

<https://doi.org/10.31891/2307-5732-2026-365-80>

УДК 004.89:004

МОНЬКО ОЛЕКСАНДР

Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»

<https://orcid.org/0009-0005-7110-1540>

e-mail: monko.oleksandr@lpi.kpi.ua

ЛЮШЕНКО ЛЕСЯ

Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»

<https://orcid.org/0000-0003-4319-5955>

e-mail: lyushenkol@gmail.com

ГІБРИДНИЙ НЕЙРО-СИМВОЛЬНИЙ МЕТОД ПРОГРАМНОГО ПРОГНОЗУВАННЯ З ВИКОРИСТАННЯМ ЧИСЛОВИХ РЯДІВ ТА НЕСТРУКТУРОВАНИХ ТЕКСТОВИХ ДАНИХ

Програмне прогнозування за допомогою сучасних методів, таких як машинне навчання, глибокі нейронні мережі тощо, критично залежить від якості вхідних даних, їх структурованості та динаміки змін. Тому особливої уваги заслуговують методи аналізу, обробки та побудови моделі даних бази знань. Від зазначених процесів залежить вибір методу прогнозування та подальші дії щодо уточнення прогнозу та інтерпретації отриманих прогнозних результатів. Такий підхід має певні труднощі в програмній реалізації, які пов'язані з тим, що реальні дані мають пропуски, інформаційний шум, зайві дані тощо. Аналіз, обробка та побудова моделі даних потребують додаткової експертної оцінки, особливо за наявності прихованих факторів впливу, які не фіксуються в явному вигляді. У статті представлено програмний гібридний метод, який реалізує нейро-символьну інтеграцію числових даних з неструктурованою текстовою інформацією, яка має певні експертні дані, для уточнення прогнозу. Запропонований підхід розглядає неструктуровану текстову інформацію, яка містить експертні знання, як засіб компенсації помилки базової моделі прогнозування. Метод використовує цифровий адаптивний механізм шлюзування, який ініціює пошук експертної інформації виключно тоді, коли оцінка прогностичної невизначеності вказує на критичний брак релевантних числових патернів, що дозволяє системі автономно визначати необхідність звернення до бази знань. Окрім того, програмний гібридний метод дозволяє перетворити знайдену текстову неструктуровану інформацію на формалізовані експертні дані, які доповнюють модель вхідних даних, що дозволяє уточнювати прогноз у випадку нерелевантних або неповних вхідних даних, що спрямовано на підвищення точності прогнозування та забезпечення інтерпретованості результатів.

Ключові слова: прогнозування часових рядів, гібридні нейро-символьні системи, оцінка невизначеності, семантичний пошук.

MONKO OLEKSANDR, LIUSHENKO LESIA

National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine

HYBRID NEURO-SYMBOLIC METHOD OF PROGRAMMATIC FORECASTING USING NUMERICAL SERIES AND UNSTRUCTURED TEXTUAL DATA

Programmatic forecasting using modern methods, such as machine learning and deep neural networks, critically depends on the quality of the input data, their structural integrity, and the dynamics of their changes. Therefore, methods of data analysis, processing, and modeling deserve special attention. The choice of the forecasting method, as well as subsequent actions for refining the forecast and interpreting the obtained predictive results, depend on these processes. Such an approach encounters certain difficulties in software implementation, which are associated with the fact that real-world data contain missing values, informational noise, redundant data, etc. Data analysis, processing, and modeling require additional expert evaluation, especially in the presence of hidden influencing factors that are not explicitly captured. This article presents a software-based hybrid method that implements the neuro-symbolic integration of numerical data with unstructured textual information containing specific expert data to refine the forecast. The proposed approach considers unstructured textual information, which contains expert knowledge, as a means of compensating for the errors of the base forecasting model. The method utilizes a digital adaptive gating mechanism that initiates the search for expert information exclusively when the estimation of predictive uncertainty indicates a critical lack of relevant numerical patterns, allowing the system to autonomously determine the necessity of accessing the knowledge base. Furthermore, the programmatic hybrid method allows for the transformation of the retrieved unstructured textual information into formalized expert features that supplement the input data model. This enables forecast refinement in cases of irrelevant or incomplete input data, which is aimed at improving forecasting accuracy and ensuring the interpretability of the results.

Keywords: time series forecasting, neuro-symbolic approach, predictive uncertainty estimation, semantic search.

Стаття надійшла до редакції / Received 24.02.2026

Прийнята до друку / Accepted 11.04.2026

Опубліковано / Published 28.05.2026



This is an Open Access article distributed under the terms of the [Creative Commons CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)

© Монько Олександр, Люшенко Леся

Постановка проблеми та її зв'язок із важливими науковими чи практичними завданнями

Сучасний світ характеризується високою динамічністю та невизначеністю, що спричиняє виникнення кризових явищ в економіці та критично важливих галузях. Сталий розвиток та дієвість управлінських рішень безпосередньо залежать від здатності своєчасно виявляти ризики, що, у свою чергу, вимагає надійного програмного прогнозування складних процесів. Проте фундаментальною проблемою при побудові таких прогностичних систем залишається якість та повнота вхідних даних. Тому задача передбачення майбутніх станів часто ускладнюється дефіцитом релевантної ретроспективної інформації: наявних числових часових рядів зазвичай недостатньо для моделювання складних нелінійних сценаріїв.

Ситуація загострюється впливом прихованих факторів – непередбачуваних зовнішніх подій та раптових структурних зрушень, які фізично не можуть бути зафіксовані у попередніх числових вибірках. У таких умовах застосування суто кількісних методів машинного навчання стає неефективним. Інформація, необхідна для розуміння подібних нештатних ситуацій, зазвичай існує у вигляді експертної текстової інформації, яка здебільшого зберігається в неструктурованому текстовому форматі, що створює об'єктивний бар'єр для її прямого використання у прогностичних моделях.

У результаті маємо, що формування прогнозів виключно на основі ретроспективних числових значень залишає поза увагою критично важливу експертну інформацію. Ігнорування цієї інформації породжує ефект «інформаційної сліпоти» моделей прогнозування [1], що призводить до різкого падіння точності саме в ті моменти, коли вона найбільш необхідна (у періоди криз, аномалій та перехідних процесів).

Отже, актуальним завданням є розробка гібридного методу прогнозування, здатного долати невідповідність числових даних шляхом залучення неструктурованої текстової інформації, яка трансформується у формалізований вигляд та динамічно доповнює базову прогностичну модель.

Аналіз досліджень та публікацій

У сучасній науковій літературі актуальною є проблематика підвищення над методів машинного навчання для задач прогнозування в умовах невизначеності. Значна частина робіт присвячена обмеженням методів глибокого навчання при роботі з малими вибірками даних. Зокрема, у порівняльному дослідженні Гевамалаге, Бергмайр та Бандара доведено, що складні архітектури, такі як рекурентні нейронні мережі, часто поступаються простішим статистичним методам при прогнозуванні коротких часових рядів через схильність до перенавчання та складність оптимізації гіперпараметрів [2]. Це підкреслює необхідність пошуку гібридних підходів, які могли б компенсувати дефіцит ретроспективних даних ефективним використанням додаткової експертної інформації.

Суттєвою перешкодою в застосуванні глибокого навчання у критичних галузях є відсутність чи недостатня інтерпретованість отриманих за його допомогою результатів. Згідно з дослідженням «Пояснюваний штучний інтелект (ХАІ) для даних часових рядів: огляд» Т. Роя та ін.: більшість сучасних методів прогнозування є занадто складними для інтерпретації людиною, що знижує довіру до результатів та унеможливує перевірку рішень у нештатних ситуаціях [3]. Суто числові методи часто виявляються неспроможними врахувати та пояснити вплив неочевидних факторів, які не відображені у числових метриках, але можуть бути зафіксовані у текстових джерелах.

Окрім проблеми інтерпретованості, фундаментальною перешкодою є низька якість ретроспективних числових даних. На практиці дані часто містять пропуски, інформаційний шум, аномалії, спотворення, що робить їх невідповідними для побудови надійних прогностичних моделей. Водночас згідно з дослідженнями в галузі аналітики великих даних, близько 95% є неструктурованими. Серед основних викликів прогнозування є гетерогенність даних, велика кількість нерелевантних даних, а також хибні кореляції [4].

Вирішення описаного вище протиріччя лежить у площині переходу до гібридних нейро-символьних систем. Як стверджують Гарсез і Ламб у своїй праці про третю хвилю нейро-символьного штучного інтелекту, саме поєднання статистичного навчання та логічних міркувань дозволяє створювати системи, що відповідають високим вимогам довіри та безпеки [5]. Актуальною науковою задачею є розробка програмного гібридного методу, який би використовував оцінку прогностичної невизначеності нейромережі як тригер для динамічного залучення текстових експертних даних, підтримуючи баланс між точністю прогнозу та його інтерпретованістю.

Формулювання цілей статті

Метою статті є математична формалізація гібридного нейро-символьного методу програмного прогнозування, який дозволяє динамічно уточнювати результати базової моделі прогнозування на основі неструктурованої текстової експертної інформації в умовах невідповідності вхідних даних.

Виклад основного матеріалу

Для досягнення поставленої мети розробляється гібридний метод M_H який поєднує класичні методи числового прогнозування з механізмом динамічного уточнення на основі експертної текстової інформації. Ключова гіпотеза полягає в тому, що використання неструктурованої експертної текстової інформації у випадках високої невизначеності базової прогностичної моделі дозволяє компенсувати брак чи недостатню якість ретроспективних числових даних, підвищити точність.

Припускається, що простір вхідних даних S є неповною репрезентацією стану об'єкта, тобто існують приховані, неочевидні фактори h_{latent} , які не відображені у числових рядах, але впливають на істинне значення цільової змінної $y_T \in R$ у момент часу T . Вважається, що множина експертних документів K містить неструктуровані текстові описи, які корелюють з h_{latent} .

У межах даного дослідження під гібридним методом розуміється загальний алгоритмічний підхід в прогнозуванні, а моделлю виступає його конкретна параметрична реалізація.

Нехай:

- $X = R^n$ – n -вимірний векторний простір числових ознак вхідних даних, де $n \in N$ – кількість ознак (наприклад, показники датчиків).
- $x_t \in X$ – вектор ознак, що спостерігається в момент часу $t \in N$.
- $w \in N$ – фіксована довжина вхідної послідовності (розмір часового вікна).
- $S = X^w$ – векторний простір усіх можливих вхідних числових послідовностей.

- $s_T \in S$ – це конкретна вхідна послідовність в поточний момент часу T .
- $Y = R$ – векторний простір цільових значень (прогнозів).
- $k \in N$ – фіксований горизонт прогнозування.
- $H = \{(s_t, y_{t+k})\}_{t=w}^{T_{max}-k}$ – скінченна множина ретроспективних пар «вхід-вихід» ($H \subset S \times Y$), що використовується для навчання та валідації.
- $K = \{d_1, d_2, \dots, d_m\} = \{d_i\}_{i=1}^m$ – скінченна множина з m експертних текстових документів, де d_i це i -й конкретний документ (напр., звіт про інциденти, технічний регламент тощо).
- J – множина всіх можливих послідовностей слів, що формують текстове обґрунтування результату прогнозування.

Запропонований гібридний метод M_H складається з двох основних етапів: підготовчого, який виконується перед початком прогнозування, та операційного, який виконується для кожного нового прогнозу. Підготовчий етап являє собою формування бази експертних знань та навчання параметричних складових гібридної моделі M_H .

1. Підготовчий етап.

Формування бази експертних знань.

З метою забезпечення семантично коректного пошуку, початкова множина неструктурованих документів K має бути оброблена та перетворена на індексовану структуру K' , а саме базу знань, що створює зв'язок між числовими та текстовими даними. Нехай \mathcal{T} – це простір усіх можливих текстів, $E_{text}: \mathcal{T} \rightarrow R^d$ – попередньо навчена модель мовних векторів, що відображає текстовий документ у d -вимірний векторний простір R^d (де $d \in N$ – розмірність вектора). Процес підготовки бази знань полягає у перетворенні множини необроблених документів $K = \{d_i\}_{i=1}^m$ у індексовану K' , яка є набором пар «вектор-документ»:

$$K' = \{(v_i, d_i) \mid v_i = E_{text}(d_i), v_i \in R^d, d_i \in K\}, \quad (1)$$

де v_i – отримане векторне представлення документа d_i . У рамках даного методу під документом d_i розуміється атомарна змістовна одиниця тексту, довжина якої не перевищує обмеження вікна контексту мовної моделі

Навчання гібридної моделі.

Перш ніж гібридна модель зможе здійснювати прогнозування на нових даних, вона проходить етап навчання, у межах якого відбувається налаштування всіх її параметричних функцій. Нехай запропонований метод M_H складається з наступних структурних елементів: базової моделі прогнозування M_b , механізму адаптивного шлюзування M_g , генератора запитів M_q , механізму скалярної проєкції M_{fact} та генератора текстових обґрунтувань M_{exp} (детальний опис кожного з них наведено далі), тоді $\theta_b, \theta_g, \theta_q, \theta_{fact}, \theta_{exp}$ – їхні власні набори параметрів, а θ є повним набором параметрів конкретної прогностичної моделі, що реалізує метод M_H . $\theta = (\theta_b, \theta_g, \theta_q, \theta_{fact}, \theta_{exp})$. Процес навчання полягає у знаходженні оптимального набору параметрів θ^* , що мінімізує функцію втрат \mathcal{L} на навчальній множині H :

$$\theta^* = \arg \min_{\theta} \sum_{t=w}^{T_{max}-k} \mathcal{L}(M_H(s_t; \theta), y_{t+k}). \quad (2)$$

Функція втрат \mathcal{L} вимірює розбіжність між згенерованими кінцевими прогнозами та істинними значеннями цільової змінної. Тут \mathcal{L} розглядається як узагальнений критерій якості функціонування запропонованого гібридного методу. Він залежить від результатів проміжних обчислень, зокрема, від активації механізму шлюзування на основі прогностичної невизначеності.

2. Операційний етап.

Операційний етап відповідає за генерування результату прогнозування для вхідної послідовності даних s_T , що є парою $(\widehat{y}_T^*, j_T) \in Y \times J$, де \widehat{y}_T^* – кінцевий прогноз, а j_T – його обґрунтування. Під час операційного етапу запропонований гібридна модель послідовно здійснює базове прогнозування та оцінку прогностичної невизначеності, активує механізм адаптивного шлюзування, генератор запитів, виконує семантичний пошук в базі знань та обчислення величини скалярного уточнення, генерує текстове обґрунтування та зрештою об'єднує ці всі дані в кінцевий прогноз з відповідним обґрунтуванням. Опис кожного з цих кроків запропонованого методу викладено нижче.

Базове прогнозування: числові вхідні дані

Першим кроком запропонованого методу є базове прогнозування, під час якого базова модель генерує початковий прогноз, спираючись виключно на числові дані, паралельно обчислюючи рівень власної прогностичної невизначеності. Нехай $N_S \in N$ – фіксована кількість ітерацій стохастичних випробувань (гіперпараметр моделі). Позначимо через \widehat{y}_{T_i} результат i -го проходу M_b для вхідної послідовності s_T :

$$\widehat{y}_{T_i} = M_b(s_T; \theta_b, \mu_i), \quad \text{для } i = 1, \dots, N_S, \quad (3)$$

де μ_i – це випадковий бінарний вектор, що визначає підмножину нейронів, які будуть тимчасово виключені на i -й ітерації. Оскільки вектори є стохастично незалежними для кожної реалізації, обчислені значення прогнозів \widehat{y}_{T_i} будуть дещо відрізнятися між собою. Початковий прогноз базової моделі \widehat{y}_T у момент часу $T + k$ розраховується як просте середнє арифметичне отриманих значень:

$$\widehat{y}_T = \frac{1}{N_S} \sum_{i=1}^{N_S} \widehat{y}_{T_i}. \quad (4)$$

Оцінка прогностичної невизначеності

Оцінка прогностичної невизначеності c_T в момент часу T визначається як дисперсія цих прогнозів, а $C = [0,1] \subset R$ є множиною допустимих значень оцінок невизначеності. Якщо всі N_S випробувань генерують щільний кластер подібних результатів, дисперсія прагне до нуля, що свідчить про низьку прогностичну невизначеність. Натомість значне розсіювання отриманих значень вказує на високий рівень невизначеності та сигналізує про дефіцит релевантної інформації в історичних числових даних. Маємо наступну формулу для визначення прогностичної невизначеності:

$$c_T = \frac{1}{N_S} \sum_{i=1}^{N_S} (\widehat{y}_{T_i} - \widehat{y}_T)^2. \tag{5}$$

Механізм адаптивного шлюзування

Замість використання жорсткого порогу для використання бази знань вводиться механізм адаптивного шлюзування експертної інформації (M_g). Він визначає ступінь необхідності використання додаткових текстових даних через обчислення скалярної ваги $z_T \in (0, 1)$, фактично ймовірність того, що результату прогнозування базової моделі необхідне уточнення [6]. Такий підхід забезпечує динамічне балансування: при низькій невизначеності $z_T \rightarrow 0$ остаточний результат формується виключно на основі прогнозу базової моделі M_b , а при високій невизначеності $z_T \rightarrow 1$ відбувається уточнення результату за допомогою бази знань.

Нехай $M_g(\cdot; \theta_g) : S \times C \rightarrow (0,1)$ – параметрична функція шлюзування, яка аналізує вектор прихованого стану базової моделі та рівень невизначеності прогнозу, генеруючи ймовірність уточнення $z_T \in (0,1)$, тоді:

$$z_T = \sigma(W_g \cdot [h_T, c_T] + b_g), \tag{6}$$

де $\sigma(\cdot)$ – логістична функція (сигмоїда), $[h_T, c_T] \in R^{d_{h+1}}$ – вектор прихованого стану базової моделі ($h_T \in R^{d_h}$) та скалярної оцінки прогностичної невизначеності c_T . $\theta_g = \{W_g, b_g\}$ – навчальні параметри шлюзу, а саме: $W_g \in R^{1 \times (d_h+1)}$ – матриця ваг шлюзу, що проектує розширений вектор стану в скалярний простір, $b_g \in R$ – скалярний зсув [7].

Генератор запитів

Генератор запитів (M_q) генерує запити до бази знань K' , базуючись на поточному числовому стані s_T , який спричинив високу прогностичну невизначеність. Оскільки вектор прихованого стану базової моделі h_T (числа) та вектори документів v_i (текст) належать до різних векторних просторів, для їх узгодження використовується шар лінійної проєкції. Нехай $h_T \in R^{d_h}$ – це вектор прихованого стану базової моделі в момент часу T , тоді:

$$q_T = M_q(h_T; \theta_q) = W_q h_T + b_q, \quad q_T \in R^d, \tag{7}$$

де $\theta_q = \{W_q \in R^{d \times d_h}, b_q \in R^d\}$ – параметри генератора, що оптимізуються, а саме матриця ваг проєкції, що відображає прихований стан моделі у векторний простір документів та вектор зсуву проєкції відповідно. Мета навчання θ_q полягає в тому, щоб вектор q_T опинився геометрично близьким до векторів v_i тих документів, які є релевантними для поточної ситуації прогнозування [8]. Фактично, згенерований векторний запит q_T є представленням «інформаційної потреби» алгоритму: він перетворює поточний «проблемний» числовий стан моделі h_T у координати семантичного простору, вказуючи напрямок для пошуку текстової інформації.

Механізм семантичного пошуку

Механізм семантичного пошуку (M_r) забезпечує цільовий пошук необхідної експертної інформації з попередньо індексованої бази знань K' . Приймаючи на вхід векторний запит q_T , функція виконує пошук документів із найвищим ступенем векторної подібності в просторі бази знань K' . Результатом пошуку є формування локального кластера документів D_T , зміст яких є максимально релевантним до поточного стану цільового показника [9]. Формально цей процес описується оператором відбору k_n документів індексованої бази знань, векторні представлення яких мають найвищий ступінь подібності до запиту:

$$D_T = M_r(q_T, K') = \text{Top}_{k_n}(\{(d_i, v_i) \in K'\}; \text{sim}(q_T, v_i)), \tag{8}$$

де $\text{Top}_{k_n}(\cdot)$ – це функція вибору, яка повертає підмножину з k пар «вектор-документ», що мають найвищу подібність [6], $\text{sim}(a, b) = \frac{a \cdot b}{|a| |b|}$, v_i – векторне представлення документа d_i . Отримана підмножина $D_T \subset K$ формує множину релевантних експертних даних для поточного прогнозу.

Отримана підмножина документів D_T становить локальний масив релевантних експертних знань, проте різні документи мають різну ступінь користі для конкретного прогнозу. Тому знайдена підмножина потребує ранжування та агрегації для формування єдиного вектора текстових ознак, який акумулює найважливішу інформацію та фільтрує можливий інформаційний шум.

Для агрегації знайденої інформації в межах M_r застосовується механізм крос-модальної уваги (attention mechanism), який обчислює внесок кожного документа відповідно до його релевантності запиту q_T . Нехай $V_D \in R^{k_n \times d}$ – матриця, рядками якої є вектори знайдених документів v_i . Оцінка ступеня релевантності здійснюється через обчислення вектора коефіцієнтів уваги $a_T \in [0,1]^{k_n}$ як нормованого скалярного добутку із застосуванням функції softmax σ_{sm} :

$$a_T = \sigma_{sm} \left(\frac{V_D q_T}{\sqrt{d}} \right), \tag{9}$$

де \sqrt{d} – масштабуючий коефіцієнт, необхідний для стабілізації градієнтів при великих значеннях розмірності d

[10]. Агрегований вектор контексту $c_{ctx} \in R^d$ визначається як зважена сума векторів документів [11]:

$$c_{ctx} = V_D^T a_T = \sum_{i=1}^{k_n} a_{T,i} v_i. \quad (10)$$

Механізм скалярної проєкції

Перетворення вектора контексту c_{ctx} у числовий простір виконує механізм скалярної проєкції M_{fact} . Ця функція, побудована за принципом багатощарового перцептрона [7], обчислює скалярне уточнення $r_T \in R$:

$$r_T = M_{fact}(c_{ctx}; \theta_{fact}) = w_f^T \cdot \phi(W_p c_{ctx} + b_p) + b_f, \quad (11)$$

де $\theta_{fact} = \{W_p, b_p, w_f, b_f\}$ – параметри, що оптимізуються під час мінімізації загальної похибки прогнозування. $W_p \in R^{d_h \times d}$ – матриця ваг проєкції вхідного вектора розмірності d у прихований простір розмірності d_h , $b_p \in R^{d_h}$ – вектор зсуву прихованого шару, $\phi(\cdot)$ – функція активації ReLU, де $\phi(z) = \max(0, z)$, $w_f \in R^{d_h}$ – вектор ваг вихідного шару, що проєктує прихований стан у скалярний простір, $b_f \in R$ – скалярний зсув вихідного шару. Значення r_T відображає отриману величину відхилення, зумовлену факторами, які формалізовані у векторі контексту.

Генератор текстових обґрунтувань

Інтерпретованість результатів забезпечується генератором текстових обґрунтувань (M_{exp}), що відображає векторні представлення запиту q_T та контексту c_{ctx} у простір текстових пояснень J . Приймаючи на вхід числові репрезентації нашого запиту та знайдених експертних знань, алгоритм трансформує їх у лінгвістичну форму, зрозумілу для кінцевого користувача. Дана функція формалізується наступним чином:

$$j_T = M_{exp}(q_T, c_{ctx}; \theta_{exp}), \quad (12)$$

де θ_{exp} – набір внутрішніх параметрів генератора, що підлягають оптимізації. Результатом є текстове обґрунтування j_T , яке описує логічний зв'язок між знайденою текстовою інформацією та величиною уточнення r_T .

Механізм адаптивного злиття

Механізм адаптивного злиття (M_f) є функцією, що об'єднує результат базової моделі (пара (\widehat{y}_T, c_T)) та набір текстових документів для формування кінцевого результату прогнозування. Як механізм злиття використовується зважена лінійна комбінація, що динамічно регулюється попередньо обчисленим коефіцієнтом шлюзування z_T .

Визначимо прогноз на основі бази знань y_T^{kn} як суму базового прогнозу та обчисленого уточнення:

$$y_T^{kn} = \widehat{y}_T + r_T, \quad (13)$$

тоді кінцевий прогноз \widehat{y}_T^* обчислюється за формулою:

$$\widehat{y}_T^* = M_f(\widehat{y}_T, r_T, z_T) = (1 - z_T) \cdot \widehat{y}_T + z_T \cdot y_T^{kn} = \widehat{y}_T + z_T \cdot r_T. \quad (14)$$

Зрештою, загальний результат роботи запропонованого гібридного методу M_H на операційному етапі формується як кортеж із двох значень – підсумкового числового прогнозу та його текстового обґрунтування:

$$M_H(s_T, \theta^*) = (\widehat{y}_T^*, j_T). \quad (15)$$

Отриманий результат \widehat{y}_T^* є кінцевим прогнозованим значенням гібридної моделі в момент часу T , що поєднує числові закономірності часового ряду з неструктурованою текстовою експертною інформацією, а згенероване текстове обґрунтування j_T забезпечує прозорість результатів для кінцевого користувача.

Відповідно до концепції запропонованого підходу, гібридний метод має звертатися до механізму семантичного пошуку M_r та формувати текстові обґрунтування за допомогою M_{exp} виключно в умовах інформаційного дефіциту, адже постійне залучення текстової інформації без вагомих підстав може призвести до збільшення інформаційного шуму та мати негативний ефект на результат прогнозування. Виходячи з цього, цільова функція втрат \mathcal{L} , введена у постановці задачі (2), є зваженою сумою помилки прогнозування та штрафу за активацію механізму шлюзування:

$$\mathcal{L}(\theta; s_T, \widehat{y}_T) = \mathcal{L}_{MSE}(\widehat{y}_T^*, y_T) + \lambda \cdot z_T, \quad (16)$$

де \mathcal{L}_{MSE} – середньоквадратична похибка остаточного прогнозу відносно істинного значення y_T , а $\lambda > 0$ – гіперпараметр регуляризації, що керує компромісом між точністю та частотою звернень до бази знань.

Завдяки оптимізації параметрів з урахуванням штрафу $\lambda \cdot z_T$, модель навчається мінімізувати вихід активації шлюзу z_T у випадках, коли базового числового прогнозу достатньо. Це дозволяє реалізувати на операційному етапі стратегію ранньої зупинки шляхом введення жорсткого порогу відсікання $\tau \in (0, 1)$. Якщо обчислене значення шлюзу $z_T < \tau$, тоді модель ігнорує більш комплексний семантичний пошук знань, приймаючи $\widehat{y}_T^* = \widehat{y}_T$ [12]. Такий підхід дозволяє уникати використання операцій пошуку в ситуаціях, коли ступінь прогностичної невизначеності базової моделі є низьким. Значення гіперпараметра τ встановлюється емпірично залежно від вимог конкретної прикладної галузі.

На рис. 1 зображено діаграму послідовностей роботи операційного етапу запропонованого гібридного методу.

Для оцінки прогностичної здатності запропонованого гібридного методу M_H у рамках даного дослідження було проведено модельний експеримент на даних часового ряду з високою волатильністю (курс біткоїну – BTC-USD) за період з 2017 по 2023 рік. Вибірку було розділено на навчальну (до 01.05.2021) та тестову

множини, що дозволило перевірити здатність моделі прогнозувати невідомі події, наприклад, регуляторні обмеження (заборона криптовалют в Китаї, травень 2021) та кризи ліквідності (листопад 2022).

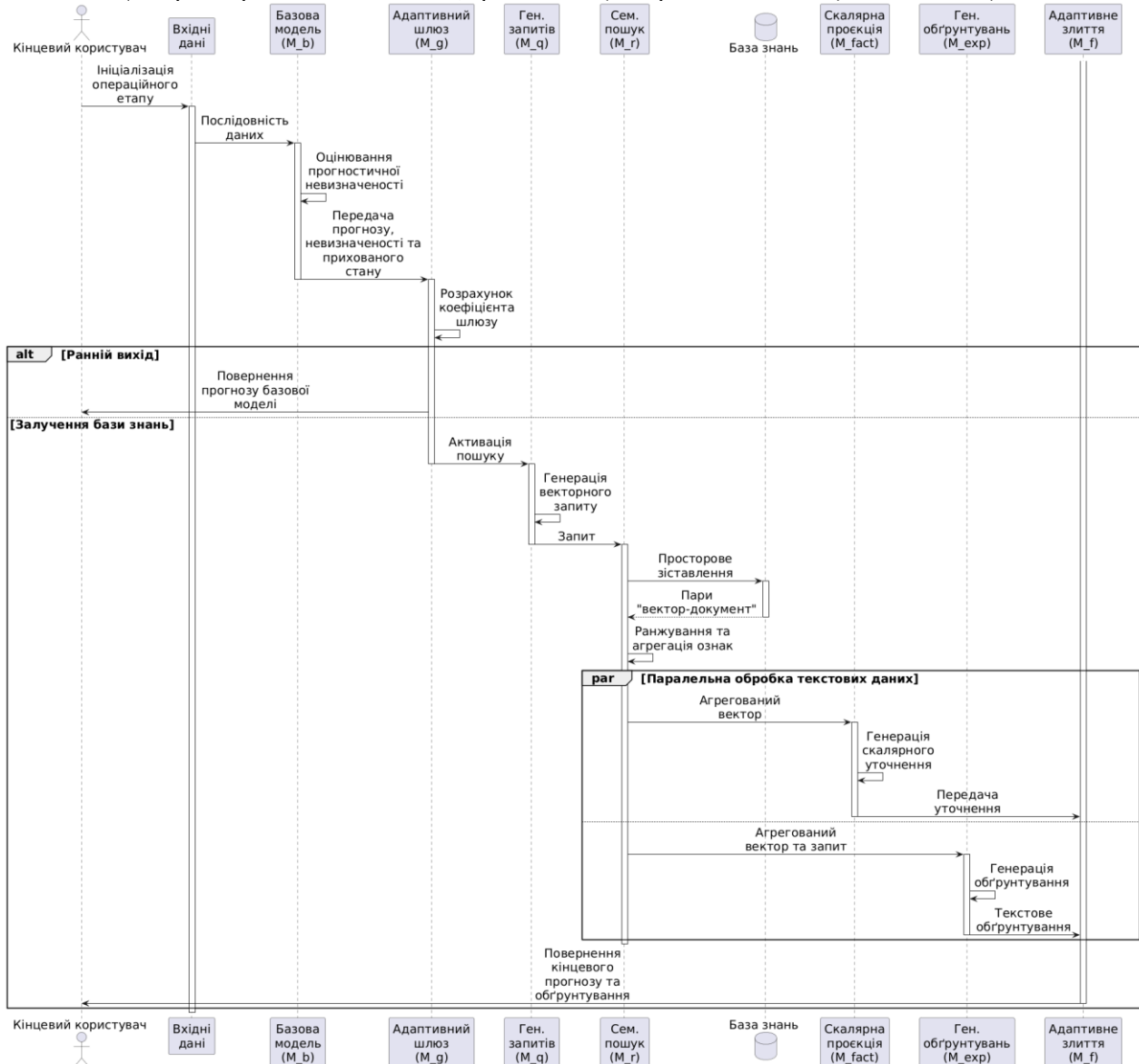


Рис. 1. Діаграма послідовностей роботи операційного етапу гібридного методу

У ході порівняльного аналізу базову модель M_b (Bi-LSTM) було протиставлено гібридному підходу, що реалізує механізм текстово-орієнтованого уточнення на основі текстової експертної інформації. Результати моделювання показали, що базова нейромережа, покладаючись виключно на ретроспективні числові патерни, демонструє ефект інерційності: вона успішно відтворює тренд, але із запізненням реагує на зміни, спричинені зовнішніми факторами. Коли невизначеність моделі перевищувала поріг відсікання τ , моделювалося уточнення прогнозу на основі текстових даних.

Оцінка якості прогнозування здійснювалася локально – для множини часових точок, які було визначено як аномальні. Ключовим показником було обрано середню абсолютну помилку у відсотках (MAPE). Результати показали, що базова модель прогнозування у цих точках показала низьку точність із середньою помилкою $MAPE_{base} \approx 25.66\%$. Застосування гібридного методу відбувалося за умови часткового відновлення бракуючої інформації на рівні 75%, але вже з таким показником середньої якості відновлення інформації вдалося знизити показник помилки до $MAPE_{hybrid} \approx 6.41\%$. Таке зменшення помилки підтверджує гіпотезу, що залучення текстових експертних даних дозволяє компенсувати похибку прогнозу в ті моменти, коли класичні методи виявляються неефективними.

Висновки

У статті викладено математичну формалізацію гібридного нейро-символьного методу програмного прогнозування M_H . На відміну від класичних підходів, запропонований метод вирішує проблему невідповідності ретроспективних даних при програмному прогнозуванні шляхом динамічного уточнення базового числового прогнозу, використовуючи неструктуровану текстову експертну інформацію. Це дозволяє компенсувати

інформаційну недостатність числових моделей та забезпечує точність прогнозування в умовах високої невизначеності.

Важливим результатом дослідження є математичний опис механізму адаптивного шлюзування, який визначає міру взаємодії між числовим прогнозом та неструктурованими текстовими даними. Введення порогового критерію τ забезпечує здійснення семантичного пошуку виключно в ті моменти, коли базова модель може помилятися через брак релевантних даних. Крім того, було формалізовано процеси перетворення агрегованого вектора контексту на скалярне уточнення базового числового прогнозу та паралельну генерацію текстового обґрунтування, що забезпечує прозорість та інтерпретованість кінцевого результату.

У даному дослідженні сформовано теоретичний базис запропонованого методу. Це створює необхідне підґрунтя для наступного етапу дослідження – програмної реалізації прогнозової системи.

Література

1. Beyond Trend and Periodicity: Guiding Time Series Forecasting with Textual Cues / J. Wang et al. *arXiv preprint*. 2024. URL: <https://arxiv.org/abs/2405.13522>.
2. Hewamalage H., Bergmeir C., Bandara K. Recurrent neural networks for time series forecasting: Current status and future directions. *International Journal of Forecasting*. 2021. Vol. 37, no. 1. P. 388–427. URL: <https://doi.org/10.1016/j.ijforecast.2020.06.008>.
3. Explainable Artificial Intelligence (XAI) on Time Series Data: A Survey / T. Rojat et al. *arXiv preprint*. 2021. URL: <https://doi.org/10.48550/arXiv.2104.00950>.
4. Gandomi A., Haider M. Beyond the hype: *Big data concepts, methods, and analytics*. *International Journal of Information Management*. 2015. Vol. 35, No. 2. P. 137–144. URL: <https://doi.org/10.1016/j.ijinfomgt.2014.10.007>.
5. Garcez A. D., Lamb L. C. Neurosymbolic AI: the 3rd wave. *Artificial Intelligence Review*. 2023. Vol. 56. P. 12387–12406. URL: <https://doi.org/10.1007/s10462-023-10448-w>.
6. Outrageously large neural networks: The sparsely-gated mixture-of-experts layer / N. Shazeer et al. *5th International Conference on Learning Representations (ICLR 2017)*. Toulon, France, 2017. URL: <https://doi.org/10.48550/arXiv.1701.06538>.
7. Goodfellow I., Bengio Y., Courville A. *Deep learning*. MIT Press, 2016.
8. Dense Passage Retrieval for Open-Domain Question Answering / V. Karpukhin et al. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 2020. P. 6769–6781. URL: <https://doi.org/10.18653/v1/2020.emnlp-main.550>.
9. Retrieval-augmented generation for knowledge-intensive nlp tasks / P. Lewis et al. *Advances in Neural Information Processing Systems 33 (NeurIPS 2020)*. 2020. Vol. 33. P. 9459–9474. URL: <https://doi.org/10.48550/arXiv.2005.11401>.
10. Attention is all you need / A. Vaswani et al. *Advances in Neural Information Processing Systems 30 (NIPS 2017)*. Long Beach, CA, USA, 2017. P. 6000–6010. URL: <https://doi.org/10.48550/arXiv.1706.03762>.
11. Neural machine translation by jointly learning to align and translate / D. Bahdanau, K. Cho, Y. Bengio. *3rd International Conference on Learning Representations (ICLR 2015)*. San Diego, CA, USA. 2015. URL: <https://arxiv.org/abs/1409.0473>.
12. Teerapittayanon S., McDanel B., Kung H. T. BranchyNet: Fast inference via early exits from deep neural networks. *23rd International Conference on Pattern Recognition (ICPR)*. Cancun, Mexico, 2016. URL: <https://doi.org/10.48550/arXiv.1709.01686>.

References

1. Beyond Trend and Periodicity: Guiding Time Series Forecasting with Textual Cues / J. Wang et al. *arXiv preprint*. 2024. URL: <https://arxiv.org/abs/2405.13522>.
2. Hewamalage H., Bergmeir C., Bandara K. Recurrent neural networks for time series forecasting: Current status and future directions. *International Journal of Forecasting*. 2021. Vol. 37, no. 1. P. 388–427. URL: <https://doi.org/10.1016/j.ijforecast.2020.06.008>.
3. Explainable Artificial Intelligence (XAI) on Time Series Data: A Survey / T. Rojat et al. *arXiv preprint*. 2021. URL: <https://doi.org/10.48550/arXiv.2104.00950>.
4. Gandomi A., Haider M. Beyond the hype: *Big data concepts, methods, and analytics*. *International Journal of Information Management*. 2015. Vol. 35, No. 2. P. 137–144. URL: <https://doi.org/10.1016/j.ijinfomgt.2014.10.007>.
5. Garcez A. D., Lamb L. C. Neurosymbolic AI: the 3rd wave. *Artificial Intelligence Review*. 2023. Vol. 56. P. 12387–12406. URL: <https://doi.org/10.1007/s10462-023-10448-w>.
6. Outrageously large neural networks: The sparsely-gated mixture-of-experts layer / N. Shazeer et al. *5th International Conference on Learning Representations (ICLR 2017)*. Toulon, France, 2017. URL: <https://doi.org/10.48550/arXiv.1701.06538>.
7. Goodfellow I., Bengio Y., Courville A. *Deep learning*. MIT Press, 2016.
8. Dense Passage Retrieval for Open-Domain Question Answering / V. Karpukhin et al. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 2020. P. 6769–6781. URL: <https://doi.org/10.18653/v1/2020.emnlp-main.550>.
9. Retrieval-augmented generation for knowledge-intensive nlp tasks / P. Lewis et al. *Advances in Neural Information Processing Systems 33 (NeurIPS 2020)*. 2020. Vol. 33. P. 9459–9474. URL: <https://doi.org/10.48550/arXiv.2005.11401>.
10. Attention is all you need / A. Vaswani et al. *Advances in Neural Information Processing Systems 30 (NIPS 2017)*. Long Beach, CA, USA, 2017. P. 6000–6010. URL: <https://doi.org/10.48550/arXiv.1706.03762>.
11. Neural machine translation by jointly learning to align and translate / D. Bahdanau, K. Cho, Y. Bengio. *3rd International Conference on Learning Representations (ICLR 2015)*. San Diego, CA, USA. 2015. URL: <https://arxiv.org/abs/1409.0473>.
12. Teerapittayanon S., McDanel B., Kung H. T. BranchyNet: Fast inference via early exits from deep neural networks. *23rd International Conference on Pattern Recognition (ICPR)*. Cancun, Mexico, 2016. URL: <https://doi.org/10.48550/arXiv.1709.01686>.