S. запропонували для прогнозування нечітких часових рядів оптимізаційну математичну модель, яка базується на алгоритмі поведінки мурах, відома як «мурашиний алгоритм». Автори відмічають підвищення ефективності прогнозування нечітких часових рядів при оптимізації мурашиної колонії в поєднанні з авторегресією. Задачі прогнозування соціальних явищ потребують обробки великих масивів даних в умовах неповноти, невизначеності, неточності інформації. Найкращим серед існуючих на цей час рішенням є використання для прогнозування нейронних

мереж. Проблеми прогнозу даних з використанням нейронних мереж розглядаються багатьма вченими. Так в [4, с. 12–19] Serikov, T. та його соавтори обґрунтували застосування нейронної мережі NARX для прогнозування поведінки мережевого трафіку та надання якісних мережевих послуг. Перевагами цієї моделі є можливість роботи з нелінійними динамічними системами, проте модель вузько спрямована та актуальна саме для довгострокового мережевого прогнозування. Ротштейн А. П. в [5] запропонував використовувати нейромережеві та генетичні алгоритми для прогнозування нелінійних часових рядів. Автори Wang L., Wang Z. в [6, с. 3–15] застосовують комбінований метод прогнозування, поєднуючи одразу декілька нейронних мереж, що приводить к покращенню ступені прогнозу. Перевагою моделей на нейромережах є висока точність прогнозу, проте, недоліком комбінованих методів є їхня мала гнучкість та адаптивність.

Аналіз наукових джерел дозволяє зробити наступні висновки. Порівняно з класичними методами прогнозування певні переваги мають прогнозні нейронні мережі. З їхньою допомогою можна моделювати нелінійні відносини у соціально-економічних системах за умов невизначеності та неповноти вхідних даних. До того ж нейронні мережі дозволяють працювати з масивами типу BigData, що мають довільний характер; здатність нейронних мереж до навчання ще одна з переваг перед традиційними моделями. Все це доводить, що застосування нейромережевих технологій для математичної обробки та прогнозування даних різних соціальних мереж є актуальним завданням. Використання нейромережевих технологій для побудови прогнозних моделей дозволить більш ефективно вирішувати актуальні соціальні та економічні проблеми розвитку сучасного суспільства.

Постановка завдання. Метою дослідження є обґрунтування використання математичних моделей для побудови прогнозів поведінки об'єктів соціальних мереж в умовах невизначеності та неповноти інформації.

Об'єктом дослідження виступають часові ряди по репостам контенту в соціальних мережах. Візьмемо для розрахунків інформацію про популярність музикальних трендів в соціальній мережі TikTok. Ці дані по характеристиках відносяться до стохастичних, мають тип BigData, формуються по невизначених правилах. Для аналізу часових рядів, побудови графічного уявлення прогнозування змін популярності контенту пропонується розробка математичної моделі на основі нейронної мережі. Це дозволить передбачати поведінку користувачів соціальної мережі та створювати прогнози просування контенту з високою точністю.

Виклад основного матеріалу дослідження. Багатим джерелом вхідних даних для

моделювання є соціальні мережі. З одного боку вони виступають джерелом великої кількості вхідних даних, з іншого боку, соціальні мережі можуть бути і споживачами даних. Використання даних із соціальних мереж дозволяє вирішувати широке коло бізнес-завдань: реклама, популяризація брендів, пошук збуту товарів та послуг та ін.

Розглянемо соціальну мережу для створення та перегляду коротких відео TikTok. Важливим чинником популяризації музичних трендів, відео, блогерів, тощо є процеси зміни кількості репостів. Поведінка цієї мережі характеризується цілеспрямованою зміною стану її показників у часі. Стан системи на виході залежить як від значень вхідних показників у визначений момент часу, так і від значень показників у попередні моменти часу. При аналізі обраної системи (соціальної мережі TikTok) було доведено, що вхідні дані у кожен час мають випадковий характер, що впливає на точність результатів роботи моделі. Тому обов'язковою умовою достовірності результатів роботи моделі, що розробляється, є використання великих вибірок даних, типу BigData.

В моделі для досягнення поставленої мети дослідник повинен мати можливість впливати на систему – керувати нею. З цією метою в модель вводяться предиктори моделі, тобто параметри керування. Змінюючи значення предикторів можна керувати процесом досягнення кінцевого стану системи. У моделі, що розробляється, як такі керуючі впливи на систему, запропоновано використовувати значення основних (вхідних) показників у нульовий момент часу. На системі TikTok ми можемо не тільки застосувати, але й отримати зворотний зв'язок, перевіривши взаємодію між об'єктами, що вивчаються, їх кореляцію. Будемо використовувати як вхідні дані індивідуальні часові ряди з кількістю репостів на конкретну дату для конкретного випадку. Це, наприклад, може бути музичний тренд, окремий користувач (блогер), відеоконтент, і т. ін. Розглянуті величини (кількість репостів) мають випадковий характер, що і є причиною виникнення невизначеності. Період моделювання не обмежений (довільний), як крок моделювання (ітерації) обрано добу. Дані про процес можуть збиратися не щодня, але у разі відсутності достовірних даних на деякий проміжок часу в моделі необхідно передбачити «відображення» аналізованих процесів, тобто зробити поправку на кількість пропущених в обстеженні днів.

Інструментом керування, як було сказано раніше, виступають предиктори. Для створення керуючого предиктора на модель на нульовому етапі обчислень задаємо наступну множину керуючих змінних: рівень успішності виконавця (залежить від рівня популярності музичного тренду на початок моделювання); можлива кількість репостів на момент початку моделювання (передбачення,

стартова величина); виправлення на помилку передбачення. Ставиться завдання – передбачити майбутнє, тобто спрогнозувати кількість репостів на наступному етапі ітерації.

Сформулюємо обмеження математичної моделі:

$$\begin{cases} pV = 0, \text{ якщо первинних даних про об'єкт не має;} \\ pV = 1, \text{ якщо первинні данні про об'єкт є.} \end{cases} \quad (1)$$

$$0 \leq dS_i \leq n_i/2, \quad (2)$$

$$dS_i = n_i - (n_i - 1), \quad (3)$$

$$DRFE_{i(n-1)} > 0. \quad (4)$$

де i – номенклатура об'єктів (окремо для кожного музичного тренду); pV – булева змінна, яка характеризує наявність початкових даних про об'єкт; n_i – дата моделювання; dS_i – кількість пропущених під час аналізу днів (для інтерполяції даних); $DRFE_{i(n-1)}$ – помилка передбачення i -го об'єкту в момент $(n-1)$.

Для вирішення завдань прогнозування кількості репостів запропоновано побудувати рекурентну нейронну мережу. У даному дослідженні базовим елементом мережі є динамічний нейрон для обчислення прогнозних показників репостів. У побудованій штучної нейронної мережі вхідні сигнали про кількість репостів відразу подаються на вихід, нейрон перетворює сигнал і видає прогноз про кількість репостів на один крок вперед. Далі нейрон обробляє ці данні з врахуванням ваги відповідного зв'язку.

Для навчання нейрона (наприклад, для передбачення репостів музичного тренду) запропоновано скористатися алгоритмом «зворотного розповсюдження помилки» [7], якій налаштовує вагові коефіцієнти для отримання більш точного прогнозу репостів. Результат навчання нейрона, який проводився в реальному часі по принципу «за кроками», наведений на рис. 1.

На рис. 1 лінія «Репости» ілюструє фактичні значення щоденної зміни репостів. Лінія «Прогноз» відтворює траєкторію передбачуваного майбутнього репостів, зважаючи на те, що реальне майбутнє обчислене з урахуванням минулих знань та помилок передбачень. Кожен наступний крок, після розрахування значення вихідної змінної, враховує поправки на вагові коефіцієнти. Тому кожна нова ітерація буде мати інші вхідні параметри (прогноз репостів), які змінюються у відповідності з поточними параметрами входів (кількість репостів) та передбаченими виходами. Якість прогнозу швидко покращується за рахунок навчання нейрона. Ціллю навчання нейрона є мінімізація помилок між розрахунковими та вхідними даними.

Нейронна мережа здатна постійно вдосконалюватися, тому в майбутніх дослідженнях можливо навчити нейрон робити максимально точні прогнози, поєднуючи декілька різних контентів. Це може бути поєднання як декількох нейронів

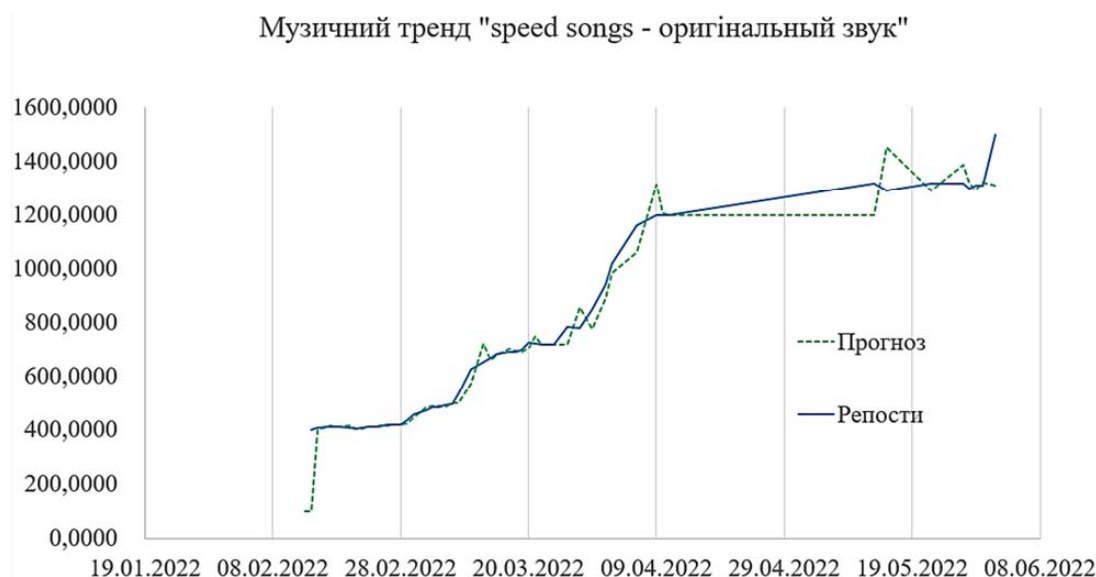


Рис. 1. Процес навчання нейромережі (лютий – березень)

Джерело: авторська розробка

у мережу (наприклад, для передбачення змін по зв'язаних контентах або для прогнозування по представникам одного бренду). Такий підхід зводиться до задачі оптимізації для прогновної моделі з побудовою цільової функції. Послідовність формування та вибір оптимальних траєкторій управління є предметом подальших розробок авторів в задачах оптимізації на основі розробленої динамічної прогновної нейронної моделі, ідеології статистичних випробувань та прийняття рішень в умовах невизначеності впливу зовнішнього середовища.

Висновки. У дослідженні показано особливості побудови математичної моделі, що призначена для прогнозування часових рядів показників, які характеризують поведінку складних систем на прикладі соціальної мережі TikTok. Базовим елементом нейронної мережі, що розробляється в даному дослідженні, є динамічний нейрон для обчислення прогнозних показників репостів музичних трендів у соціальній мережі TikTok. Процес навчання нейрона відбувається з урахуванням помилки передбачення за попередній період. Завдяки такому механізму нейрон в режимі реального часу вчиться передбачати яка музика буде в тренді завтра, і навпаки, які тік-ток тренди не зростатимуть і т.д. Описаний алгоритм застосовується до побудови індивідуальних прогнозів поведінки досліджуваного елемента (наприклад, музичного тренду) за умов, коли вихідні індивідуальні дані носять стохастичний характер, а прогноз будується на крок уперед. Загальна працездатність моделі перевірена на великих вибірках даних музичних трендів мережі Тік-Ток.

Подальша розробка моделі передбачає додавання нових нейронів (наприклад, на відео,

на блогерів тощо) на різних шарах мережі, їх взаємопов'язане та спільне навчання, а також вирішення завдання довгострокового прогнозування. Практична цінність моделі полягає в можливості її використання в будь-якому середовищі, що вимагає отримання знань на перспективу, передбачень даних. Розробка математичних моделей з використанням нейромереж є новим інструментом для вивчення таких соціально-економічних явищ як соціальні мережі, телеграм-канали, інтернет-спільноти, спільні чати, форуми і т. ін.

БІБЛІОГРАФІЧНИЙ СПИСОК:

1. Балабанов О. С. Аналітика великих даних: принципи, напрямки і задачі (огляд). *Проблеми програмування*. 2019. № 2. С. 47–68. DOI: <https://doi.org/10.15407/pp2019.02.047>.
2. Снитюк В. Є. Прогнозування. Моделі. Методи. Алгоритми : навчальний посібник. Київ : «Маклаут», 2008. 364 с.
3. Cai Q., Zhang D., Zheng W., Leung S. C.H. A new fuzzy time series forecasting model combined with ant colony optimization and auto-regression. *Knowledge-Based Systems*, V. 74, (2015) pp. 61–68. DOI: <http://doi.org/10.1016/j.knosys.2014.11.003>.
4. Serikov, T., Zhetpisbayeva, A., Mirzakulova, S., Zhetpisbayev, K., Ibrayeva, Z., Soboleva, L., Tolegenova, A., & Zhumazhanov, B. Application of the NARX neural network for predicting a one-dimensional time series. *Eastern-European Journal of Enterprise Technologies*, 2021. Vol. 5(4 (113)), pp. 12–19. DOI: <https://doi.org/10.15587/1729-4061.2021.242442>.
5. Ротштейн А. П. Интеллектуальные технологии идентификации: нечеткие множества, генетические алгоритмы, нейронные сети. Винница: «УНІВЕРСУМ-Вінниця», 1999. 320 с. URL: http://pdf.lib.vntu.edu.ua/books/2019/Rotshtejn_1999_320.pdf.

6. Wang, L., Wang, Z., Qu, H., Liu, S. (2018). Optimal Forecast Combination Based on Neural Networks for Time Series Forecasting. *Applied Soft Computing*, 2018. Vol. 66. P. 1–17. DOI: <http://doi.org/10.1016/j.asoc.2018.02.004>.

7. Калініна І. О. Дослідження нейромережових методів у задачах прогнозування. *Наукові праці. Випуск 93. Том 106. С. 132–138*. URL: <https://lib.chmnu.edu.ua/pdf/naukpraci/computer/2009/106-93-17.pdf>.

REFERENCES:

1. Balabanov O. S. (2019) Analitika velykykh danykh: pryntsyry, napriamky i zadachi. *Problemy prohramuvannia*, no. 2, pp. 47–68. DOI: <https://doi.org/10.15407/pp2019.02.047>.

2. Snytiuk V. Ye. (2008) Prohnozuvannia. Modeli. Metody. Alhorytmy [Forecasting. Models. Methods. Algorithms]. Kyiv: «Maklout», 364 p.

3. Cai Q., Zhang D., Zheng W., Leung S. C. H. (2015). A new fuzzy time series forecasting model combined with ant colony optimization and auto-regression. *Knowledge-Based Systems*, vol. 74, pp. 61–68. DOI: <http://doi.org/10.1016/j.knosys.2014.11.003>.

4. Serikov, T., Zhetpisbayeva, A., Mirzakulova, S., Zhetpisbayev, K., Ibrayeva, Z., Soboleva L., Tolegenova, A., & Zhumazhanov, B. (2021). Application of the NARX neural network for predicting a one-dimensional time series. *Eastern-European Journal of Enterprise Technologies*, vol. 5(4(113)), pp. 12–19. DOI: <https://doi.org/10.15587/1729-4061.2021.242442>.

5. Rothstein A. P. (1999) Yntellektualnie tekhnolohy ydentyfikatsyy: nechetye mnozhestva, henycheskye alhorytmi, neironnie sety [Intelligent identification technologies: fuzzy sets, genetic algorithms, neural networks]. Vinnytsia: "UNIVERSUM-Vinnytsia", 320 p. Available at: http://pdf.lib.vntu.edu.ua/books/2019/Rotshtejn_1999_320.pdf.

6. Wang, L., Wang, Z., Qu, H., Liu, S. (2018). Optimal Forecast Combination Based on Neural Networks for Time Series Forecasting. *Applied Soft Computing*, vol. 66, pp. 1–17. DOI: <http://doi.org/10.1016/j.asoc.2018.02.004>.

7. Kalinina I.O. Doslidzhennia neiromerezhevykh metodiv u zadachakh prohnozuvannia [Research of neural network methods in forecasting tasks]. *Naukovi pratsi*, vol. 93, tom 106, pp. 132–138. Available at: <https://lib.chmnu.edu.ua/pdf/naukpraci/computer/2009/106-93-17.pdf>.