

ВЕРИФІКАЦІЯ ЗНАННЯ-ОРІЄНТОВАНОЇ ТЕХНОЛОГІЇ ЗМЕНШЕННЯ НЕВИЗНАЧЕНОСТІ ТА ПІДВИЩЕННЯ УЗГОДЖЕНОСТІ ПРОГНОЗІВ У СЦЕНАРНОМУ ПРОГНОЗУВАННІ ЧАСОВИХ РЯДІВ

Вінницький національний технічний університет

Анотація

У роботі подано подальшу верифікацію знання-орієнтованої ієрархічної мультиагентної інтелектуальної системи сценарного прогнозування часових рядів на основі великих мовних моделей. Для дослідження використано часовий ряд добової кількості випадків COVID-19 в Україні за 2020-2023 роки. На основі локальних ознак динаміки сформовано сценарний набір історичних вікон, що охоплює режими зростання перед піком, спаду після піку, відновлення після спаду та аномальної динаміки. У межах цього набору проведено порівняння спрощеної архітектури системи з повноцінною сценарною ансамблевою архітектурою. Показано, що повноцінна ансамблева архітектура забезпечує більш узгоджені результати: середня WAPE зменшується на 15,7 %, середня ширина прогнозних інтервалів зменшується на 13,7 %, а значення інтегрального критерію J підвищується на 22,4 % порівняно зі спрощеною. Отримані результати підтверджують доцільність подальшого розвитку сценарного ансамблювання та засобів оцінювання невизначеності в межах IMAIC.

Ключові слова: часові ряди, сценарне прогнозування, великі мовні моделі, мультиагентна система, невизначеність, ансамблювання, COVID-19.

Abstract

The paper presents a further verification of the knowledge-oriented hierarchical multi-agent intelligent system for scenario-based time series forecasting built on large language models. The study uses the time series of daily COVID-19 cases in Ukraine for the period 2020-2023. Based on local dynamic features, a scenario-oriented set of historical windows was formed, covering growth before a peak, decline after a peak, recovery after decline, and anomalous dynamics. Within this set, a simplified system architecture was compared with a full scenario-based ensemble architecture. The results show that the full scenario-based ensemble architecture provides more consistent forecasting performance: mean WAPE decreases by 15.7%, mean prediction interval width decreases by 13.7%, and the value of the integral criterion J increases by 22.4% compared with the simplified architecture. These results confirm the feasibility of further developing scenario-based ensemble forecasting and uncertainty estimation mechanisms within IMAIS.

Keywords: time series, scenario forecasting, large language models, multi-agent system, uncertainty, ensemble forecasting, COVID-19.

Вступ

Аналіз сучасних публікацій показує, що LLM-орієнтовані підходи до прогнозування часових рядів розвиваються досить інтенсивно, однак отримані результати суттєво залежать від постановки задачі, способу подання даних і схеми інтеграції мовної моделі з часовими ознаками [1-3]. Саме тому додаткова верифікація сценарно-орієнтованих архітектур на історичних фрагментах ряду є актуальною. У попередніх роботах авторами було запропоновано технологію декомпозиції та ансамблювання моделей для зменшення прогнозної невизначеності [4] і мультиагентну архітектуру IMAIC для сценарного прогнозування часових

рядів на основі LLM [5]. Для подальшої верифікації такої технології необхідне оцінювання на багатьох історичних ділянках, що відображають різні режими динаміки рядів.

Метою даної роботи є порівняння спрощеної архітектури A1 та повноцінної сценарної архітектури A3 на множині історичних ділянок з різними режимами динаміки часового ряду. Такий підхід дозволяє перейти до систематичної перевірки переваг ІМАІС.

Результати дослідження

Для дослідження використано часовий ряд добової кількості випадків COVID-19 в Україні за 2020-2023 роки. Після попередньої обробки даних було сформовано набір усіх можливих вікон, що містять 28 днів історії та 14 днів майбутнього для прогнозування. Для кожного вікна обчислювалися ознаки локальної динаміки, після чого кожному ділянці було віднесено до одного з таких сценаріїв: “зростання перед піком”, “спадання після піку”, “відновлення після спаду” та “аномальна динаміка”. На основі цієї класифікації було сформовано сценарно-орієнтований бенчмарк, призначений для подальшої багатовіконної верифікації архітектур A1 та A3. До нього увійшло 20 вікон, по 5 для кожного типу.

Подальше порівняння виконувалося у термінах формалізації ІМАІС, запропонованої у попередній статті. Множину всіх можливих архітектур подано як $\Omega = \Omega_1 \cup \Omega_2 \cup \Omega_3$, де Ω_1 відповідає агентам підготовки даних і формування ознак, Ω_2 — агентам побудови прогнозних моделей та альтернативних сценаріїв, а Ω_3 — агентам кооперації прогнозів і оцінювання невизначеності. Щоб уникнути конфлікту позначень із назвами архітектур A1 та A3, далі конкретну реалізацію ІМАІС описуватимемо через три рівні: L_1 , L_2 та L_3 , причому $L_i \in \Omega_i$, $i = 1, 2, 3$. У цій постановці рівень L_1 формує датасет $D = \{(x_p, y_p)\}$ на основі очищення та передоброблення даних, рівень L_2 формує множину сценаріїв $S = \{s^{(k)}\}_{k=1}^K$ з горизонтом H , рівень L_3 забезпечує кооперацію прогнозів, ансамблювання та оцінювання невизначеності.

Якість роботи архітектури оцінювалась за інтегральним критерієм $J = \sum_{i=1}^3 w_i U_i$, де $\sum_{i=1}^3 w_i = 1$, $w_i > 0$. Тут U_1 характеризує якість підготовки даних, U_2 — адекватність відтворення динаміки та точність прогнозу, а U_3 — зменшення і каліброваність зони невизначеності [5].

За результатами експерименту A3 зменшує середню WAPE з 2.128 до 1.794, підвищує емпіричне покриття з 0.704 до 0.736 та збільшує середнє значення J з 0.514 до 0.629. Для критерію J отримано статистично значущий вигравш архітектури A3 ($p = 0.0005$), що свідчить про більш стійкий інтегральний ефект від сценарної кооперації агентів. На рисунку 1 наведено порівняння архітектур A1 та A3 за складовими інтегрального критерію якості U_1 , U_2 , U_3 та підсумковим значенням J .

Найбільший інтерес становлять сценарії спадання після піку та аномальної динаміки. Для таких сценаріїв A3 показує менше значення WAPE на 17.0 % і більше значення інтегрального критерію J на 15.7 %. На рисунку 2 наведено порівняння архітектур A1 та A3 для різних типів сценарних вікон.

Порівняння архітектур A1 та A3 за складовими інтегрального критерію

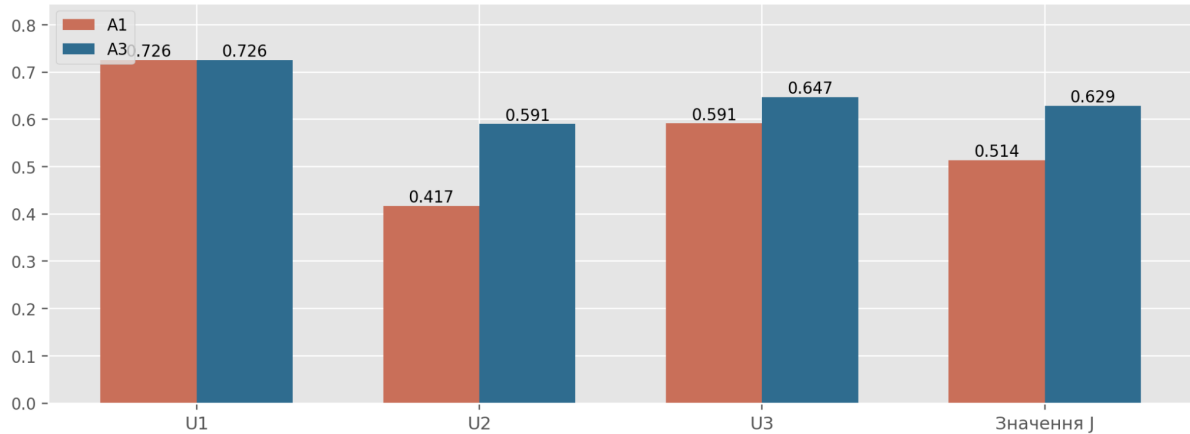


Рисунок 1. Порівняння архітектур A1 і A3 за значеннями складових інтегрального критерію

Порівняння A1 та A3 на різних типах сценарних вікон

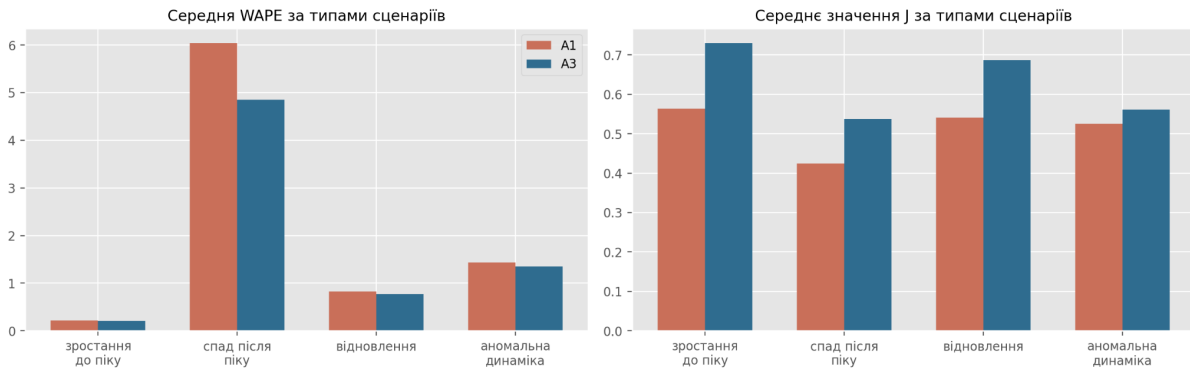


Рисунок 2. Порівняння ефективності архітектур A1 і A3 за типами сценаріїв

Висновки

Запропонований сценарно-орієнтований бенчмарк дозволив виконати подальшу верифікацію технології IMAIC на множині автоматично відібраних історичних ділянок одного часового ряду. Показано, що повноцінна архітектура A3 забезпечує кращий інтегральний результат, ніж спрощена архітектура A1: середнє значення WAPE більше на 15,7 %, а значення інтегрального критерію J більше на 22,4 %.

Подальші дослідження доцільно спрямувати на підключення зовнішніх текстових контекстів, розширення бібліотеки агентів та перевірку технології на інших нестационарних рядах.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Jin M. Time-LLM: Time Series Forecasting by Reprogramming Large Language Models [Електронний ресурс] / M. Jin, S. Wang, L. Ma et al. – 2024. – Режим доступу: <https://arxiv.org/abs/2310.01728>.
2. Wang X. From News to Forecast: Integrating Event Analysis in LLM-Based Time Series Forecasting with Reflection / X. Wang, M. Feng, J. Qiu, J. Gu, J. Zhao // Advances in Neural Information Processing Systems. – 2024. – Vol. 37. – P. 58118–58153. – DOI: 10.52202/079017-1853.
3. Pan Z. S²IP-LLM: Semantic Space Informed Prompt Learning with LLM for Time Series Forecasting / Z. Pan, Y. Jiang, S. Garg, A. Schneider, Y. Nevmyvaka, D. Song // Proceedings of the 41st International Conference on Machine Learning. – 2024. – Vol. 235. – P. 39135–39153. – (Proceedings of Machine Learning Research).

4. Варер Б. Ю. Технологія декомпозиції та ансамблювання моделей для зменшення невизначеності та підвищення узгодженості прогнозів на основі LLM / Б. Ю. Варер // Наукові праці Вінницького національного технічного університету. – 2025. – № 4. – С. 129–137. – DOI: 10.31649/2307-5376-2025-4-129-137.
5. Варер Б. Ю. Знання-орієнтована ієрархічна мультиагентна інтелектуальна система сценарного прогнозування часових рядів на основі LLM / Б. Ю. Варер, В. Б. Мокін // Вісник ВПІ. – 2026. – № 1.

Мокін Віталій Борисович — д-р техн. наук, професор, завідувач кафедри системного аналізу та інформаційних технологій; e-mail: ybmokin@vntu.edu.ua;

Варер Борис Юхимович — аспірант кафедри системного аналізу та інформаційних технологій; e-mail: androbor17@gmail.com

Mokin Vitalii B. — Dr. Sc. (Eng.), Professor, Head of the Chair of System Analysis and Information Technologies, e-mail: ybmokin@vntu.edu.ua;

Varer Borys Yu. — Post-graduate Student of the Chair of System Analysis and Information Technologies, e-mail: androbor17@gmail.com